ASSOCIATION FOR THE ADVANCEMENT OF ANIMAL BREEDING AND GENETICS

Proceedings of the Eighteenth Conference

Matching Genetics and the Environment a New Look at and Old Topic

Barossa Valley, S.A. Australia
28th September – 1st October 2009

ISSN 1328-3227
ISBN 978-0-646-52103-9

Produced by:
Association for the Advancement of Animal Breeding and Genetics
C/- AAABG Secretary
Animal Science
University of Adelaide, Roseworthy Campus
Roseworthy
South Australia , 5371
Australia

Internet web site: http://www.aaabg.org

Available from:
AAABG Distribution Service
c/- AGBU
The University of New England
Armidale, NSW 2351
Australia

PRESIDENT'S MESSAGE

Welcome to the 18th Conference of the Association for the Advancement of Animal Breeding and Genetics (AAABG). I am proud to introduce these proceedings and trust that you enjoy their rich depth of endeavour and the breadth of topics.

Our 18th Conference is a time for us to celebrate our 30th anniversary. Over the last 3 decades world population has grown from 4.4 to 6.8 billion , a growth rate of 1.5% p.a. and comes during the period after the large gains in crop production from the green revolution. Food production growth can no longer come from expansion, but must come from increasing efficiency.

Despite increases in demand, food production has grown and is also meeting the challenge of declining terms of trade that are inconsistent across products. For example, in the past 30 years in real terms, wool prices have dropped by roughly 30% and lamb prices have risen by roughly 30% leading to major structural change in our sheep industry. While slowed by the current global financial crisis, we are on the verge of a livestock revolution with expected large increases in demand for animal foods resulting from increased affluence in developing economies.

Genotype by environment interactions (GxE) have been of interest to animal breeders for as long as we have records. As production systems are pushed further, the GxE becomes increasingly important. We have some excellent papers on the topic herein and I am certain this will continue to be a focus of the next generation of breeders.

As we achieve a greater understanding of biology and inheritance, we learn of greater complexity and the wonders of the world we live in. We still cannot predict with certainty what an individual animal will look like or how they will perform. However, we do have good industry recording programs. Relatives do resemble each other and by selection of superior parents, genetic progress has been substantial, is ongoing and has potential to accelerate significantly.

We are in a very exciting period as we observe the transition of molecular tools from descriptors of individual gene variants to that of whole genomes. Indeed, genomic selection has increasingly become part of the animal geneticist's vocabulary. The challenge ahead is to use these tools in multidisciplinary approaches for increasing production efficiency. To that end, Cooperative Research Centres have had a profound impact on livestock research and postgraduate training. Benefit-cost analyses of their impact have always been favourable for both producers and consumers and there are many papers from CRCs scientists and students in these proceedings.

Our food and fibre producers manage the biology of the system overlaid by micro- and macro-economic forces, environmental conditions and changing public perception of ethics in relation to animal welfare. A major aim in establishing AAABG is to increase the dialogue between producers, breeders and geneticists. This is a major challenge for each Organising Committee and we have tried some new strategies such as providing discounts for all producers and students and added formal Breeders' Days.

I trust that you enjoy the program we have put together and look forward to the rich experience of discussing the application of research results with close friends and colleagues.

Wayne S. Pitchford
President

ASSOCIATION FOR THE ADVANCEMENT OF ANIMAL BREEDING AND GENETICS

EIGHTEENTH CONFERENCE

COMMITTEE

President	Wayne Pitchford
President-Elect	Johan Greeff
Vice President	Forbes Brien
Secretary	Cynthia Bottema
Treasurer	David Rutley
Editorial Committee	Alex Safari
	Bill Pattie
	Barrie Restall

Organising Committee
David Adelson
Cynthia Bottema
Forbes Brien
Carolyn Fitzsimmons
Stefan Hiendleder
Greg Nattrass
Bill Pattie
Wayne Pitchford
David Rutley
Bill Richardson
Alex Safari
Jim Walkley

Breeders' Day Committees
Lynton Arney
Shirley Barker
Forbes Brien
Malcolm Buckby
Jean Evans
John Fox
Joe Grieve
Ben Hebart
Wayne Pitchford
Graeme Pope
Bill Richardson
Ian Rowett
Jamie Withers

SPONSORS OF THE 30TH ANNIVERSARY CONFEENCE

The financial assistance of the following organisations is gratefully acknowledged.

GOLD SPONSOR

SILVER SPONSORS

BRONZE SPONSOR

AAABG was formerly known as the Australian Association for Animal Breeding and Genetics. Following the 1995 OGM the name was changed when it became an organisation with a joint Australian and New Zealand membership. The Association for the Advancement of Animal Breeding and Genetics is incorporated in South Australia.

THE ASSOCIATION FOR THE ADVANCEMENT OF ANIMAL BREEDING AND GENETICS INCORPORATED

OBJECTIVES

(i) to promote scientific research on the genetics of animals;
(ii) to foster the application of genetics in animal production;
(iii) to promote communication among all those interested in the application of genetics to animal production, particularly breeders and their organisations, consultants, extension workers, educators and geneticists.

To meet these objectives, the Association will:

(i) hold regular conferences to provide a forum for:
(a) presentation of papers and in-depth discussions of general and industry-specific topics concerning the application of genetics in commercial animal production;
(b) scientific discussions and presentation of papers on completed research and on proposed research projects;
(ii) publish the proceedings of each Regular Conference and circulate them to all financial members;
(iii) use any such other means as may from time to time be deemed appropriate.

MEMBERSHIP

Any person interested in the application of genetics to animal production may apply for membership of the Association and, at the discretion of the Committee, be admitted to membership as an Ordinary Member.

Any organisations interested in the application of genetics to animal production may apply for membership and, at the discretion of the Committee, be admitted to membership as a Corporate member. Each such Corporate Member shall have the privilege of being represented at any meeting of the Association by one delegate appointed by the Corporate Member.

Benefits to Individual Members

- While it is not possible to produce specific recommendations or "recipes" for breeding plans that are applicable for all herd/flock sizes and management systems, principles for the development of breeding plans can be specified. Discussion of these principles, consideration of particular case studies, and demonstration of breeding programs that are in use will all be of benefit to breeders.
- Geneticists will benefit from the continuing contact with other research workers in refreshing and updating their knowledge.
- The opportunity for contact and discussions between breeders and geneticists in individual members' programs, and for geneticists in allowing for detailed discussion and appreciation of the practical management factors that often restrict application of optimum breeding programs.

Benefits to Member Organisations

- Many of the benefits to individual breeders will also apply to breeding organisations. In addition, there are benefits to be gained through coordination and integration of their efforts. Recognition of this should follow from understanding of common problems, and would lead to increased effectiveness of action and initiatives.
- Corporate members can use the Association as a forum to float ideas aimed at improving and/or increasing service to their members.

General Benefits

- Membership of the Association may be expected to provide a variety of benefits and, through the members, indirect benefits to all the animal industries.
- All members should benefit through increased recognition of problems, both at the level of research and of application, and increased understanding of current approaches to their solution.
- Well-documented communication of gains to be realised through effective breeding programs will stimulate breeders and breeding organisations, allowing increased effectiveness of application and, consequently, increased efficiency of operation.
- Increased recognition of practical problems and specific areas of major concern to individual industries should lead to increased relevance of applied research.
- All breeders will benefit indirectly because of improved services offered by the organisations which service them.
- The existence of the Association will increase appreciably the amount and use of factual information in public relations in the animal industries.
- Association members will comprise a pool of expertise – at both the applied and research levels – and, as such, individual members and the Association itself must have an impact on administrators at all levels of the animal industries and on Government organisations, leading to wiser decisions on all aspects of livestock improvement, and increased efficiency of animal production.

CONFERENCES

One of the main activities of the Association is the Conference. These Conferences will be structured to provide a forum for discussion of research problems and for breeders to discuss their problems with each other, with extension specialists and with geneticists.

ASSOCIATION FOR THE ADVANCEMENT OF ANIMAL BREEDING AND GENETICS

FELLOWS OF THE ASSOCIATION

"Persons who have rendered eminent service to animal breeding in Australia and/or New Zealand or elsewhere in the world, may be elected to Fellowship of the Association…"

Elected February 1990

R.B.M. Dun
F.H.W. Morley (deceased)
A.L. Rae
H.N. Turner (deceased)

Elected September 1992

K. Hammond

Elected July 1995

C.H.S. Dolling
J.R. Hawker
J. Litchfield

Elected February 1997

J.S.F. Barker
R.E. Freer

Elected June 1999

J. Gough
J.W. James

Elected July 2001

J.N. Clarke
A.R. Gilmour
L.R. Piper

Elected September 2005

B.M. Bindon
M.E. Goddard
H.-U. Graser
F.W. Nicholas

Elected September 2007

K.D. Atkins
R.G. Banks
G.H. Davis

Elected September 2009

N. Fogarty
A. Fyfe
J. McEwan
R. Mortimer
R. Ponzoni

HONORARY MEMBERS OF THE ASSOCIATION

"Members who have rendered eminent service to the Association may be elected to Honorary Membership…"

Elected September 2009

W.A. Pattie
J. Walkley

NEAL M. FOGARTY

Neal began his career in the sheep industry in 1964 as a university trainee in Wool and Pastoral Science at the University of New South Wales. At the completion of his studies he started work at Trangie Research Station as a Wool Development Trainee. He obtained his M.Sc. degree from the University of New South Wales in 1977 and his Ph.D. from the University of Nebraska, USA in 1981.His scientific standing and proficiency has been recognised by university colleagues through his co-supervision of post graduate students.

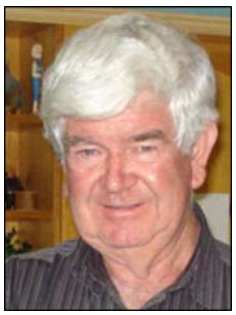
Neal's research career has been clearly focused on sheep meat research. He has long maintained an interest in sheep reproduction, maternal genetics, dual purpose sheep and the uptake of these findings by the lamb industry. One of Neal's earliest projects was concerned with the reproductive performance and health of Border Leicester sheep including studying the genetics of pneumonia. He conducted much of the early research on the genetic improvement of the Dorset Horn breed in Australia, particularly the relationship between pelvic size and dystocia which was, at the time, a significant industry issue. The development of the 'Hyfer' breed, a Dorset Merino cross infused with the Booroola gene, was another project in his research portfolio. The project focus was to supply the lamb industry with a sheep that bred all year round and was capable of producing a lot of lambs.

In 1995 Neal published a definitive review of genetic parameters for liveweight, fat and muscle measurements, wool production and reproduction in sheep. This review highlighted the extent and limitations of published information and outlined the requirements for future work on estimating genetic parameters. This work provided many of the genetic parameters that were fundamental to the early development of LAMBPLAN (and which are still used today).

Other projects including the diverse genotype experiment and elite lamb project in which Neal played integral roles successfully linked biology and genetics to production systems. Neal led and coordinated the national Maternal Sire Central Evaluation from 1997 which identified large amounts of genetic variation in maternal meat sheep (over $40 gross margin/ewe/year in profitability). During 2006 and 2007, Neal, as a Program Leader, led the establishment of the Sheep CRC's Information Nucleus at 8 sites around Australia. The Information Nucleus, which will further accelerate genetic improvement into the future,is a fundamental resource for the Sheep CRC and Australian sheep industries. Neal's work over the past 40 years has contributed to the so-called 'Lamb Revolution' which saw the lamb industry transformed from a poorly profitable by-product of the wool industry with genetic merit based on show ring and visual assessment into a vibrant and highly profitably industry in its own right that embraces new technology which has dramatically improved productivity and consumer appeal.

Neal was listed as a member of AAABG on 31 October 1979; he contributed at least 30 papers to fourteen conferences including being an invited speaker on Meat Sheep for the 18th Conference in 2009. He served AAABG as a contributing member, referee and for the 12th Conference in Dubbo he was Vice President and a member of the Editorial Committee. Neal officially retired from NSW DPI in March 2008; since then he has continued to be involved in the sheep industry as a part-time Post-Retirement Research Fellow.

ALAN R. FYFE



Alan Fyfe is a purebred pig breeder from Yelmah Stud, Hamley Bridge in South Australia. He has had a major influence on the Australian pig industry serving the industry and its members selflessly and with distinction for many decades. Furthermore, he is a long standing member of AAABG including being on 1984 and 1995 organising committees and has been an advocate of genetic improvement in Australia and New Zealand.

Alan's generosity of spirit has seen him share his experiences with other breeders. His strong working relationships with SARDI, PIRSA , the University of Adelaide especially the Roseworthy Campus research and extension staff and the Animal Breeding and Genetics Unit at Armidale has included access to his herd's performance data and the practices he has implemented in his breeding program – willing sharing of intellectual property at no cost to others.

Alan has been a member of the SA Swine Compensation Fund Committee which advised on investment of funds to benefit industry and on choice of research projects. As a member of the Swine Compensation Funds' successor, the SA Pig Industry Advisory Group he, and colleagues, provide advice to the Minister for Agriculture, Food and Fisheries on property registration, industry codes of practice, vendor declarations and relevant regulations for the pig industry as well as research priorities. For many years he has contributed to industry policy and development via the SA Farmers Federation's Commercial Pig Section. In the mid 1980s, as inaugural Chairman, he played a pivotal role in establishing a boar testing facility, SABOR Ltd, with the aid of Swine Compensation funding. Subsequently Alan became Chairman and Director of the SABOR Artificial Breeding Centre Ltd based in Clare and has overseen its growth from a Government supported unit to a fully privatised, commercially viable pig AI centre marketing semen nationwide. His expertise in pig genetic improvement resulted in a consultancy to assess genetic merit of a NZ pig herd.

Over the years Alan moved from showing to embracing objective measurement, genetic evaluation systems and index selection with Yelmah's success as a national leader reflected in both live pig and semen sales to all Australian states. Alan has written papers on using PIGBLUP in the Yelmah herd; "Practical experience with PIGBLUP" for the Australasian Pig Science Association (APSA 1989), "Selection Program Implementation on the Farm" (AAABG 1984) and "Efficient use of Recording Systems" (AAABG 1988). Alan has also made his data available for research; for example Horst Brandt used the Yelmah data to estimate genetic correlations between purebred and crossbred performance and Tom Long used Yelmah as a case study for genetic and financial evaluations of commonly used breeding systems.

He was the first licensed PIGBLUP user, signing a contract in November 1989, and became a member of various Pig Genetics Consultative Groups which guided the development of PIGBLUP during the early to mid 1990s and again from 2001 until 2006. His stud Yelmah is part of the National Pig Improvement Scheme - the across herd genetic evaluation system in pigs. Alan was one of the founding members of this scheme. Alan is truly a servant leader, an innovator and early adopter of new genetic technologies.

JOHN C. McEWAN

John Colin McEwan grew up on a sheep and beef farm near Tokanui, Southland, New Zealand (NZ), part of the property his great grandfather settled and developed from native bush after emigrating from Islay Scotland. His family also developed a stud that was one of the earliest registered Romney flocks. He received a BSc(Hons) in biochemistry from the University of Otago in 1978. After a year teaching, he joined the Ministry of Agriculture and Fisheries (now AgResearch) as a technician at Woodlands in Southland. In 1985 he transferred to An Foras Taluntais (now Teagasc), Ireland. In 1986, he returned to Invermay, near Dunedin, where he has remained until present, rising through the ranks to the highest science level within AgResearch. John has been a major force behind many initiatives aimed at improving the genetic merit of livestock, particularly sheep, in NZ, Australia and globally.

Sheep Improvement Ltd (SIL) undertakes genetic evaluations for the NZ sheep industry. John's involvements include its establishment in 1999, designing the computational framework, writing much of the code and management. A national across-breed analysis (advanced central evaluation; ACE) was introduced in 2003 and currently involves around 4 million animals. In 2002 John was instrumental along with Neville Jopson in the establishment of the Central Progeny Test to evaluate industry sires and provide across breed linkage to underpin ACE. In the early 1990's, John recruited farmers to record parasite resistance traits, leading to the WormFEC breeding service. John developed a carcass trait genetic evaluation system in 1996 based on the InnerVision CT facility (a joint venture with Landcorp Farming Ltd). For both these services, John initially collected and checked the data, ran the genetic evaluations, and generated client reports. They have now been incorporated into SIL. Perhaps John likes the challenge of difficult-to-measure traits; he is currently turning his attention to feed efficiency and green house gas emissions.

John helped elucidate the Inverdale gene. He helped Landcorp to verify, map (using his own code) and select the 'Carwell' gene affecting *L. dorsi* area, leading to the LoinMAX® test. John also helped to confirm and map a gene for muscling effects, with three Texel breeders, leading to the MyoMAX® test. John identified the region for microphthalmia, a recessive defect in NZ Texels resulting in blindness, using only 45 animals. This led the i-Scan® test. John's research also identified a region associated with a favourable combination of production traits and parasite resistance, leading to the WormSTAR™ test. John's efforts in obtaining global co-funding were instrumental in NIH moving cattle to the top of its livestock sequencing priorities. A project sequencing sheep at a comparatively low depth (3x) was able to use cattle as an assembly guide, leading to an ovine 50k SNP chip. John also established and led the highly regarded bioinformatics group for AgResearch at Invermay during 2000-1. Other examples of John turning his hand to computational and analytical tasks include methods such as Peddrift (AAABG 1997), TIPS (AAABG, 2001) and MELD, the latter used for assembling the skim sequence of the sheep genome.

John has still found time and energy to contribute to the wider scientific community through societies, international committees, mentoring and co-supervising graduate and post graduate students from a variety of universities. He has been a member of AAABG since 1990, served as vice president for 1999-2001, and contributed as a reviewer, session chair and author (29 papers). The NZ Society of Animal Production awarded him their McMeekan Memorial Award in 2003. He has 51 refereed journal articles and over 200 other publications.

ROBERT MORTIMER

Robert grew up on the family farm “Rosedale” working in a family partnership with his bother (Ted) and father (Les) who was a passionate tractor driver. Les‘s passion for wheat growing left an opening for the boys to develop their interest in sheep breeding without too much parental guidance. In 1969 the partnership expanded with the purchase of “Devondale” and Robert moved in with his new wife (Pam) to an undeveloped farm and the opportunity to start work on his own merino flock.

The story of how Robert tackled this challenge - “to start work on his own Merino flock” – is outlined in some detail on the Centre Plus website. What that story reveals, albeit written between the lines, are some of the traits that Robert, with assistance from family members and some like minds, bought to bear on the task (http://www.centreplus.com.au/about_us/our_story). These are fundamental to his achievements, and resonate with the traits that have characterised master breeders back to Robert Bakewell, and which underpin all effective animal breeding.

Those traits include: openness to new ideas and approaches coupled with careful assessment of all information and observations, extremely close and comprehensive observation and recording of performance, and very careful thought about what the animals are being asked to do. In short, sample wide diversity and choose the best (maximise selection intensity), assess carefully (heritability), target the right traits (breeding objective), and replace older, less effective ideas with better ones (optimise i/L). Anyone who has ever spoken to Robert about sheep breeding will recognise these traits, and the humble diligence and patience with which he applies them.

The outstanding example of how these traits have been expressed is the careful learning involved in applying BLUP methods to breeding Merinos: listening, thinking, adjusting, and being prepared to share that learning with anyone who cares to listen. In years to come this will be recognised as a pioneering achievement.

The other trait that Robert brings to his lifelong passion, and which he has clearly passed on both vertically and horizontally, is deep respect for all who are genuine and have something to offer. The Centre Plus philosophy is essentially open source innovation in real life, with no fuss or fanfare, and represents in very honest form the value that the diversity of human talents can add: many people contribute to the Centre Plus story on one way or another, and are in turn encouraged in their own sheep breeding or farming endeavours.

Robert’s contribution to industry development has included membership of a range of organisations focussed on the advancement of animal and Merino breeding, including the Federation of Performance Breeders, Industry Liaison Committees for the Q-Plus Research project, Macquarie CTSE site, the Dubbo AAABG Conference, Merinotech and Merino Benchmark.

Robert Mortimer is indeed a master breeder.

RAUL PONZONI

Raul Ponzoni has a B.Agric.Sc.(Uruguay) and in 1975 obtained his Ph. D. from the School of Wool and Pastoral Sciences, University of N.S.W., Australia supervised by Professor John James. Raul's international impact has no doubt been greatly assisted by his exemplary communication skills; he is fluent in oral and written English and Spanish with a good working knowledge of four other languages. From 1977 till 2002 he held positions of Senior Livestock Research Officer in the SA Department of Agriculture and Principal Research Scientist in the South Australian Research and Development Institute after which he joined his current organisation the WorldFish Center in Penang.

The outputs from Dr. Ponzoni's very distinguished and highly productive research in animal breeding and genetics impact at both scientific and industry levels. He is internationally recognised professionally bringing his un-doubted professional expertise to bear on genetic improvement in Australian industries including wool and sheep meat, beef cattle, alpacas, goats and abalone. More recently he has expanded this professional repertoire to fish breeding and genetics which includes dissemination of superior seed stock and promotion of integrated livestock and aquaculture systems for poverty alleviation in tropical environments.

Raul's scientific excellence has been acknowledged by his appointment as an Affiliate Associate Professor, University of Adelaide where he taught animal breeding and genetics while supervising post-graduate students. Numerous overseas consultancies in Uruguay, Brazil, Bangladesh and Lesotho plus his membership on an FAO Expert Panel on management of genetic resources are further evidence of his international standing. He presented invited papers and co-ordinated sessions on defining traits in breeding objectives and their associated economic values for sheep, goats and beef cattle at more industry focussed World Congresses on sheep and cattle. He wrote a booklet "Genetic improvement of hair sheep in the tropics "(1992,FAO) where his assiduousness produced an extremely clear guide on the essential principles of designing and conducting breeding programs.

In the 1980s Dr. Ponzoni was a key technical member of an Animal Production Committee (APC) Working Party examining the development of national performance recording schemes for wool and meat sheep to replace a number of State or University based and private systems. The Final Report to APC was a strong foundation on which Woolplan and Lambplan were developed. Features included a choice of traits in breeding objectives, choice of economic values including breeder specification and various combinations of characters which could be used to estimate overall genetic merit in dollars. These schemes provided options for breeders to submit pedigree records, age of dam, type of birth and management group for their sheep and have appropriate adjustments made. Quality assurance features included provision of test samples by the Australian Wool Testing Authority for round-robin trials of participating wool testing laboratories and regular processing of standard data sets by service providers. Woolplan subsequently evolved into Rampower and then MerinoSelect, a service offered by Sheep Genetics.

His outstanding work ethic is reflected by his publications: Books, Manuals and Proceedings (8), Scientific Literature (186 as senior or co-author), Other Reportings (62) with at least 6 in press, submitted or as drafts. AAABG has been enriched by Raul's involvement in all of our Conferences. He presented invited or contributed papers at 17 meetings and refereed papers for one Conference proceedings he does not appear to have attended. A Corresponding Member for the Inaugural Conference (1979) Steering Committee, he chaired the Editorial Committee (1984, Adelaide) and was on the 1987 Perth Conference Committee.

HELEN NEWTON TURNER MEDAL TRUST

The Helen Newton Turner Medal Trust was established in 1993 following an anonymous donation to the Animal Genetics and Breeding Unit.

The Helen Newton Turner Medal is awarded to provide encouragement and inspiration to those engaged in animal genetics. The Medal is named after Dr Helen Newton Turner whose career with CSIRO was dedicated to research into the genetic improvement of sheep for wool production. The Medallist is chosen by Trustees from the ranks of those persons who have made an outstanding contribution to genetic improvement of Australian livestock.

The Helen Newton Turner Medal was first awarded in 1994 to Associate Professor JohnJames and a list of all recipients to date is given below. The recipient of the Medal is invited to deliver an Oration on a topical subject of their choice. The Oration of the 2007 Medal recipient, Mrs Lucinda Corrigan, is reproduced in these proceedings.

Trustees of the Helen Newton Turner Trust are:

- Dr Richard Sheldrake (Chairman), representing NSW Department of Primary Industries
- Professor Brian Kinghorn, representing the University of New England
- Dr. Scott Dolling, representing the Association for the Advancement of Animal Breeding and Genetics
- Dr Roly Nieper, representing the National Farmers' Federation
- Dr. Hans-Ulrich Graser, Director, Animal Genetics and Breeding Unit

MEDALLISTS

1994	J.W. James	*2001*	G.A. Carnaby
1995	L.R. Piper	*2003*	F.W. Nicholas
1997	J. Litchfield	*2005*	K. Hammond
1998	J.S.F. Barker	*2007*	Lucinda Corrigan
1999	C.W. Sandilands		

Honorary members Elected 2009

WILLIAM A. PATTIE

William Anthony Pattie (Bill) was born at Leura NSW, and graduated from the University of New South Wales with a B.Sc. in 1960. He was appointed Livestock Research Officer at Trangie Agricultural Research Station, NSW, working with Bob Dun, where he conducted research into inheritance of growth and wool production of Merino sheep and crossbreeding systems in prime lamb production. Following postgraduate studies, supervised by Helen Turner, on realised selection responses in single character selection lines established at Trangie by Fred Morley, he was awarded a PhD in animal breeding by UNSW in 1966. In 1968, he was appointed Principal Livestock Research Officer, NSW Department of Agriculture, and was responsible for the direction of livestock production research in that Department.

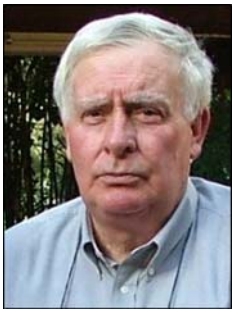
From 1971 to 1974, Bill served as Head of the Department of Animal Husbandry, Hawkesbury Agricultural College, NSW, where he was responsible for the establishment of a new department, developing curricula, teaching courses in genetics, animal breeding and wool technology. Then from 1974 to 1991, he was successively Lecturer, Senior Lecturer then Reader in Animal Breeding in the Faculties of Veterinary Science and Agricultural Science at the University of Queensland, teaching Animal Genetics and Breeding. He also supervised research Masters and PhD students and developed computer assisted learning programs for Genetics and animal Breeding. Bill's research activities involved studies of the inheritance of growth, reproduction and cashmere production in Australian goats and the design of cattle, sheep and goat breeding programs.

As well as his work in Australia, Bill has worked in Indonesia, Thailand, Vietnam, China, France, The Netherlands and the South Pacific. He has made significant contributions in Australia and overseas in the areas of university management and development, academic staff development, academic program and curriculum development in agriculture, distance education, animal breeding and genetics.

From 1991 to 1997, Bill was Pro Vice-Chancellor, Professor of Agriculture and Head of the School of Agriculture at the University of the South Pacific. As Pro Vice-Chancellor, he was responsible for the Alafua campus and its associated teaching farms, the School of Agriculture and the Institute of Research, Extension and Training for Agriculture in Samoa. As Professor of Agriculture, he was responsible for the academic program in Agriculture for the University of the South Pacific, which serves 12 countries of the region. During his tenure, all academic programs in Agriculture were revised and restructured, while courses in the B.Agr. program were made available by distance education.

Since his retirement at the end of 1997, Bill has maintained his interest in Animal Breeding and Production and has devoted much of his time to assisting with the organisation of the last three AAABG conferences and producing the conference proceedings. He has developed the AAABG internet web site to be the main channel of communication and operational support for the Association and its members. This includes on-line databases and secure financial transaction facilities for membership subscriptions and conference registrations. The web site also provides a full range of editorial support for the conference proceedings from receipt of papers, review and revision to production of the proceedings. Now, to mark the 30th anniversary of the Association, Bill is putting the proceedings of all the AAABG conferences on the Association's web site, fully indexed and searchable.

JIM WALKLEY



James Richard Whenan Walkley (Jim) was born at Moonta on SA's Yorke Peninsula graduating from the University of Adelaide with a B.Agric.Sc. in 1968. From 1969 to 1977 he worked with Scott Dolling as a Livestock Research Officer in the SA Department of Agriculture managing research projects to estimate phenotypic and genetic parameters for Merino sheep and to extend technical information on sheep breeding to wool growers.

During 1977/78 he studied for a Master of Science (Animal Breeding) at the University of Edinburgh, and supervised by the late Prof. Charles Smith prepared a thesis on the use of traits measured in males and females to estimate breeding values for reproductive rate in sheep.

During the 1980s he was Executive Officer, Woolplan the national sheep performance recording scheme developed on behalf of Animal Production Committee/Standing Committee of Agriculture.

In 1987 he was a member of the Organising Committee for the Australian Wool Corporation's "Merino Improvement Programs in Australia" conference at Leura which advised their Production Research Advisory Committee on R, D and E priorities for Merino breeding and genetics.

Jim held various research management positions in the SA Department and SARDI from 1987 to 1994 including: A/Principal Livestock Research Officer leading the sheep and wool research group consisting of six scientists and their technical staff. He was also Principal Officer, Research Management and a foundation member of SARDI's "Management Committee" and briefly their Interim Chief Livestock Scientist managing 36 State and 12 externally funded staff.

In May 1995 he started work with the Queensland Department of Primary Industries as a Principal Geneticist then was Acting Industry Manager, Beef and Sheep for several years responsible for some 80 staff and then as a Program Leader in the Queensland Beef Industry Institute and its successors.

From mid-2002 till the present he has been involved in Beef CRCs as a Program Manager, a member of the Re-bid Team for the CRC for Beef Genetic Technologies and again as a Program Manager.

During his career he has been actively involved with both the sheep and later the Northern beef industries in the use of genetic technologies. He assisted the SA Stud Merino Sheep Breeders' Association organise conferences in regional SA and Broken Hill and in Queensland he was Secretary of the South and SE Regional Beef Research Committees which provide advice on R, D & E priorities to the Northern Australian Beef Research Council and on project funding to MLA's Northern Beef Program Industry Committee .

Jim joined AAABG in 1979 attending 13 conferences up to and including 2009. He authored/co-authored and refereed papers and in 1984 was on the Editorial, Newsletter and Social Committees. Until he left SARDI in July 1994 he was Secretary of the 1995 conference Organising Committee. For the 16th AABG conference in Noosa Lakes he was Vice-President and is a member of the Organising Committee for the 2009 meeting in the Barossa Valley.

Jim has provided an outstanding example through his commitment to our employers, professional associations and livestock industries and in mentoring junior staff.

HELEN NEWTON TURNER MEDALLIST ORATION 2007

CREATING VALUE FROM GENETIC IMPROVEMENT

Lucinda Corrigan – 2007 Medal Recipient

"Old Renny Lea" Bowna via Albury NSW 2642

Trustees of the Helen Newton Turner Medal Trust, Professor Van Der Werf, ladies and gentlemen.

My challenge is "where to start?". When I read the orations from previous medal recipients, it was clear I needed to talk about where I spend my time, which is at the interface of science and industry.

Harold Wilson said, "Courage is the art of being the only one who knows you're scared to death".

This may be true on public occasions, but to lead into the future, Mark Twain captures the sentiment aptly;

"It is curious that physical courage should be so common in the world and moral courage so rare"

I looked up the word oration, thinking it may give some further clues on how to approach this event. "Speech of a ceremonial kind or harangue" from the Oxford dictionary. And a harangue is "A speech to an assembly, a loud or vehement address."

I promise not to harangue!

I could spend this valuable time, describing the world I was brought up in, on a Riverina sheep station, working large mobs of merinos in dusty yards, sheep sales where tweeded Merino breeders parted long, Peppin staples with loving hands, never ending discussions on sheep breeding between my father, Martin Bell, and his close friends, Corriedale breeder Peter Sloane from 'Neyliona' and Merino studman Raymond Taylor from 'Pooginook'.

Or, following university, joining the newly established cashmere goat industry, under the tutelage of wool and mohair supremo, Fred Moylan, breeding 4,000 does at Kinross Cashmere Company, eventually starting a cashmere group breeding scheme with Doug Winter, who had worked with Dr. Barrie Restall at Wollongbar Research Station and then the Australian Merino Society. Or, joining a fourth generation farming business in 1986, where Bryan had started selling small numbers of Angus bulls a few years earlier; to expanding the herd to its current scale.

1. However, I thought I would start with maternal genetics!

Being female and accepting a great female scientist's legacy, it seemed an appropriate place to start. Helen Newton Turner's name was well known to me by the time that Dr. Frank Nicholas taught us genetics at university, her efforts hotly debated by the tweeded merino men and not always with reverence!

Traditionally, the role of females in genetic improvement systems can be overlooked, when the contribution of genes by the male of the species is so much greater.

Consider the value of the following group of cows, and their attributes in 2009. (All these cows are searchable on the Angus Australia website under these identities)

NOR V97 – born in 2000, this high performing cow has had 57 progeny, and grossed over $90, 000 in income. She has high performing carcase genetics, in the top 1% of the breed for marbling and top 5% for the Certified Australian Angus Beef, Long Fed index. She is an Angus Performance Registered cows.

NOR W449 – Born in 2001, this elite cow has had 60 progeny, producing sale topping bulls a few times. Her performance is at the very top of the breed in carcase and fertility genetics. She is also an Angus Performance Registered cow.

VBB M14 – Born in 1992, we purchased this cow when she was young and she has produced over $100,000 income with a total of 112 progeny. She is an Angus Performance Registered cow.

VLY Y5 – Born in 2003, this is a high performance cow with 18 progeny, whom we purchased when young.

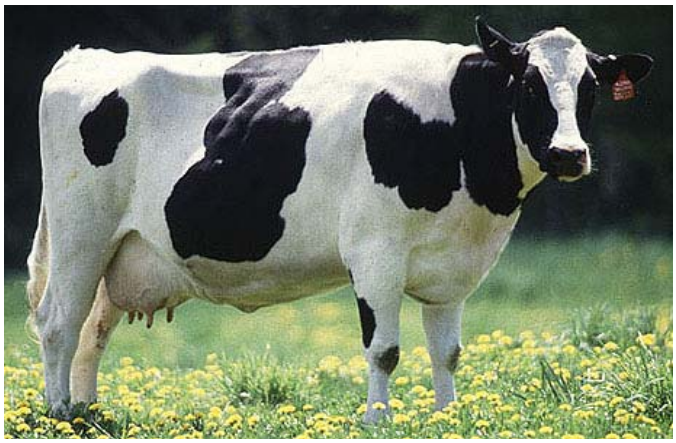
These are all elite, proven breeding cows, the first three are Angus Performance Registered, and the VLY Y5 is a Herdbook registered cow. On market value, the Y5 cow is worth much more than the other cows collectively, based on her registration status. On a return on capital basis, the others are way out in front and their genetic merit is judged by the performance of their offspring and reflected in their own breeding values.

That is solely due to their genetic description, created by the collective efforts of the Australian industry in genetic improvement. The APR cows would have remained anonymous and undescribed, had it not been for the ability to measure and record their performance in our national evaluation system Breedplan. You may agree with me, that it is an anachronistic system, where

value is created by a studbook entry category, rather than any relationship to their value in the supply chain.

2. Costs and a Serious Dilemma

The average Australian dairy cow eats over 200MJ of energy a day compared with an average Bos Taurus beef cow who consumes around 120MJ. Contrary to some proponents, I don't believe the future for the beef industry is ever increasing cow size, like the dairy industry, because we are competing internationally and with countries who have very different cost structures to our own. We have an increasingly variable climate, where water has become a precious input, where dry matter production for the cowherd competes with other uses for fodder and grain, such as energy. As emissions trading is introduced into this country, the urgency to improve the efficiency of beef production in variable climates is an immense challenge. More than ever we need the accurate tools to select animals with the right traits for this challenge.



I will come back to the dairy cow a little later on.

3. Data Quality

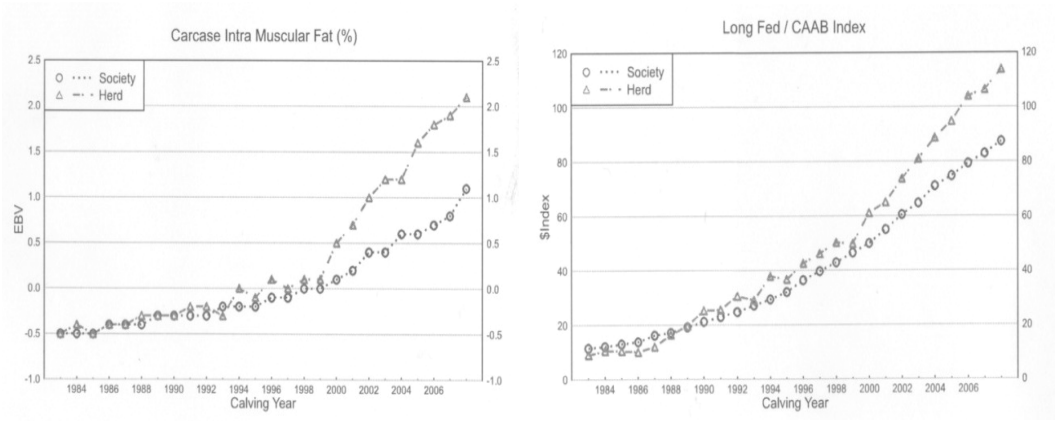
Many breeders forget about the females, and there is very poor data collection on females in many herds. When Dr. Peter Parnell and I gathered data from ABRI in 2006 for our paper, on "The Uptake of Genetic Technologies in Southern Australia", the low level of recording was an eye opener. In some high profile herds in this country, scan data is only collected on the male progeny and not the female. Many herds are not collecting and submitting birthweights, mature cow weights, days to calving and other simple to collect data. For the commercial client, that should set the alarm bells ringing, because the breeder uses Breedplan for marketing but has not made the connection between using Breedplan for genetic selection and improvement.

At the time of the 2007 AAABG conference, one major Bos Taurus breed had stopped reporting mature cow weight, due to the lack of data. At this stage, this is the most important trait for selecting for maturity pattern. The relationship between maturity pattern, fat cover and fertility in southern Australian is the topic of investigation in the Beef CRC's Maternal Efficiency project. Feed efficiency in an ad libitum situation is also being investigated. Being able to use these data for accurate selection is a matter of urgency for the industry.

The beef cow in a carbon constrained world will need to breed quickly, produce progeny that grow evermore quickly to meet market specifications, with a moderate maturity pattern that minimises her own carbon emissions.

January 2009 ANGUS GROUP BREEDPLAN
Herd Rennylea – Comparison to Breed Genetic Trends

These genetic trends, illustrate what the progress that is possible using embryo transfer. The key tool required is to be able to monitor inbreeding. These tools are available at the press of a button in 2009.



4. Industry Professionalism

I have always had a soft spot for dairy cows. In the early 1960s, Saturday work was dropped from the Federal Pastoral Industry award. Hence my father had a problem, who would milk the cow on the weekend? He discovered his six year old could fulfil the job, and after an hour or so I would appear back at the house with a gallon of milk in the bucket and more on me!

When I married Bryan, ever the romantic, he gave me a milking cow, and I milked Jane and her successors for ten years, escaping from young children to the cowshed, where the first rays of the sun would hit my face, tucked into the cow's flank!

Sometimes in this industry, the specialist genetics producer can feel like a milking cow. We pay to collect data, we submit this to the sole provider of genetic services, we pay to put it in, we have to pay to get it out again, we are bound by many useless regulations that do not create value, and do not take into account business risk, our commercial clients may not access the data on the bulls they purchase from us if they have a certain category of membership, in a system paid for by the levy payers and tax payers of Australia.

The Beef Genetic Improvement systems were commercialised without minimum standards, which continues to reduce the professionalism of the genetics industry. There are others who comment on this. Industry expert Bob Freer tells the story about visiting the surgeon who is also a stud breeder and while he is eminently qualified to be surgeon, has no training to breed cattle. It is inconceivable that such a situation would be tolerated in any other profession. It is an industry with no barriers to entry.

5. Value for the Commercial Industry

As Bos Taurus cattle are used across Australia in a greater range of environments, the need for accurate selection is intensified. We supply bulls to producers in the centre of Australia, west of the Simpson desert where the competition for energy is life threatening in some seasons. Their

specifications for low milk, moderate maturity pattern, positive carcase quality are important for this environment. If the simple traits were well recorded on 100% of the database, the reliability would deliver greater accuracy and commercial value to these producers, and the supply chain.

After some serious self examination, and reading the contributions of the previous awardees, I am humbled by the honour that the selection committee conferred on me. It is not always a comfortable place, to decide to challenge the status quo. In that role I have some admirable role models. Recently I was listening to Mr. James Litchfield, who received the Helen Newton Turner medal in 1997. His final statement was extremely telling "Nothing has changed", he said "the conservatism we faced when we introduced performance testing, from the early 1960s, is alive and well today."

In his presentation in 2003, Professor Frank Nicholas mentioned his inherited colour blindness as a major factor in his choice of career. I cannot claim the same, that my webbed toes (two on each foot) had any similar effect!

Bryan's breeding skills, our love of livestock, and my background in applied research and communication combined, have been an insightful journey. I acknowledge his generosity and the partnership we have been able to create together. As we face the challenge of running a beef genetics business in an increasingly variable climate, we need to keep drawing on the intellectual discovery in the science community, to work together to solve these problems.

While I was not fortunate to meet Dr. Helen Newton Turner, she has been an inspiration throughout my life. She led CSIRO's sheep breeding programme from 1956 to 1973, and laid the foundations for the selection and improvement of merino sheep using objective measurement. After she retired from the CSIRO, she dedicated her time to animal production in developing countries, and travelled all over the world. Colleagues tell a story, after working together one day in the 1960s, they retired to the Canberra Club. Dr. Newton Turner was not allowed to enter and they went elsewhere! Times have changed.

Our ruminant industries require productivity increases that will keep up with declining terms of trade. Genetic improvement is permanent and cumulative, using an array of tools and clear breeding goals. Our industries need these improvements as we face new challenges and come under constant scrutiny from the very urban Australia of the 21st Century.

Reference

Corrigan, L. and Parnell, P.F. (2006); Application of Genetic Technologies in the Temperate Australian Seedstock Industry, in 'Australian Beef The Leader, The Impact of Science on the Beef Industry', CRC for Beef Genetic Technologies, pp 81-90.

HELEN NEWTON TURNER

TABLE OF CONTENTS

CITATION OF PAPERS

Papers in this publication should be cited as appearing in the *Proceeding of the Association for the Advancement of Animal Breeding and Genetics* (abbreviation: Proc. Assoc. Advmt. Anim. Breed. Genet.).

For example:

Young, M.J. and Amer, P.R. (2009) Rates of Genetic Gain in New Zealand Sheep. *Proc. Assoc. Advmt. Anim. Breed. Genet.* **17**:422.

REVIEWERS

All papers, invited and contributed, were subject to peer review by two referees who are listed below. The allocation of papers to oral or poster presentation was based on the facilities available, the conference program and themes.

The following people are acknowledged and thanked for their efforts in reviewing papers:

Peter Amer
Jason Archer
Paul Arthur
Kevin Atkins
Rob Banks
Steve Barwick
Kevin Beard
Hugh Blair
Forbes Brien
Daniel Brown
Kim Bunter
Heather Burrow
Yizhou Chen
Schalk Cloete
Ron Crump
Mick Deland
Ken Dodds
Sonja Dominik
Kath Donoghue
Sandra Eady
Nick Elliott
Peter Fennessy
Troy Fisher
Malcolm Fleet
Neal Fogarty
Ian Franklin
Mick Garrick
John Gibson
Arthur Gilmore
Mike Goddard
Cedric Gondro
Hans Graser
Johann Greef
Mekonnen Haile-Mariam
Sue Hatcher
Ben Hayes
Michele Hebart
John Henshall
Susan Hermesch
Jay Hetzel
Stefan Hiendleder
Paul Hughes
Phil Hynd
John James
David Johnston
James Kijas
Brian Kinghorn
Kon Konstantinov
Peter Kube
Greg Lee
Brian Luxford
Michel MacBeth
Oliver Mayo
John McEwan
Brian McGuirk
Will Mckiernan
Karin Meyer
Chris Moran
Chris Morris
Sue Mortimer
Greg Natrass
Scott Newman
Frank Nicholas
Belinda Norris
Peter Parnell
Laurie Piper
Wayne Pitchford
Raul Ponzoni
Kishore Prayaga
Ian Purvis
Simon Quigley
Hermann Raadsma
Tony Reverta
Nick Robinson
Ross Shepherd
Bolorama Sunduimijid
Andrew Swan
Ross Tallam
Bruce Tier
Ari Verbyla
Peter Visscher
Rob Woolaston
Yuandan Zhang
Julius van der Werf

AUTHOR INDEX

Proc. Assoc. Advmt. Anim. Breed. Genet. 18:1-6

IMPRINTED GENES IN MAN AND MOUSE ARE MODEL SYSTEMS IN COMPARATIVE EPIGENOMICS

Martina Paulsen

Saarland University, FR 8.3 Life Sciences, PO Box 151150, D-66041 Saarbruecken, Germany.

SUMMARY

In mammals, epigenetic modifications are key players in gene regulation and genome stability. Consequently, the epigenetic protein machinery and epigenetically modified regulatory elements, such as promoter regions, are highly conserved among mammals. Hence, comparative studies of mammalian epigenomes may help to understand the mechanisms and functions of epigenetic gene regulation. Imprinted genes that are mono-allelically expressed due to allele-specific epigenetic modifications of their regulatory sequences have been recognized as ideal model systems in epigenetics. For this reason, detailed comparative studies in epigenomics were firstly initiated on imprinted genes. These analyses have resulted in the identification of new imprinted genes and regulatory elements and have highlighted complex patterns of conservation that includes not only sequence conservation but also structural elements such as the presence of tandem repeats and retrotransposed elements. Comparative studies on these genes have been extended to other topics such as the conservation of tissue-specific gene expression patterns. These analyses show that the tight conservation of epigenetic regulation of imprinting does not prevent the divergence of tissue-specific gene expression patterns that might be associated with new species-specific functions of imprinted genes.

EPIGENOMICS - RELATIONSHIPS BETWEEN GENETIC INFORMATION AND CHROMATIN STRUCTURE

In eukaryotic species, inheritance of information is not only based on the sequence of the DNA but also on epigenetic modifications of the chromatin. These modifications include modification of histone proteins and modifications of the DNA that do not affect the DNA sequence (Jenuwein and Allis 2001). Both types of modifications determine the structure of the chromatin and thereby, influence the expression of genes. Especially for DNA methylation, it has been shown that methylation patterns once they are established can be transmitted through numerous cell divisions. Nevertheless, especially in mammals, epigenetic modifications are substantially changed during development. Epigenetic reprogramming affects development of germ cells, embryonic stem cell development and differentiation processes, thereby indicating that epigenetic research has a substantial input into stem cell and reproduction research (Hemberger et al. 2009, Dean et al. 2001).

During the last few years, the genomic sequences of numerous mammalian species and an increasing number of human individuals have been made available. In parallel, experimental techniques have been established to analyse epigenetic modifications on a genome-wide level (Weber et al. 2005, Zhang et al. 2009, Roh et al. 2005, Bernstein et al. 2004). Current research in epigenomics focuses on the understanding of interactions between proteins involved in epigenetic processes and regulatory sequences, such as CpG islands in promoter regions, that attract a distinct type of epigenetic modifications. For example, the presence of specific DNA sequence motifs, the structure of the double helix, and overlap with repetitive elements have a strong influence on the chromatin structure of CpG islands (Bock et al. 2006, Bock et al. 2007).

The comparison of genetic and epigenetic features in different species allows, on one hand, the efficient identification of conserved genetic elements involved in epigenetic gene regulation, and on the other hand, highlights epigenetic differences that may restrict the usage of model

organisms, such as the mouse, in fields like epigenetic topics in medical research. On the level of species, it has become clear that mammalian species possess a well-conserved epigenetic machinery that encompasses evolutionary conserved proteins, such as DNA methyltransferases, histone modifying enzymes, and proteins that bind to specific chromatin structures (Yokomine et al. 2006, Bestor 2000). Also regulatory sequences, such as promoters, are highly similar in different species. Nevertheless, comparisons of the human and mouse genomes have shown that these species may exhibit some differences in their epigenetic regulatory elements. For example, CpG islands in the human are longer than in the mouse, and due to the acquisition of lineage-specific Alu elements, the human genome possesses more CpG islands that reside in repetitive elements (Hutter et al. 2009, Zhang et al. 2004). Besides differences between species, differences among (human) individuals receive more focus in epigenetic research, since such these differences might contribute to phenotypic diversity and might represent additional risk factors in human disease.

IMPRINTED GENES ARE MODELS FOR EPIGENETIC GENE REGULATION

In therians, a number of genes are expressed only from one of the two parental chromosomes. These so-called imprinted genes acquire different epigenetic modifications in the parental germlines that are maintained after fertilization. Subsequently, these different epigenetic marks result in silencing of one gene copy whereas the other copy remains active. As the mono-allelic expression of imprinted genes depends solely on the differential epigenetic modifications of the parental gene copies, these genes represent an ideal model system in epigenetics.

Imprinted gene expression is a specific way of gene regulation that is seen in eutherian species and to some extent also in marsupials, but appears to be absent in other vertebrate species (Killian et al. 2001). Therefore, imprinted genes are intensively investigated in human and mouse, and to some extent also in cattle and marsupials (Gebert et al. 2006). The comparison of these genes in different species has several purposes. Firstly, it aims to identify features of the DNA sequence that are responsible for establishment and maintenance of mono-allelic gene expression. A second goal is the identification of all genes that are prone to be imprinted. Last but not least, the comparison of imprinted genes in different species should help to understand the evolution of imprinting in therian species and should highlight conserved functions of these genes in this clade.

To date, more than mammalian 150 transcripts imprinting effects have been noted (Morison and Reeve 1998, http://igc.otago.ac.nz/home.html). With few exceptions, genes whose allele-specific expression patterns have been analysed in human as well as in mouse are imprinted in both species, indicating strict conservation of imprinting among eutherian species. Most imprinted genes are organized in so-called imprinted regions in the mammalian genome, i.e. imprinted genes are neighboured by other imprinted genes (MRC Harwell, UK, http://www.har.mrc.ac.uk/research/genomic_imprinting/index.html). In these regions, central differentially methylated regions (DMR) control the mono-allelic expression patterns of neighbouring genes. Strict conservation of the overall physical structures of imprinted regions, as indicated by the conserved presence and order of orthologous imprinted genes, suggests that an evolutionary conserved gene arrangement is required for cis allele-specific interactions between genes and regulatory elements (Paulsen et al. 2005, Paulsen et al. 2001). A typical feature of imprinted regions is the presence of evolutionary conserved non-coding RNAs. This includes small RNAs, such as microRNAs and snoRNAs, and longer non-translated transcripts that often represent antisense-transcripts of protein-encoding genes. The lack of conserved non-coding RNA genes in the vicinity of the orthologs of imprinted genes in non-mammalian vertebrates indicates that the evolution of these non-coding RNAs coincided with the evolution of imprinted gene expression in the respective genomic regions (Paulsen et al. 2005, Edwards et al. 2008, Smits et al. 2008). Especially for long imprinted non-translated transcripts, it has been shown that they

mediate cis long-range epigenetic silencing of overlapping or neighbouring protein-encoding genes (Fitzpatrick et al. 2002). Hence, the establishment of non-coding transcripts in these regions might have been a crucial event in the evolution of imprinted gene expression.

REPETITIVE ELEMENTS IN IMPRINTED REGIONS

In mammalian genomes, retrotransposable elements are usually epigenetically silenced, thereby preventing retro-transposition events especially in the germlines. Since LINE1 elements are enriched on eutherian X chromosomes, it has been proposed that repetitive elements might be involved in epigenetic silencing processes (Bailey et al. 2000). Consistent with this hypothesis, unusual densities of repetitive elements have been observed in imprinted regions in human and mouse (Greally 2002, Walter et al. 2006). This affects mostly SINE elements that are under-represented in imprinted regions. However, so far, there is no experimental proof that these elements might indeed exhibit special allele-specific patterns of epigenetic modifications in imprinted regions. Similarly, there are no mouse models available in which such elements have been deleted or amplified in order to show that they attract indeed germ-line specific DNA methylation or histone modification patterns.

The unusual densities and distribution of repetitive elements in imprinted genes have been exploited by bioinformatic approaches addressing the identification of new imprinted genes (Luedi et al. 2005, Luedi et al. 2007) In the meantime, other studies used experimental approaches for the systematic genome-wide discovery of imprinted genes (Nikaido et al. 2003). Unfortunately, these studies showed little overlap of candidates and subsequent studies which tried to validate imprinted gene expression of predicted candidates showed that only few of them were indeed imprinted (Ruf et al. 2007). This indicates that the specific characteristics of retrotransposed elements in imprinted genes are not a feature that efficiently distinguishes imprinted genes from bi-allelically expressed genes.

SPECIAL FEATURES OF IMPRINTED CPG ISLANDS AND DIFFERENTIALLY METHYLATED REGIONS

Allele-specific gene regulation in imprinted regions is mediated by a few DMRs. These elements often contain direct repeats and tend to be CpG rich, i.e. they frequently overlap with CpG islands. Though CpG islands are key elements in epigenetic gene regulation and, therefore, are believed to be strictly conserved, the CpG islands of imprinted genes show some species-specific features. Due to the depletion of SINE elements, imprinted genes in human possess less CpG islands than randomly selected genes (Hutter et al. 2006). A similar effect is not seen in the mouse, since in this species, repetitive elements rarely overlap with CpG islands. In the mouse, many imprinted genes possess intronic CpG islands that may serve as promoters of antisense transcripts, whereas a similar enrichment does not reach statistical significance in the human.

Nevertheless, in both species, CpG islands of imprinted genes are enriched in direct repeats. Interestingly, the repeated motifs are highly divergent, in a way that the CpG islands or DMRs of different genes contain different repeated motifs (Hutter et al. 2006). In addition, they are not conserved in orthologous CpG islands or DMRs in different species (Paulsen et al. 2005). Hence, the presence of direct repeats in the CpG islands and DMRs of imprinted region represents a conservation rather of DNA structure than of DNA sequence. Though the DNA sequences of DMRs are obviously not highly conserved, the presence of repeats seems to be sufficient for conserved regulatory functions of these elements. It has been shown for several DMRs that the allele-specific patterns of DNA methylation are conserved in several species. Targeted deletions of DMRs in mouse and deletions of DMRs in human patients suffering imprinting disorders show that these elements fulfill conserved regulatory functions in both species, and for some DMRs, it

has been shown that their functions (for example, as promoters of non-coding transcripts) are conserved between mouse and human (Lee et al. 1999, Fitzpatrick et al. 2002).

TISSUE-SPECIFIC EXPRESSION PATTERNS OF IMPRINTED GENES INDICATE FUNCTIONAL DIVERGENCE

The evolution of imprinting in therians is believed to be related to the evolution of the placenta as a permeable interface between the embryo and its mother, and imprinted genes are supposed to function predominantly in regulation of embryonic growth and nutrient supply during prenatal development. However, some imprinted genes are only weakly expressed in the placenta or are expressed at pronounced levels in other organs during postnatal stages. In order to address conservation of tissue-specific expression patterns of imprinted genes, we have evaluated tissue-specific micro-array expression data of these genes in human and mouse, thereby showing that strong inter-species conservation of tissue-specific expression is restricted to few imprinted genes and few tissues (Steinhoff et al. 2009). Among these are organs such as adrenal gland, pancreas and pituitary that are involved in endocrinal functions. This suggests that a major evolutionary conserved function of these genes is indeed in regulating growth and nutrient uptake. The identification of conserved binding sites for tissue-specific transcription factors in the promoter regions of genes that are expressed in the corresponding tissues highlights the usefulness of comparative approaches for the detection of tissue-specific regulatory elements.

The pronounced expression of imprinted genes at postnatal stages suggests that postnatal functional aspects of imprinted gene expression will be an interesting field for future research. In addition, the unexpected evolutionary divergence in expression patterns of imprinted genes indicates that these genes are capable of establishing species-specific functions. Hence, future comparative investigations on imprinted genes might also require additional mammalian species as alternative model systems.

REFERENCES

Bailey, J.A., l. Carrel, A. Chakravarti, E. E. Eichler. 2000. Molecular evidence for a relationship between LINE-1 elements and X chromosome inactivation: the Lyon repeat hypothesis. *Proc Natl Acad Sci U S A* **97**: 6634.

Bernstein, B.E., E.L. Humphrey, C.L. Liu, and S.L. Schreiber. 2004. The use of chromatin immunoprecipitation assays in genome-wide analyses of histone modifications. *Methods Enzymol* **376**: 349.

Bestor, T.H. 2000. The DNA methyltransferases of mammals. *Hum Mol Genet* **9**: 2395.

Dean, W., F. Santos, M. Stojkovic, V. Zakhartchenko, J. Walter, E. Wolf, and W. Reik. 2001. Conservation of methylation reprogramming in mammalian development: aberrant reprogramming in cloned embryos**.** *Proc Natl Acad Sci U S A* **98**: 13734.

Edwards, C.A., A.J. Mungall, L. Matthews, E. Ryder, D.J. Gray, A.J. Pask, G. Shaw, J.A. Graves, J. Rogers, I. Dunham, M.B. Renfree, and A.C. Ferguson-Smith. 2008. The evolution of the DLK1-DIO3 imprinted domain in mammals. *PLoS Biol* **6**: e135.

Fitzpatrick G.V., P.D. Soloway, M.J. Higgins. 2002. Regional loss of imprinting and growth deficiency in mice with a targeted deletion of KvDMR1. *Nat Genet.* **32**:426.

Gebert, C., C. Wrenzycki, D. Herrmann, D. Groger, R. Reinhardt, P. Hajkova, A. Lucas-Hahn, J. Carnwath, H. Lehrach, and H. Niemann. 2006. The bovine IGF2 gene is differentially methylated in oocyte and sperm DNA. *Genomics* **88**: 222.

Greally, J.M. 2002. Short interspersed transposable elements (SINEs) are excluded from imprinted regions in the human genome. *Proc Natl Acad Sci U S A* **99**: 327.

Hemberger, M., W. Dean, and W. Reik. 2009. Epigenetic dynamics of stem cells and cell lineage commitment: digging Waddington's canal. *Nat Rev Mol Cell Biol* **10**: 526.

Hutter, B., V. Helms, and M. Paulsen. 2006. Tandem repeats in the CpG islands of imprinted genes. *Genomics* **88**: 323.

Hutter, B., M. Paulsen, and V. Helms. 2009. Identifying CpG Islands by Different Computational Techniques. *Omics*.

Jenuwein, T. and C.D. Allis. 2001. Translating the histone code. *Science* **293**: 1074.

Killian, J.K., C.M. Nolan, N. Stewart, B.L. Munday, N.A. Andersen, S. Nicol, and R.L. Jirtle. 2001. Monotreme IGF2 expression and ancestral origin of genomic imprinting. *J Exp Zool* **291**: 205.

Lee M.P., M.R. DeBaun, K. Mitsuya, H.L. Galonek, S. Brandenburg, M. Oshimura, and A.P. Feinberg. 1999. Loss of imprinting of a paternally expressed transcript, with antisense orientation to KVLQT1, occurs frequently in Beckwith-Wiedemann syndrome and is independent of insulin-like growth factor II imprinting. *Proc Natl Acad Sci U S A* **96**:5203.

Luedi, P.P., A.J. Hartemink, and R.L. Jirtle. 2005. Genome-wide prediction of imprinted murine genes. *Genome Res* **15**: 875.

Luedi, P.P., F.S. Dietrich, J.R. Weidman, J.M. Bosko, R.L. Jirtle, and A.J. Hartemink. 2007. Computational and experimental identification of novel human imprinted genes. *Genome Res* **17**: 1723.

Morison, I.M. and A.E. Reeve. 1998. A catalogue of imprinted genes and parent-of-origin effects in humans and animals. *Hum Mol Genet* **7**: 1599.

Nikaido, I., C. Saito, Y. Mizuno, M. Meguro, H. Bono, M. Kadomura, T. Kono, G.A. Morris, P.A. Lyons, M. Oshimura, Y. Hayashizaki, and Y. Okazaki. 2003. Discovery of imprinted transcripts in the mouse transcriptome using large-scale expression profiling. *Genome Res* **13**: 1402.

Paulsen, M., S. Takada, N.A. Youngson, M. Benchaib, C. Charlier, K. Segers, M. Georges, and A.C. Ferguson-Smith. 2001. Comparative sequence analysis of the imprinted Dlk1-Gtl2 locus in three mammalian species reveals highly conserved genomic elements and refines comparison with the Igf2-H19 region. *Genome Res* **11**: 2085.

Paulsen, M., T. Khare, C. Burgard, S. Tierling, and J. Walter. 2005. Evolution of the Beckwith-Wiedemann syndrome region in vertebrates. *Genome Res* **15**: 146.

Roh, T.Y., S. Cuddapah, and K. Zhao. 2005. Active chromatin domains are defined by acetylation islands revealed by genome-wide mapping. *Genes Dev* **19**: 542.

Ruf, N., S. Bahring, D. Galetzka, G. Pliushch, F.C. Luft, P. Nurnberg, T. Haaf, G. Kelsey, and U. Zechner. 2007. Sequence-based bioinformatic prediction and QUASEP identify genomic imprinting of the KCNK9 potassium channel gene in mouse and human. *Hum Mol Genet* **16**: 2591.

Smits, G., A.J. Mungall, S. Griffiths-Jones, P. Smith, D. Beury, L. Matthews, J. Rogers, A.J. Pask, G. Shaw, J.L. VandeBerg, and J.R. McCarrey, SAVOIR Consortium, M.B. Renfree, W. Reik, I. Dunham. 2008. Conservation of the H19 noncoding RNA and H19-IGF2 imprinting mechanism in therians. *Nat Genet*. **40**:971.

Steinhoff, C., M. Paulsen, S. Kielbasa, J. Walter, and M. Vingron. 2009. Expression profile and transcription factor binding site exploration of imprinted genes in human and mouse. *BMC Genomics* **10**: 144.

Walter, J., B. Hutter, T. Khare and M. Paulsen. 2006. Repetitive elements in imprinted genes. *Cytogenetic and Genome Research* **113**: 109.

Weber, M., J.J. Davies, D. Wittig, E.J. Oakeley, M. Haase, W.L. Lam, and D. Schubeler. 2005. Chromosome-wide and promoter-specific analyses identify sites of differential DNA methylation in normal and transformed human cells. *Nat Genet* **37**: 853.

Yokomine, T., K. Hata, M. Tsudzuki, and H. Sasaki. 2006. Evolution of the vertebrate DNMT3 gene family: a possible link between existence of DNMT3L and genomic imprinting. Cytogenet *Genome Res* **113**: 75.

Zhang, Z., N. Carriero, and M. Gerstein. 2004. Comparative analysis of processed pseudogenes in the mouse and human genomes. *Trends Genet* **20**: 62.

Zhang, Y., C. Rohde, S. Tierling, H. Stamerjohanns, R. Reinhardt, J. Walter, and A. Jeltsch. 2009. DNA methylation analysis by bisulfite conversion, cloning, and sequencing of individual clones. *Methods Mol Biol* **507**: 177.

Proc. Assoc. Advmt. Anim. Breed. Genet. 18:7-10

THE INFLUENCE OF ANIMAL BREEDING ON THE ORIGIN OF SPECIES BY CHARLES DARWIN

O. Mayo

CSIRO Livestock Industries, PO Box 10041, Adelaide BC, SA 5000

SUMMARY

Charles Darwin based his theory of evolution on the analogy of 'nature' selecting individuals for reproduction in the same way that plant and animal breeders chose parent stock in their breeding. However, the influence of animal breeding on the theory was all pervasive. The paper illustrates how Darwin used results from animal breeding. Advantages and disadvantages are emphasized.

INTRODUCTION

Any reader of the *Origin* in this, its sesquicentenary year, will be struck, first, by the metaphor of natural selection. In Darwin's words, '[the] preservation of favourable variations and the rejection of injurious variations, I call Natural Selection. … No one objects to agriculturalists speaking of the potent effects of man's selection; and in this case the individual differences given by nature, which man for some object selects, must of necessity first occur. … In the literal sense of the word, no doubt, natural selection is a misnomer; but who ever objected to chemists speaking of the elective affinities of the various elements? - and yet an acid cannot strictly be said to elect the base with which it will in preference combine.' This metaphor has been explored frequently, notably by Evans (1984), who followed the development of Darwin's thought through his notebooks and correspondence, and concluded that 'Darwin's recognition of the power of selection in changing organisms was almost entirely due to what he learned of plant and animal breeding.' (p. 133)

Darwin's all embracing theory needed data on the nature and origin of variability in natural populations, on the nature of inheritance, on the effects of the environment on variability and inheritance, on the fixity of species and on the effects of crossing species and varieties and on change over time. He recognised that breeders could provide much of this information.

VARIABILITY, INHERITANCE, EFFECTS OF ENVIRONMENT

Data. Darwin corresponded with breeders for over a decade, collecting materials. He wrote: 'About breeding, I know of no one book. I did not think well of Lowe, but I can name none better. Youatt I look at as a far better and *more precise* authority; but then his views and facts are scattered through three or four thick volumes. I have picked up most by reading really numberless special treatises and all agricultural and horticultural journals; but it is a work of long years. *The difficulty is to know what to trust.*' (Darwin to T. H. Huxley 29 Nov 1859 (Darwin, F. 1887, vol. 2, p. 281)) Published and privately communicated data were inadequate. Accordingly, he began his own breeding work using pigeons, which had many advantages: small size, low cost, clutch size, generation interval, life pairing, widespread use.

Table 1 shows quantitatively how Darwin drew on information from domesticated species. The different intensity of citation for the two main topics reflects both Darwin's interests and the state of the science at the time.

Variability. Darwin was particularly interested in the changes brought about under domestication

and then by selection on the domesticated species, so that, for example, domesticated flying birds such as ducks had reduced wing size and increased leg size compared with their wild progenitors. This showed both how directional selection could occur and its magnitude.

Table 1. Mention by Darwin of domesticated species in 300 pages dealing with variation, heredity and environmental effects and 60 pages dealing with embryology etc. (in 2nd ed., 475 pages in toto) in the *Origin of Species*

Species	Variation etc.	Embryology etc.
Pigeons	28	3
Other poultry	24	
Sheep	10	1
Cattle	12	3
Other ruminants	4	
Horses	14	3
Dogs	15	2
Other Mammals	18	
Bees	18	
Other insects	1	
Total	144	12

He noted phenomena such as reversion, i.e. appearance of an ancestral type, which we would explain today as segregation of a Mendelian recessive, and prepotency, which we would explain as another aspect of dominance and perhaps epistasis. Here, as elsewhere, Darwin was handicapped by his lack of a workable quantitative model for inheritance (see e.g. pp. 160-1 of Darwin 1859). Thus, he wrote 'The laws governing inheritance are quite unclear; no one can say why a peculiarity in different individuals of the same species, or in individuals of different species, is sometimes inherited and sometimes not so; why the child often reverts to its grandfather or grandmother or other more remote ancestor…' (Darwin 1859, p. 10). He did, however, carefully note reversion to the wild type in pigeons. See also Bartley (1992).

On prepotency, we should note that Yarrell's law (that an older breed would be prepotent when crossed with a newer breed, i.e. the progeny (F1) would resemble the older breed, a concept that interested Darwin: see Nicholls 2009) is a prediction of Fisher's theory of unconscious selection for dominance during the domestication of a species (Fisher 1935, 1938).

Darwin did not spell out details in the *Origin*, which he repeatedly referred to as an 'abstract' of a larger work which he began but did not complete. In later books, he presented much detail but not in the form of experimental results as in his plant hybridisation or earthworm studies. For example, in the lengthy pigeon chapters in Darwin (1890, vol. 1, pp. 166-7), he gave modal numbers of tail feathers for many breeds, but not the range or the shape of the distributions. This reflected both the state of small sample statistical methods and his own preferences for dealing with numerical data (see e.g. Leach and Mayo 2005). He used the same approach for data on number of vertebrae across breeds (Darwin 1890, p. 175). In presenting data on allometric growth, which he regarded as an important constraint on the direction of natural selection, so that human success in changing proportions was strong evidence of potential selective forces in nature, he presented the magnitude of the disproportion using the mode or arithmetic mean, with no indication of within-population variability (p. 181).

Use, disuse, nutrition and other environmental factors Darwin all saw as sources of variability on which natural selection could act. Domestication and breeding led to better nutrition, for example, so that increased fertility was manifested.

Darwin also devoted some space to a consideration of the deleterious effects of inbreeding, to which he returned later, at which later time he demonstrated both these effects and the (converse) beneficial effects of outcrossing (see Leach and Mayo 2005 for discussion and references).

Speciation. Though Darwin did not formally use the biological species definition, and indeed recognised many of its problems (Darwin 1859, p. 259, 269, 276), he regarded sterility on crossing as a good indicator of distinct species. He noted that many breeders held, for no good reason, that each major breed of a domesticated animal as having arisen from a different species, and hence used pigeons as a strong counter example. From his own work and that of others, he concluded: 'All the domestic races pair readily together, and, what is equally important, their mongrel offspring are perfectly fertile.' (Darwin 1890, p. 201) In his crosses between pure breeds, Darwin found that the F1 generally resembled one parent, with segregation in the F2. However, as is well known, his model for inheritance was not particulate, let alone Mendelian, and he did not present numerical data on pigeons that others might have been able to analyse.

What he was able to show was that a particular species (*Columba livia*) was highly variable and that the variations had been selected to produce varieties or races that were highly distinctive and could be, given isolation, incipient species. He was also able to reject the widespread confused belief that every major strain of domestic livestock had arisen from a separate species: 'May not those naturalists who, knowing far less of the laws of inheritance than does the breeder, and knowing no more than he does of the intermediate links in the long lines of descent, yet admit that many of our domestic races have descended from the same parents - may they not learn a lesson of caution, when they deride the idea of species in a state of nature being lineal descendants of other species?' (Darwin 1859, p. 29)

Darwin frequently took the word of scientists or breeders whom he had found to be reliable, and this could lead to absurd outcomes as he laboured to condense vast bodies of non-numerical data into a meaningful form. For example, he wrote that the Brahma breed of poultry had arisen recently 'by a cross, which can be truly propagated' (Darwin 1868, p. 258), with additional comments on the origin of the breed which so aroused the ire of a professional poultry man, L. Wright, that he pursued Darwin's 'error' for five pages of his huge book on poultry, even though Darwin corrected the passage to that quoted in his second edition (1875) and had been dead eight years when Wright's volume appeared (Wright, 1890, pp. 245-9). The key unusual evolutionary point, of a 'true-breeding' cross (F1 onwards) was never referred to again, but the fossilized posthumous debate remains in the literature. No usable data were presented by Wright in his discussion, and this contrasts particularly with Darwin's customary practice; he presented data so carefully that they can still be useful and can be analysed with statistical tools that Darwin never dreamt of (see Leach and Mayo 2005, Mayo 2009). In the present context, Darwin (1887, vol. 2, p. 71) presented good Mendelian single factor segregation data from snapdragon without recognising their significance, except insofar as they provided evidence of reversion.

CONCLUSIONS

Darwin used breeders' information and his own experiments wisely, to strengthen his arguments about the variability of populations, the mutability of species, and the power of selection.

His defective theory of inheritance and lack of statistical tools meant that he overstated some arguments and was wrong on other points e.g. reversion but, overall, evidence from animal

breeding was central to his theory and has been supported by almost all later work.

ACKNOWLEDGMENTS

I thank C. R. Leach for helpful comments on a draft of this paper and two referees for several improvements.

REFERENCES

Bartley, M. M. (1992) *J. Hist. Biol.* **25**:307.

Darwin, C. R. (1859) *On the Origin of Species by Natural Selection or the Preservation of Favoured Races in the Struggle for Existence*. John Murray, London.

Darwin, C. R. (1868, 1875, 1890) *The Variation of Animals and Plants under Domestication*. John Murray, London. (Page references from volume 1 of 1890 reprint of 1875 2nd edition.)

Darwin, F., editor 1887 *Life and Letters of Charles Darwin*. John Murray, London.

Evans, L. T. (1984) *J. Hist. Biol.* **17**:113.

Fisher, R. A. (1935) *Phil. Trans. R. Soc. B* **225**:197.

Fisher, R. A. (1938) *Proc. R. Soc. B* **125**:25.

Helms, J. A. and Brugmann, S. A. (2007) *Integrative and Comparative Biology,* doi:10.1093/icb/icm051 accessed 20 March 2009

Leach, C. R. and Mayo, O. (2005) *Outbreeding Mechanisms in Flowering Plants An evolutionary perspective from Darwin onwards.* J. Cramer (Schweizerbart'sche), Berlin.

Mayo, O. (2009) *Trans. R. Soc. SA* **133** (in press).

Nicholls, H. (2009) *Nature* **457**:790.

Secord, J. (1981) *Isis* **72**:163.

Wright, L. (1890) *The Illustrated Book of Poultry*. Cassell, London.

Proc. Assoc. Advmt. Anim. Breed. Genet. 18:11-13

CHARLES DARWIN IN AUSTRALIA: HIS ZOOLOGICAL OBSERVATIONS[1]

F.W. Nicholas

Faculty of Veterinary Science, University of Sydney NSW 2006

SUMMARY

Early in 1836, Charles Darwin spent two months in Australia as part of his round-the-world voyage on HMS *Beagle*. During this time he visited Sydney, travelled on horseback to Bathurst, visited Hobart and called in to King George Sound. In musing in his diary on the "strange character of the Animals of this country as compared to the rest of the World", for the first time he writes down questions that have begun to puzzle him, e.g. has there been two creators? His extensive zoological collections and observations in Australia formed part of the wealth of information he collected over several decades from throughout the world ─ information that led, in 1859, to the publication of ideas that still, 150 years later, underpin our understanding of life on earth.

INTRODUCTION

In December 1831, the Royal Navy's HMS *Beagle* left England on its second surveying voyage. On board was a young man called Charles Darwin. During the voyage, he took every opportunity to examine the geological formations and the myriad forms of plants and animals, both living and fossil, in the different parts of the world visited by the *Beagle*. The many observations made by Darwin during the voyage led him to question conventional wisdom on the origin of species, and sowed the seeds for his thinking about evolution. Combining his *Beagle* observations with masses of information collected after his return to England, Darwin gradually developed an idea as to how evolution could have occurred, and in 1859 he published his revolutionary book, "On the Origin of Species by Means of Natural Selection". By providing Darwin with the initial impetus for the development of his far-reaching ideas, the *Beagle*'s voyage has become an important event in world history.

During this voyage, the *Beagle* visited Australia between January and March 1836, giving Darwin an opportunity to examine and explore the infant colony. He was very active and observant during his visit, he collected numerous specimens of animals and rocks, and he made a number of observations that played a role in the development of his ideas on evolution.

This paper briefly reviews the most interesting zoological aspects of Darwin's visit to Australia.

THE ANT-LION EPISODE

The *Beagle*'s first Australian port of call was Sydney, where it stayed for 18 days. As was his custom whenever the *Beagle* stayed in a port for any length of time, Darwin took an inland excursion, in this case to Bathurst. He "hired a man & two horses" and followed more-or-less the present line of the Great Western Highway, over-nighting at Emu Ferry, Wentworth Falls, Blackheath, and at Wallerawang (then a sheep station, whose homestead was located near the Cox's River). As he records in his diary, "In the dusk of the evening, I took a stroll along a chain of ponds (which in this dry country represents the course of a river) & had the good fortune to see several of the famous Platypus" which reminded him of the northern hemisphere water vole.

[1] This paper is abridged from "Charles Darwin in Australia" (Anniversary edn) by F.W. and J.M. Nicholas (Cambridge University Press, Melbourne, 2008). Quotations are from Darwin's *Beagle* diary.

Earlier in the day, he had seen a kangaroo rat, occupying a similar ecological niche (with similar behaviour) to the northern hemisphere rabbit. On this and previous days, he had seen "plenty of Crows, like our jack daws, & another bird, something like the magpie." Earlier that same evening, he had observed the "conical pitfall of a Lion-Ant", noticing that it is similar to, but not the same as, northern hemisphere species. Not surprisingly, all these sightings caused him to reflect "on the strange character of the Animals of this country as compared to the rest of the World."

So it was that in the space of just a few hours in the middle of January 1836, on an isolated property in inland Australia, the 26-year-old Darwin had been confronted with three clear illustrations of the fact that similar environments in completely different parts of the world seemed to be inhabited by animals having similar adaptations, but obviously belonging to different species. In the two most striking cases, the similarly adapted animals belonged to different genera, families, orders and subclasses as well. To modern biologists, this phenomenon is called convergent evolution, and is seen as providing evidence of the power of natural selection as an adaptive force. To Darwin, it was a puzzle: "A Disbeliever in everything beyond his own reason, might exclaim, "Surely two distinct Creators must have been [at] work". In relation to the ant-lion pits, he asks, "Would any two workmen ever hit on so beautiful, so simple & yet so artificial a contrivance? I cannot think so.— The one hand has worked over the whole world.—"

These words, written in his diary on board the *Beagle* in the last week of January 1836, when Darwin had returned to Sydney, represent one of the earliest times he raises the issue of the origin of species in writing. The last sentence in the quotation above is tantalisingly ambiguous. Given that Darwin was writing the diary at least partly for the benefit of his friends and family back in England, many of whom still probably hoped that he would enter the Anglican ministry, this ambiguity may have been intentional. If by the 'one hand' Darwin means a single universal natural phenomenon, then this is one of the earliest glimpses we have, if not the earliest glimpse, of the theory that he was to use much later to explain the mechanism of evolution. It is quite possible, however, to interpret the 'one hand' as referring to God the creator who 'hit on' the one 'artificial . . . contrivance' and used it on a slightly different scale in the two different hemispheres of the world. In writing for a very mixed audience, Darwin chose words that would offend no one.

ZOOLOGICAL COLLECTIONS IN AUSTRALIA

The *Beagle* called in at Hobart and King George Sound (Albany) on its way home. At both places, as well as at Sydney, Darwin and/or his servant Syms Covington collected many natural history specimens.

Sydney. In Sydney, they collected an oyster, a mud whelk, several air breathers, a sand snail, and a trochid or top shell, plus a crab, a snake and various frogs and lizards (all of which had been described previously). The sole mammal amongst the Sydney specimens, a mouse, turned out to be previously undescribed. Initially called *Mus gouldii* when named by George Waterhouse in 1837, it was later renamed *Pseudomys gouldii*. Unfortunately, it is now extinct. Darwin and Covington between them also captured ninety-seven different insect species representing five orders in the environs of Sydney. Forty-two were previously unknown. Included among these new species were a leaf beetle (*Idiocephala darwini*), a seed bug (*Ontiscus darwini*), a gasteruptiid wasp (*Foenus darwinii*) and a bee (*Halictus darwiniellus*) that were each named after Darwin. The remaining novel insects comprised six leaf beetles (Chrysomelidae), four stink bugs (Pentatomidae), a seed bug (Lygaeidae), an assassin bug (Reduviidae), a water boatman (Corixidae), a leafhopper (Cicadellidae), a cicada (Cicadidae), a flatid planthopper (Flatidae), a froghopper or spittlebug (Cercopidae), three parasitic wasps (Chalcididae), an encyrtid (Encyrtidae), five eucaratids (Eucharitidae), a eulophid (Eulophidae), four seed chalcids (Eurytomidae), five Lamprotatidae, and one torymid (Torymidae).

Hobart. Here Darwin observed at close quarters a blotched blue-tongued lizard, *Tiliqua nigrolutea*, and what turned out to be a venomous snake, either a tiger snake (*Notechis ater*) or a copperhead (*Austrelaps superba*), which he thought was harmless.

What would have happened to the history of biological science if the young Charles Darwin had died of a snake bite in Hobart in 1836?

Also collected were shells, including rock barnacles, mesodesma (a bivalve), a whelk, an amber shell, and some bulimoid land shells, periwinkles, top shells, and air breathers. He also collected some free-living flatworms or planaria (later named *Planaria tasmaniana*) which he kept alive on the *Beagle* for the next eight weeks, collecting information on their reaction to light and their amazing regenerative powers. Among the Hobart collection, there were also at least 119 species of insects, 63 of which were previously unknown. Included among these new species were dung beetles, leaf beetles, ladybird beetles, weevils, ptinid beetles and parasitic wasps, together with a new water scavenger beetle, a new spider beetle and a new bee. Of particular interest was one species of dung beetle that had adapted from kangaroo dung to cattle dung in only 33 years.

King George Sound. One of the most important finds in the environs of Albany was another previously unknown species of native Australian rodent: the bush rat, *Rattus fuscipes*. This animal is unusual in that it is not a marsupial, but it is a native of Australia. The species inhabits a narrow coastal strip in the south-western corner of Western Australia and much of the eastern seaboard of Australia. Another interesting find was a southern frog, *Crinia georgiana*. Ten species of fish were collected, two of which were new to science: the Longhead Flathead (*Leviprora inops*), and the Common Jack Mackerel (*Trachurus declivis*). Also collected were an air-breathing limpet, a nerite, a littorinid, a periwinkle, a physa from a freshwater lake (Lake Seppings), several bulimoid land snails including two species from Bald Head, and some barnacles. The insect haul comprised at least sixty-six species, including forty-eight that were previously unknown. Six of the new species were subsequently named after their discoverer: two planthoppers (*Haplodelphax darwini* and *Alleloplasis darwini*), a predaceous diving beetle (*Hydroporus darwinii*), a small-headed fly (*Ogcodes darwini*; later renamed *Ogcodes basalis*), a seed bug (*Ontiscus darwini*) and a parasitic wasp (*Anipo darwini*; later renamed *Ipoella darwini*).

CONCLUSION

There was no Eureka moment during Darwin's visit to Australia. Instead, what he saw in Australia, and his subsequent continual gathering of information from Australia, contributed to the wealth of evidence he assembled from around the world showing that (a) species have evolved and (b) evolution can be explained by a combination of natural selection, correlated responses to selection, and chance events, all operating on the vast store of heritable variation that exists in nature.

ACKNOWLEDGMENTS

Readers are directed to the acknowledgements in "Charles Darwin in Australia" for a complete list of people who provided invaluable help in enabling the story of Darwin's visit to Australia to be compiled.

THE GENETICS OF SEXUAL DIMORPHISM IN SHEEP

B.W. Gudex[1], C. Gondro[1], K. Marshall[2] and J.H.J. van der Werf[1]

[1]School of Environmental and Rural Science, University of New England, Armidale, NSW 2351
[2]The International Livestock Research Institute, PO BOX 30709-00100, Nairobi, Kenya

SUMMARY

This study estimated the magnitude of sexual dimorphism, i.e. sex specific heritabilities and correlations between ewes and rams for weaning, post weaning, yearling and hogget weights in 3 sheep breeds (Coopworth, Poll Dorset and White Suffolk). Rams were heavier than ewes at all ages, but higher heritability estimates were observed for ewes, although few of these differences were significant. Genetic correlations between ewes and rams varied from 0.97 to 0.59 and decreased with increasing age at measurement.

INTRODUCTION

While it is well known that rams are larger than ewes (sexual dimorphism), there has been little effort to quantify the genetic factors that contribute to this phenomenon and how sexual dimorphism may be utilised in breeding programs. Unequal heritabilities for traits measured on ewes and rams and/or correlations between the sexes of less than unity could have implications for the accurate estimation of breeding values and the optimal multiple trait selection response in each sex. Moreover, sexual dimorphism could potentially be exploited in breeding programs because it may be beneficial to divergent selection between sexes within breeds for growth and maternal traits.

MATERIALS AND METHODS

Data. Pedigree and performance data for the Coopworth, Poll Dorset and White Suffolk breeds were extracted from the Sheep Genetics database. This database consists of pedigree and performance records from Australian and New Zealand studs and is used for genetic evaluation purposes. Data for weaning (wwt - 40 to 120 days), post weaning (pwt - 80 to 340 days), yearling (ywt - 290 to 430 days) and hogget (hwt - 410 to 550 days) weights were extracted. Analysis of adult weights is not reported due to insufficient data. Only data that met the following criteria was used: 1) dates of birth and measurement were recorded, 2) both sire and dam were known, 3) the sex of the animal was known, 4) the animal resulted from natural conception (not artificial insemination or embryo transfer) and 5) the contemporary group (based on flock, birth year and trait management group) the animal belonged to had more than 10 records from the same sex recorded. The number of records used for analysis for each breed and trait is shown in table 1.

Table 1: Size of the Coopworth, Poll Dorset and White Suffolk datasets for each sex

	Coopworth		Poll Dorset		White Suffolk	
	ram	ewe	ram	ewe	ram	ewe
wwt	17271	18719	51956	48413	50935	51415
pwt	4409	4070	56326	38737	47495	38549
ywt	9959	9673	44379	30165	25053	19104
hwt	433	5357	19341	10164	7290	3924

Analysis. The mixed model analysis was carried out using the ASReml program (Gilmour *et al.* 2006). All models included the interaction between birth and rearing types, the age of the dam (years), the age of the animal at the time of measurement (days) and contemporary group as fixed factors. These fixed factors are the same as those used in the OVIS program used by Sheep Genetics for the genetic evaluation of industry studs (Brown *et al.* 2000). Random effects fitted in the model included additive genetic and, when significant, maternal additive genetic and maternal permanent environmental effects. The significance of additional random effects was established using Log Likelihood ratio tests. In the across sex multitrait analyses, the residual covariance was set to zero as the same animal could not be represented in both sexes.

The heritability (h^2), phenotypic standard deviation (σ_p) and genetic correlation between rams and ewes (rg) obtained from this analysis, along with selection intensities (i) for each breed and sex that were obtained from 10 year averages (1999-2008) in the Sheep Genetics Database (Sheep Genetics, *unpublished data*, 15/5/2009), were used to predict the response to mass selection in each sex (R) using the following equation from Cheverud *et al.* (1985):

$$R_{Sex\ A} = (0.5)\,(h^2_{Sex\ A}\ \sigma_{pSex\ A}\ i_{Sex\ A} + h_{Sex\ A}\ h_{Sex\ B}\ rg\ \sigma_{p\ Sex\ A}\ i_{Sex\ B})$$

Using this equation, the response of sexual dimorphism for a trait can be defined as the difference between the responses obtained for ewes and rams (Cheverud *et al.* 1985).

RESULTS AND DISCUSSION

The sex specific heritabilities for each of the 3 breeds and at each of the 4 ages are summarised in Table 2. Although there is a consistent trend in all analyses that ewes have larger heritability estimates than rams, the difference was only significant (95% confidence interval) in 2 out of 13 analyses. These findings are consistent with those found in literature (e.g. Pattie 1965, Baker *et al.* 1979). Comparison of variance components indicated that ewes had relatively lower residual and higher additive genetic variances than rams. Rams had significantly more phenotypic variation (Vp) than ewes, which was partly due to strong correlations between phenotypic variation and weight.

The genetic (r_g) correlations between the sexes for each of the 3 breeds and at each of the 4 ages are summarised in Table 3. The genetic correlation estimates between the sexes were high and varied between 0.97 at weaning in the Poll Dorset and White Suffolk breeds to 0.59 at 1 year of age in the White Suffolk breed. Estimates for the Coopworth breed were contained within this range. There was a noticeable decrease in the across sex genetic correlations between weaning and post weaning for the Poll Dorset and White Suffolk breeds and between post weaning and yearling for the Coopworth breed. This decrease could be possibly attributed to the fact that ram lambs reach sexual maturity before ewe lambs as reported in Taylor (1968).

Between sex genetic correlations of less than 1 allow response to selection for sexual dimorphism and examples of the sex specific responses are given in Table 2. Divergent selection objectives could be used to optimise selection for the sex specific roles present in a self replacing production system. An example of this would be to have a higher selection emphasis on growth in rams than ewes, thus producing male progeny that reach sale weight earlier and restrict or reducing mature ewe weight. Mature weight has a correlated influence on feed intake and the feed intake of the ewe can comprise over 80% of the total feed costs in a production system (Thompson, 1991). The responses shown in Table 2 are merely illustrative as they are based on single trait mass selection, whereas usually responses are optimized in a multiple trait objective. However, these results do suggest that a reasonable amount of response for sexual dimorphism is possible.

There was little variation in maternal heritability (m^2) between rams and ewes with estimates larger in rams in 7 out of the 9 analyses where this effect was estimated, but only on one occasion was this significant. No identifiable trend in maternal permanent environmental effects (pe) was found. The lack of variation in the maternal heritability and permanent environmental effects

between the sexes is not surprising and indicate that the differences between dams affect ewe and ram lambs equally. For the analyses where the maternal additive genetic and maternal permanent environmental effects were estimable, the correlations between the sexes were high and positive.

Table 2: Phenotypic variance (Vp), heritability (h^2), maternal heritability (m^2), maternal permanent environmental effect (pe) and the predicted response to selection for each trait in ewes or rams.

Trait	Breed	Sex	Vp	h^2	m^2	pe	Response
wwt	Coopworth	ram	20.12[a]	0.308[a]	0.181[a]	0.119[a]	1.85
		ewe	15.94[b]	0.394[a]	0.134[a]	0.122[a]	1.67
	Poll Dorset	ram	30.94[a]	0.280[a]	0.138[a]	0.131[a]	3.57
		ewe	24.09[b]	0.338[a]	0.148[a]	0.128[a]	2.72
	White Suffolk	ram	31.71[a]	0.506[a]	0.265[a]	0.130[a]	2.35
		ewe	23.83[b]	0.517[a]	0.191[b]	0.131[a]	1.99
pwt	Coopworth	ram	22.25[a]	0.272[a]	0.220[a]		1.72
		ewe	17.06[b]	0.284[a]	0.104[a]		1.47
	Poll Dorset	ram	52.36[a]	0.298[a]	0.189[a]	0.092[a]	2.13
		ewe	30.98[b]	0.447[b]	0.160[a]	0.086[a]	1.71
	White Suffolk	ram	43.97[a]	0.416[a]	0.185[a]	0.121[a]	2.60
		ewe	27.62[b]	0.482[a]	0.210[a]	0.105[a]	1.75
ywt	Coopworth	ram	26.99[a]	0.323[a]	0.198[a]		2.00
		ewe	22.89[b]	0.364[a]	0.113[a]		1.92
	Poll Dorset	ram	62.68[a]	0.295[a]	0.236[a]		2.68
		ewe	31.17[b]	0.350[a]	0.220[a]		1.74
	White Suffolk	ram	49.31[a]	0.285[a]	0.195[a]	0.072[a]	2.25
		ewe	32.87[b]	0.444[b]	0.133[a]	0.125[b]	1.79
hwt	Coopworth	ram	60.75[a]	0.120[a]			2.57
		ewe	31.15[b]	0.411[a]			2.44
	Poll Dorset	ram	76.14[a]	0.324[a]			3.42
		ewe	30.38[b]	0.368[a]			2.23
	White Suffolk	ram	56.16[a]	0.249[a]			2.08
		ewe	33.37[b]	0.344[a]			1.69

Identical superscripts indicate that the sexes are not significantly different within trait (95% confidence interval) from each other.

Table 3: Genotypic correlations between ewes and rams

	Coopworth		Poll Dorset		White Suffolk	
wwt	0.93	±0.03	0.97	±0.01	0.97	±0.01
pwt	0.94	±0.09	0.82	±0.03	0.80	±0.03
ywt	0.84	±0.06	0.74	±0.04	0.59	±0.06
hwt	0.78	±0.52	0.80	±0.06	0.75	±0.16

Corrected weights for each sex at each of the 4 ages were obtained from the ASReml program (Gilmour *et al.* 2006) and are shown in figure 1. In each of the 3 breeds and at each of the 4 ages, rams were significantly (95% confidence interval) heavier than ewes and the degree of sexual dimorphism (ram weight/ewe weight) increased with age, from around 1.10 at weaning in all 3 breeds to 1.26 in the Coopworth breed, 1.27 in the White Suffolk breed and 1.30 in the Poll Dorset

breed in the hogget age group. Based on this evidence, it appears that not only do rams grow faster than ewes up to puberty but they continue to do so until hogget age.

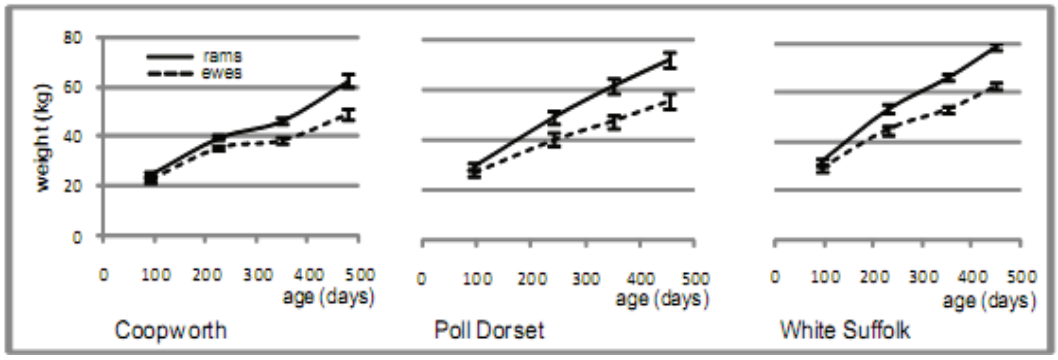


Figure 1: Growth curves for ewes and rams. Error bars represent a 95% confidence interval for each value.

CONCLUSIONS

Sexual dimorphism (rams larger than ewes) clearly exists in sheep and could be increased by selection. Weight traits measured on ewes had slightly higher heritability estimates compared with rams, but lower phenotypic variation, and the genetic correlations between the sexes were significantly less than 1. The combination of these results potentially have implications for determining the optimal models required for the accurate estimation of breeding values and in determining the optimal selection response in each sex, particularly in a scenario where one sex is recorded preferentially. Sexual dimorphism could also be combined with the sex specific roles found in self replacing lamb production systems to reduce the total feed requirements of the production system.

ACKNOWLEDGMENTS

This research was supported by Meat and Wool New Zealand (MWNZ), Meat and Livestock Australia (MLA) and the University of New England (UNE). The assistance of Arthur Gilmour is also gratefully acknowledged.

REFERENCES

Baker, R. L., Clarke, J. N., Carter, A. H., and Diprose, G. D. (1979) *N.Z. J. Ag. Res.* **22**:9

Brown, D.J., Tier, B., Reverter, A., Banks, R. and Graser, H.U. (2000) *Wool Tech. Sheep Breed.* **48**:285

Cheverud, J.M., Dow, M.M. and Leutenegger, W. (1985) *Evolution* **39**:1335

Gilmour, A.R., Gogel, B.J., Cullis, B.R. and Thompson, R. (2006). ASReml User Guide Release 2.0 VSN International Ltd, Hemel Hempsteed, HP11ES, UK.

Pattie, W. A. (1965). *Aus. J. Exp. Ag. Anmi. Husb.* 5:353

Taylor, ST C.S. (1968) *Anim. Prod.* **10**:157

Thompson, J.M. (1991) In "Australian Sheep and Wool Handbook", p.243, editor D. Cottle, Inkata Press, Melbourne.

SEXUAL DIMORPHISM OF GENES REGULATING SKELETAL MUSCLE GROWTH IN BALI CATTLE

Nattrass G.S.[1], Dahlanuddin[2], Poppi, D.P.[3] and Quigley S.P.[3]

[1]SARDI – Livestock and Farming Systems, Roseworthy, SA 5371
[2]University of Mataram, Faculty of Animal Science, NTB, Indonesia 83125
[3]University of Queensland – The School of Animal Studies, Gatton, QLD 4343

SUMMARY

Quantitative reverse transcriptase polymerase chain reaction (qRT-PCR) was used to measure the mRNA levels of a selection of genes involved in skeletal muscle growth in Bali cattle. The findings of this study provide preliminary evidence for differences in mRNA levels in the skeletal muscle of male and female Bali cattle at approximately 100 kg, 190 kg and 250 kg live weight.

INTRODUCTION

Domestic beef consumption is increasing at approximately 4% per annum in Indonesia. This increased demand is unable to be met from local supply alone under the prevailing beef cattle production systems. Bali cattle (*Bos sondaicus*) are the predominant cattle species across the eastern islands of Indonesia. Anecdotal evidence suggests that Bali cattle are sexually dimorphic in growth rate, mature size and coat colour. However, little work has been conducted on the growth, physiology and underlying biology of Bali cattle under controlled experimental conditions and this information will be useful in improving Bali cattle production. The expression of genes implicated in the regulation of skeletal muscle cell activity, metabolism and protein accretion was investigated in the *Semitendinosus* (ST) muscle of male and female Bali cattle at approximately 100 kg, 190 kg and 250 kg average live weight.

MATERIALS AND METHODS

Animals and muscle sampling. All procedures were conducted in accordance with the guidelines of the Australian Code of Practice for the Care and Use of Animals for Scientific Purposes and were reviewed and approved by the University of Queensland Animal Ethics Committee. Bali cattle (n=12 males and n=6 females) approximately six months of age and 72 ± 2 kg in live weight were located at the University of Mataram research and teaching farm, Indonesia. The animals were maintained in individual pens throughout the study and were offered a diet consisting of sesbania (*Sesbania grandiflora*) *ad libitum* from six to 24 months of age and rice bran (5 g DM/kg live weight/day) plus sesbania *ad libitum* from 24 to 28 months of age. Biopsies were collected from the ST muscle of each animal under local anaesthesia (20 mg/mL Lignocaine hydrochloride) as the animals were approximately 100 kg, 190 kg and 250 kg in live weight (approximately 10, 18 and 27 months of age, respectively). The biopsy was immediately placed in RNAlater (Ambion) and stored at 4^{o}C for 24 h prior to storage at -20^{o}C for up to 15 months.

RNA purification and cDNA synthesis. Total RNA was isolated with the MagMAX kit and treated with DNaseI (Ambion). Six hundred nanograms of total RNA from 38 ST muscle samples was reverse transcribed using an oligodT_{20} primer (2.5 μM) and the Omniscript RT kit (Qiagen). The 38 samples used in the subsequent real-time PCR studies comprised n=16 (n=10 males; n=6 females) at 100 kg, n=15 (n=10 males; n=5 females) at 190 kg and n=7 (n=4 males; n=3 females) at 250 kg live weight.

Pre-amplification and real-time PCR. A volume of cDNA from each sample, equivalent to 12.5 ng of total RNA, was pre-amplified for 14 cycles using the 2x PreAmp kit (Applied BioSystems) and a primer pool (100 nM) containing all the primers shown in Table 1. The pre-amplified cDNA was diluted 5-fold in TE buffer (10mM Tris (pH8.0) and 0.1mM EDTA). Seven consecutive 2-fold dilutions were performed on a pooled batch of pre-amplified cDNA; this dilution series was used to generate the standard curves for each gene assay. Real-time PCR was performed with Fast-Start Probe Master (Roche) and 2 μM of Universal Probe Library (UPL; Roche) probe as described in Table 1. Real-time PCR was performed with a 48.48 Dynamic Array (Fluidigm) on the BioMark system as follows: 95°C/10 min (1 cycle), 95°C/15 sec, 72°C/5 sec and 60°C/1 min (40 cycles). The BioMark system allowed the simultaneous measurement of mRNA abundance of 16 genes across 48 samples (unknowns, standards, calibrators and NTC) in triplicate, in real-time.

Data normalisation and statistical analysis. The geometric average of three stably expressed reference genes (EEF1A2, RPL19 & UCHL15) was used to normalise the real-time PCR data. Statistical assessments were performed with a General Linear Model (GLM; SAS Institute v9.1).

Table 1: Oligonucleotides designed against bovine reference sequences (RefSeqs) and GenBank cDNAs, and the corresponding UPL probe number used to detect each gene.

Gene	Forward primer (5' – 3')	Reverse primer (5' – 3')	UPL Probe	Reference sequence
ANDR	ggaacttgatcgtatcattgcat	gaactgatgcagctctcgtg	15	XM_001253942
EEF1A2	acacgccccacagacaag	ggatccctgtctccactcg	59	NM_001037464
ESR1	ccaactcctcctcatcctctc	agagaggcaccacgttcttg	124	NM_001001443
FST	cgtgctgctggaagtgaag	aggaaagctgtagtcctggtctt	101	NM_175801
GDF8	tcttgctgtaaccttcccaga	tggtgtgtctgttaccttgacttc	2	NM_001001525
GHR	gatctctggcagctgctgtt	ctctgagatgctctgacaagga	59	NM_176608
GRB10	gccatgctgtcccacttc	cattgccacaagggagttct	109	XM_602256
HMBS	ttaccaaggagctggagcat	gcgaatgaacaactaggtcca	55	NM_001046207
IGF1	tgcttttgtgatttcttgaagc	gcacagggccagatagaaga	101	NM_001077828
IGF1R	cctctgtgaacccggagtat	ggctcatggtgatcttctcc	10	XM_606794
MyHC2X	actgaggaggaccgcaaga	tgcttgcagtttatccacca	1	AB059399
RPL19	cgaatgcccgagaaggta	ccttcaggtacaggctgtgat	101	AY158223
SF3A1	cagagcgatgacgaggtgta	cacaccaaagatgtctgtacgc	108	NM_001081510
UCHL15	agcagagttacaaagacaacttgc	gcatctgatttttggcaacc	122	NM_174481

RESULTS AND DISCUSSION

Live weight change. There was no significant difference in live weight change between male (0.34 kg/day) and female (0.29 kg/day) Bali cattle between six and 12 months of age. Male Bali cattle (0.41 kg/day) grew faster than female Bali cattle (0.26 kg/day) between 12 and 27 months of age ($P<0.01$).

Relative mRNA levels. Follistatin (FST)and myostatin (GDF8) mRNA levels were significantly lower in the ST muscle of Bali cattle at 190 kg live weight compared with 100 kg live weight ($P<0.05$; Table 2). The negative regulation that GDF8, the gene responsible for double muscling in the Belgium Blue and Piedmontese cattle breeds, exerts on the proliferation of satellite cells during postnatal growth is antagonised by FST. The observed reduction in FST and GDF8 mRNA levels when the animals were approximately 190 and 250 kg may coincide with decreased number and activity of satellite cells, typically observed in mammals during the latter stages of postnatal growth and into adult life. Complicating the interpretation of these findings are the existence of

GDF8 (Jeanplong and McMahon 2005) and FST (Shimasaki *et al.* 1988) splice variants, which introduce alternate 3' exons that change the C-terminus of both proteins. The relative abundance of these splice variants is likely to alter their biological functions in skeletal muscle, and may also change the degree of antagonism that FST exerts on GDF8.

The mRNA level of the IGF type 1 receptor (IGF1R) was higher in males than females ($P<0.01$; Table 3). In the present study, as the live weight of animals increased, the average daily gain was higher in males compared with females and this may be attributed in part to the higher IGF1R mRNA levels and their role in mediating the action of circulating Insulin-like growth factors (IGFs) on skeletal muscle protein accretion.

Hydroxymethylbilane synthase (HMBS) mRNA levels were higher in males than females ($P<0.01$; Table 3) and as the animals reached heavier live weights ($P<0.01$; Table 2). Myosin heavy chain 2X (MyHC2X) mRNA levels remained constant in females but there was a 30% decline in males between 190 and 250 kg live weight ($P<0.01$; Figure 1). Hydroxymethylbilane synthase catalyses the third enzymatic step in the heme biosynthetic pathway, which regulates haemoglobin and myoglobin production. The changes observed for HMBS may indirectly indicate that male Bali cattle have a greater proportion of myoglobin-rich myofibres than females and that as both sexes get closer to mature weight there is a shift towards a more oxidative metabolism in the myofibres. This hypothesis is supported by the observation that males at 250 kg had reduced MyHC2X mRNA levels, which suggests a reduction in glycolytic metabolism. In the *Longissimus dorsi* (LD) muscle of sheep, MyHC2A-positive myofibres increased in abundance with age while MyHC2X-positive myofibres decreased over the same period (Greenwood *et al*. 2007). In addition, the LD muscle of bulls contained more αR (MyHC2A-positive) myofibres and less αW (MyHC2X-positive) myofibres than age-matched steers and implanted steers (Clancy et al., 1986), supporting a sex effect in the regulation of muscle fibre type and muscle growth. We intend to measure mRNA levels of MyHC2A and MyHCslow in these samples to confirm our hypothesis.

No significant differences were observed in mRNA levels of the androgen (ANDR), estrogen (ESR1) or growth hormone (GHR) receptors, IGF-1 (Insulin-like growth factor 1) and growth-factor receptor bound-10 (GRB10) between the sexes or at the different live weights.

Table 2: Relative mRNA levels of follistatin (FST), myostatin (GDF8) and hydroxymethylbilane synthase (HMBS) in the *Semitendinosus* muscle of Bali cattle at increasing live weights.

Gene	Live weight (kg)		
	100	190	250
FST	1.00 ± 0.08^{a}	0.73 ± 0.08^{b}	0.76 ± 0.12^{ab}
GDF8	1.00 ± 0.06^{a}	0.82 ± 0.06^{b}	0.79 ± 0.09^{ab}
HMBS	1.00 ± 0.05^{a}	1.40 ± 0.06^{b}	1.25 ± 0.08^{b}

Table 3: Relative mRNA levels of Insulin-like growth factor 1 receptor (IGF1R) and Hydroxymethylbilane synthase (HMBS) in the *Semitendinosus* muscle of female and male Bali cattle.

Gene	Sex	
	Female	Male
IGF1R	1.00 ± 0.10^{a}	1.34 ± 0.08^{b}
HMBS	1.00 ± 0.05^{a}	1.20 ± 0.05^{b}

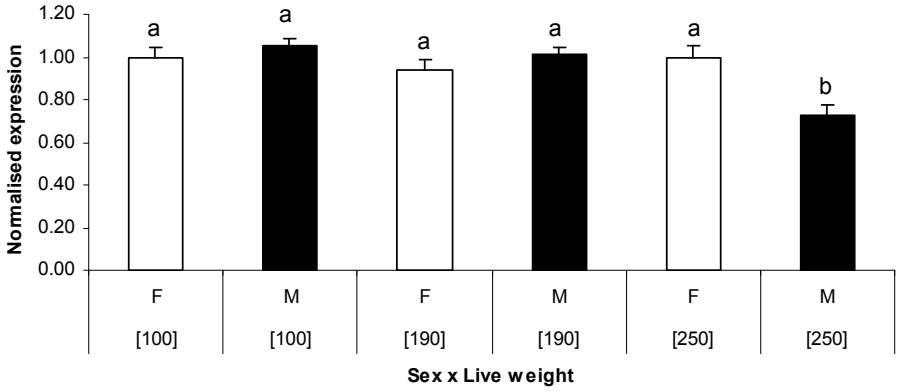


Figure 1: Interaction between sex (female, F, or male, M) and live weight (100, 190 and 250 kg) on the relative mRNA level of Myosin heavy chain 2X in the *Semitendinosus* muscle of Bali cattle.

CONCLUSIONS

The levels of IGF1R and HMBS mRNA were implicated in the regulation of sexually dimorphic growth of Bali cattle. Characterising the transcriptional profiles of alternatively spliced genes such as GDF8 and FST in greater detail will provide clearer insights into the biological functions of these genes. Our inability to definitively interpret the reduction in GDF8 and FST mRNA levels at later stages of post-natal skeletal muscle growth in Bali cattle highlights the need for further refinement of transcriptional profiling through the use of Laser Capture Microdissection (LCM). As skeletal muscle is composed of a heterogeneous mix of cells, LCM-based studies would enable the characterisation of gene expression profiles in the main cell types which constitute skeletal muscle. The information obtained from LCM would aid the interpretation of whole tissue findings. The unrivalled assay capacity of the Biomark system enabled rapid screening of large numbers of DNA samples and assays. The BioMark is well suited to undertake LCM based studies and to investigate functional associations between single nucleotide polymorphisms and mRNA levels of splice variants in nearby genes (Kwan et al., 2008).

ACKNOWLEDGMENTS

This project was funded by the Australian Centre for International Agricultural Research, the University of Queensland and SARDI. The authors gratefully acknowledge the technical assistance of Dr Muhammad Ali, Ms Baiq Susmaritha Ningsih and Mr Kasri.

REFERENCES

Clancy, M.J., Lester, J.M. and Roche, J.F. (1986) *J. Anim. Sci.* **63**: 83.

Greenwood, P.L., Harden, S. and Hopkins, D.L. (2007) *Aust. J. Exp. Agric.* **47**: 1137.

Kwan, T., Benovoy, D., Dias, C., Gurd, S., Provencher, C., Beaulieu, P., Hudson, T.J., Sladek, R. and Majewski, J. (2008) *Nat. Genet*. **40**: 224.

Shimasaki, S., Koga, M., Esch, F., Mercado, M., Cooksey, K., Koba, A. and Ling, N. (1988) *Biochem. Biophys. Res. Commun.* **152**: 717.

Jeanplong, F. and McMahon, C.D. (2005) U.S. Patent Publication Number: 20080118487.

PROGRESS IN DEVELOPMENT AND IMPLEMENTATION OF A STRATEGY FOR COMMERCIALISATION OF DNA MARKER TECHNOLOGY FOR THE AUSTRALIAN BEEF INDUSRTY

R.G. Banks[1], H. Burrow[2] and H.-U. Graser[3]

Cooperative Research Centre for Beef Genetic Technologies
[1]MLA, c/- Animal Science, University of New England, Armidale 2351
[2]Beef CRC, University of New England, Armidale, NSW 2351
[3]AGBU, University of New England, Armidale 2351

SUMMARY

This paper outlines the strategy which is being implemented to support commercialisation of DNA technologies for the Australian beef industry. The strategy includes a national database for phenotypic and genotypic information, information nucleus population for validation and research, integration of DNA results with BREEDPLAN evaluations and a close cooperation between commercial companies and other industry players. Various risks of this strategy and their control are discussed.

INTRODUCTION

Genotyping technologies have evolved rapidly over the last 10-15 years, and offer significant potential benefits for the beef (and other) industries. Australian research effort for this technology has a focus within the CRC for Beef Genetic Technologies, with significant activity within both the public and private sectors. The Beef CRC developed a strategy for the commercialisation of its DNA marker technology, in consultation with industry, research, development & extension (RDE) partners and providers, and with service companies. That strategy has four core elements:

- that information on the value of marker tests of various sorts, ranging from single genes or single nucleotide polymorphisms (SNPs) through to multi-snp panels or whole genome evaluations, should be made public,
- that industry will need an independent reference population to support the evaluation of markers,
- that a national database and matching DNA bank will be required to store phenotypes, genotypes and DNA from commercial and seedstock cattle; and
- that where possible, results of marker tests should be integrated within the existing genetic information framework provided by BREEDPLAN.

Recognising that the emerging technologies would require more than the resources of the Beef CRC alone for effective industry implementation, Meat and Livestock Australia (MLA) and Beef CRC agreed in 2008 to work together firstly to determine the feasibility of the proposed strategy, and, subject to that feasibility evaluation, on its implementation. A working party was established initially to consult on the feasibility of the strategy, and more recently to progress implementation. This paper outlines progress being achieved, plans for the immediate future, and briefly discusses some risk considerations.

FEASIBILITY ASSESSMENT

Two dimensions of feasibility assessment were explored: the technical requirements, and the implementation requirements (including longer-term governance).Technical requirements include:

- methods and skills in place to estimate the effects of markers, and
- whether or not methods and software exist for the integration of marker information into BREEDPLAN genetic evaluations.

Both questions were answered by the success of the SmartGene project (Johnston and Graser 2008), where effects of the DNA marker tests derived principally from the CRC program, were estimated for a number of breeds, and for a limited set of markers, integration into marker-enhanced EBVs was achieved (Johnston *et al.* 2008).

A continuing challenge with regard to the integration into EBVs will be the rapid evolution of the DNA tests themselves, and the format of the results. The latter is likely to vary between providers, at least in the short-term.

The SmartGene project included considerable consultation with industry and one genotyping provider, and provided a basis for addressing a second set of questions relating to implementation challenges. These implementation challenges include:

- what infrastructure is to underpin regular estimation of marker effects,
- what incentives exist, or might be established, to encourage integration of marker results with EBVs, and
- whether new governance systems might be required.

It was concluded that the core infrastructure required is of two forms: a) a database which can receive and hold data on ID, pedigree, performance and genotype, and b) suitable reference populations. The Beef CRC I&II research databases provide at least an initial version of the first requirement. It is clear that consolidation and development of the types of genotype and genotype test data that can be stored, will be needed.

Consideration of the requirements for reference population(s) is complicated by the fact that as research proceeds, the effects of or due to individual SNPs have in general become smaller, meaning in broad terms that larger numbers of performance recorded animals will be needed for "validation" or calibration of effects, and further by the fact that considerable differences may exist between populations in terms of allele effect(s) and frequencies.

Not withstanding this complication, a number of industry groups, focussed mainly amongst breed societies, are keen to explore and establish appropriate populations. The challenge will be to fund the necessary populations. The MLA Donor Company mechanism, whereby Commonwealth Government funds can match commercial funds in appropriate projects, was identified as a potential avenue for at least some of the support needed.

Consultation around the issues of engagement and governance suggested that in the short term at least, no specific new regulatory approach is required. However, ongoing consultation will be needed on data transfer and exchange protocols, and how results are reported to industry.

The overall assessment was that the strategy is technically feasible, but that new investment will be required, particularly in phenotypes, and that coordination of information flow and communications will be essential.

PLANS FOR NEXT PHASE IMPLEMENTATION AND INDUSTRY DEVELOPMENTS

From October 2008 to mid-2009, most activities have focussed on developing proposals for establishment of reference populations building on the large industry-based progeny tests in Beef CRCI and II (Upton *et al*. 2001; Burrow and Bindon 2005), and drawing on the model established in the sheep industry known as Information Nucleus (IN) (Banks *et al.* 2006). At the time of writing this paper, one breed-based project proposal for a Beef IN had been submitted to the MLA Donor Company for approval, and four others were in preparation. Each is likely to involve utilisation of 300-500 cows in a mix of stud and commercial herds, with male progeny planned to enter a research feedlot for collection of finishing and processing phase trait data, and female progeny retained for recording reproduction traits. To enhance impact on genetic progress, sires sampled will wherever possible be elite and young. In addition, there has been progress in:

- appointment of a manager for the national database, part of whose role is to consult with data and genotyping providers aiming to ensure smooth and secure data exchange;

- continuing analysis of suitable datasets to provide estimates of effects for various marker tests, and posting of the results obtained to an industry website

SmartGene and Pfizer MVP results have been posted on an Australian results website (http://www.beefcrc.com.au/Aus-Beef-DNA-results), and Igenity to follow, and a common reporting format has been developed with NBCEC in North America (www.nbcec.com).

RISK CONSIDERATIONS

The strategy proposed by the Beef CRC seeks to maximise the benefits for the industry. That strategy carries its own internal risks, faces potential external risks, and its implementation will be within the context of some wider challenges facing the industry. Each of these risks is examined briefly, together with discussion of potential responses.

The main "internal" risk is that no markers (or marker-based tests) will be discovered that offer any real value, expressed as the proportion of genetic variance explained for one or more traits in one or more breeds. Although this risk is considered small, Beef CRC is mitigating this risk by undertaking specific research aimed at understanding reasons for the failure of most existing markers to provide value across breeds, by investing in development of larger SNP panels to overcome the problem and using its international collaborations to share resources and agreeing on common approaches to overcoming problem areas..

The second internal risk is that whatever tools become available are too expensive to be funded within the financial capacity of the Australian beef industry. This reflects the relatively low margins for seedstock operations, the almost complete absence of vertical integration in the industry, and the relatively poor flow of price signals from consumer through the value chain to the breeder. Should this risk become real, industry as a whole will have a serious strategic issue, particularly if beef industries in competitor countries are able to adopt the technology. Careful assessment of the degree of market failure and potential responses using collective funding will be required.

A third internal risk is that availability of marker tests undermines confidence in BREEDPLAN. The CRC's mitigating strategy aims to inform the market by placing full results of objective assessment of marker tests into the public domain and integrating marker data into EBVs. But the second element of the strategy depends on choices made by breeders. Accordingly, clear communication about the value of integration into EBVs is essential.

The fourth internal risk is that the amount of phenotype data required for useful calibration of markers proves to be too large for industry to achieve. This risk will be mitigated through development of denser SNP panels that allow discovery of markers that are closer to the causative mutation and where effects are more consistent across breeds.

There are two external risks. The first is that commercial genotyping companies offer products that have insufficient value for industry or provide a diversity of products with a wide range of claimed attributes, causing confusion in industry. The strategy that has been developed by the working group aims to encourage independent testing of all markers by the companies with information on test results made publicly available. The commercial companies may choose not to use such an approach and hence, industry's use of the independent information will be critical to the success of this strategy.

This approach is regarded as a win-win where industry phenotyping resources are used as a pre-release screening or "validation" step, in a manner analogous to progeny testing of young dairy bulls. This approach reduces risk for both buyers and sellers, but may take some time to evolve.

A second external risk is that the joint effects of R&D provider, industry and commercial company behaviour are inadequate in providing customer protection. In that case, some form of independent regulatory approach may be required ("independent" in this sense meaning

independent of industry, with government as the obvious possibility). It would be preferable that this does not become necessary, since introduction of an external regulatory framework would introduce substantial new costs, both financial and time.

Finally, there are wider challenges facing industry, related to the effectiveness with which existing technology is used. There is evidence that the Australian beef industry makes no better use of BLUP technology than its competitors (McDonald 2008), and that this reflects less than optimal data collection coupled with lower than potential sire selection differentials (Johnston 2007). While marker technologies may assist in making genetic improvement easier, this is automatic, since that progress ultimately depends on recording and selection efficiency: if these are not optimised, marker technology could just as easily lead to reduced progress. This risk must be addressed through coordinated and effective extension and training at the seedstock and commercial levels.

The other wider risk is that the combination of accelerating genetic progress using either existing technologies, or those augmented by marker technology, coupled with the potential explosion of DNA-based products and likely new value-chain relationships, will result in a period of heightened confusion and turbulence in the market for genes, resulting in reduced investment effectiveness. In addition, it is very likely that new industry structures will emerge, potentially around co-investment in phenotyping and genotyping, and these will challenge existing ways of delivering genetic evaluation, of delivering genetic material, and of both investing in and managing the industry genetic improvement "system".

The risk of confusion is real, and would be very damaging for the Australian beef industry, since the industry does not in general have large reserves of capital, and needs to accelerate progress in both production efficiency and product quality differentiation in order to maintain or achieve higher prices than competitors in world markets. Minimising this risk from cheaper undifferentiated beef requires effective coordination of effort across all sectors of industry. Finally, the strategy must be flexible enough to allow response to the changing technology.

CONCLUSIONS

The Australian beef industry, along with others, is investing heavily in development and implementation of DNA-based technologies. The strategy developed initially by the Beef CRC for that implementation is now moving out of a feasibility evaluation phase and into implementation. The strategy calls for development and utilisation of new information and infrastructure, and its success will depend in part of wide and effective coordination across industry coupled with capacity for rapid response to evolving technology. The risks around the implementation are manageable but significant. The Australian beef industry seems inevitably to be entering a phase of rapid and potentially valuable evolution, but at the same time one which could lead to great loss if these risks are not managed effectively.

REFERENCES

Banks R. G, van der Werf J., Gibson J.P. (2006) *8th World Cong. Genet. App. Livest. Prod.* **30 :**12.

Burrow H.M. and Bindon B.M. (2005) *Aust. J. Exp. Agric.,***45:** 941.

Johnston, D.J. (2007) *Proc. Assoc. Advmt. Anim. Breed. Genet*. **17**:8

Johnston, D.J. and Graser, H.-U. (2008) SmartGene for Beef, accessed online at http://agbu.une.edu.au/smartgene.php

McDonald (2008) International Beef Genetic Trend Benchmarking (MLA Report)

Upton W, Burrow HM, Dundon A, Robinson DL, Farrell E (2001) *Aust. J. Exp. Agric.* **41**:943.

GENOMIC SELECTION BASED ON DENSE GENOTYPES INFERRED FROM SPARSE GENOTYPES

M. E. Goddard[1,2] and B.J. Hayes[1]

[1] Department of Primary Industries and [2] University of Melbourne

SUMMARY

Genomic breeding values (GEBVs) predicted from dense marker panels are now being used in the dairy industry for bull selection. However the cost of genotyping these dense marker panels is too high to justify genomic selection for dairy cows, beef bulls and rams for meat and wool production. However if potential selection candidates were genotyped for a standard panel of less expensive sparse single nucleotide polymorphisms (SNPs) and key ancestors within their breed were genotyped for a dense panel of markers, it would be possible to trace the chromosome segments in the selection candidates back to the key ancestors and thus infer their genotypes at all markers assayed on the key ancestors. In this way genomic selection could be practiced using a large number of markers while the cost of genotyping was kept low by the widespread use of a standard SNP chip.

INTRODUCTION

Despite great advances in molecular genetics in recent years, we have identified very few of the genes that cause variation in economically important traits. These genes are known as quantitative trait loci or QTL, because most of the traits of importance are quantitative traits controlled by many genes and by environmental factors. An important recent advance has been the availability of panels of markers consisting of 10,000s to 1,000,000s of single nucleotide polymorphism (SNP) markers. This has made practicable the use of genomic selection as described by Meuwissen *et al.* (2001). Genomic selection refers to the use of a dense panel of genetic markers covering the whole genome to estimate the breeding value of selection candidates (Meuwissen et al 2001). Because the markers are closely spaced, a QTL located anywhere in the genome will be in linkage disequilibrium with at least one marker. Using the Illumina panel of approximately 50,000 SNPs, VanRaden *et al*. (2009) could estimate the breeding value for milk production traits of Holstein bull calves with an accuracy of approximately 0.8. As a result, most developed countries with large dairy industries are now implementing genomic evaluation to calculate estimated breeding values (EBVs). It is anticipated that this trend will spread to beef cattle and to other species because genomic evaluations potentially allow animals to be selected for any trait at birth with an accuracy approaching that of a progeny test. Therefore genomic selection should lead to great increases in the rate of genetic gain in livestock (Schaeffer 2006).

However, there are at least two problems with the implementation of genomic selection. Firstly, the 'SNP chips' used for genotyping SNPs are still too expensive (~$A400) to use on most animals. Secondly, there are likely to be many new SNP chips introduced to the market and many other markers, that are not on any SNP chip, discovered to be useful for selection for some trait. To estimate the effects of SNPs on a trait requires a large dataset of animals that have been measured for the trait and genotyped for the markers (a reference population). Therefore, each new panel of SNPs will have to be genotyped across a reference population of tens of thousands of animals. Even then, many animals will have been genotyped with the old SNP chip and will not be able to benefit from the new SNPs. It could even be necessary to genotype animals with multiple SNP chips to obtain EBVs with the maximum accuracy.

In this paper we outline an alternative approach that uses a standard SNP chip of moderate size and cost and infers the genotypes of an animal at all known markers so that they can be used for

genomic selection. Firstly, we consider the number of SNPs needed for genomic selection, and then we describe the imputation of SNP genotypes at ungenotyped SNPs.

NUMBER OF MARKERS NEEDED FOR GENOMIC SELECTION

There are two approaches to genomic selection that can be distinguished. In the first, the same panel of markers is used to genotype selection candidates as was used to discover which markers are useful for predicting breeding value. For instance, the Illumina 50k panel might be used to genotype a reference population and from this a prediction equation is derived that predicts breeding value from 50,000 SNP genotypes. Selection candidates are then genotyped for the same 50,000 SNPs and the prediction equation used. In the second approach, an experiment genotypes a reference population for 50,000 SNPs but the data is used to find a subset of the SNPs that can be used to predict breeding value. The selection candidates are genotyped for only a subset of markers. This raises two questions: How many markers are needed in the discovery experiment and how many of these are needed in the final prediction equation?

In the discovery experiment we do not know where the QTL are so we need to cover the whole genome. The density of SNPs needed depends on the extent of LD. The extent of LD in turn depends on the effective population size (N_e). In a breed with a small N_e, such as Holstein (N_e<100), LD extents over considerable distances and so 50,000 markers may be enough. However, if many breeds are used, the LD is only consistent between breeds of *Bos taurus* cattle if markers are <10kb apart (de Roos et al 2008), implying the need for >300,000 markers. Even for a particular breed, the number of markers needed is not an absolute number because the accuracy of predicting breeding value increases slowly as the number of markers increases.

The number of markers needed to predict breeding value in selection candidates should depend on the number of QTL affecting the trait. If there were only a small number of QTL one would expect that a few SNPs near each QTL would be enough. Unfortunately we have not identified the QTL so we cannot answer this question directly. The methods that have been tried for genomic selection vary in the assumptions they make about the number of QTL. The method called BLUP by Meuwissen et al (2001) assumes that all markers have an effect and that these effects are drawn from a single normal distribution. By contrast a method called BayesB assumes that only a fraction of the markers are needed and the others have zero effect. When compared on real data these methods do almost equally well suggesting that there are a large number of QTL affecting most traits. However, it may be possible to choose a small panel of markers that does almost as well as the full 50,000 SNPs. Even a slight loss of accuracy is undesirable but might be acceptable if the cost of genotyping the small panel was much less than the cost of the 50,000 SNP panel. For instance, a panel of 384 SNPs might give acceptable accuracy for one trait but, since many traits are important to most breeders, this still implies a 'small' panel of >1500 SNPs.

The accuracy of genomic selection improves as the number of animals in the reference population (animals with both genotypes and phenotypes) increases, and reference populations in the order of 10s of thousands of animals are required to subsequently predict GEBV with accuracy say greater than 0.7. The only practical way to obtain a very large reference population is to include the animals on which the SNP panel is used commercially. For instance, animals that are selection candidates may, in the future, have phenotypes recorded and could be added to the reference population so that the effects of SNPs on traits recorded could be continually re-estimated and updated. This is analogous to updating the EBV of a sire as his number of recorded offspring increases. However, if the selection candidates are genotyped for only a small panel, it is only the effects of these SNPs that can be re-estimated. Conversely, if the selection candidates had been genotyped for all SNPs, the choice of the best SNPs could be continually improved.

Therefore, we would like to use a small panel of SNPs to minimise the cost of genotyping, but we would like to use all available SNPs to maximise the accuracy of estimated breeding values.

The next section describes a proposal to achieve this. The method put forward has been described by Goddard and Hayes (2008), Goddard (2008) and Habier *et al*. (2009).

INFERRING SNP GENOTYPES FROM A STANDARD PANEL OF SNPS

In most breeds of livestock Ne is quite small (100-200). This means that all the animals of the breed trace back to the equivalent of 100-200 ancestors in each generation. In practice, we find that there are a small number of ancestors that contribute disproportionally to the modern breed. For instance, in Holsteins there are 6 sires who each contribute over 12% of the genes in modern cattle. These 'key ancestors' are only a few generations removed from the current population. Consequently, modern animals inherit large chromosome segments from key ancestors without any recombination. If a key ancestor is t generations removed from a modern animal, the average length of a chromosome segment inherited from the key ancestor is 1/(t+1) M. For instance, if the key ancestor is 3 generation ago, the average chromosome segment inherited is 25 cM. If this segment contained ~7 SNPs, it should be possible to trace the segment from the modern animal back to the key ancestor. If we know the alleles that key ancestor carried at other markers on this 25 cM segment, then we know that the modern animal must have inherited these alleles along with the rest of the segment. Thus we can infer the genotype of a modern animal at all markers that have been genotyped on the key ancestors provided we have genotyped enough markers on the modern animal to trace his or her chromosome segments back to a key ancestor. There is an additional requirement – the genotypes of the key ancestors must be phased so that we know which alleles are carried on the same chromosome. This can be achieved by a linkage analysis using known relatives or by a non-pedigree method relying on LD between markers such as fastPhase (Scheet and Stephens 2006).

KEY ANCESTORS

What is the smallest set of animals that can be genotyped with the dense SNP panel? We have called these animals key ancestors to indicate a set of animals from which most of the genes in modern animals are descended. A method to find the optimum set of key ancestors is now described. Suppose that you wanted to predict the mean breeding value (m) of the modern population from knowledge of the breeding values of a set of key ancestors (**g**). A regression equation could be derived, m = **b'g,** where $\mathbf{b} = \mathbf{A}^{-1}\mathbf{c}$, **A** = numerator relationship matrix among the key ancestors, **c**= vector of relationships between key ancestors and the modern population. However, m could also be predicted from a weighted average of the breeding values of the key ancestors where the weights are the proportion of genes contributed directly from a key ancestor to the modern population ie genes not contributed through another key ancestor. In fact, **b'g** is exactly this weighted average and so the sum of the elements of **b** (**1'b**) is the proportion of genes in the modern population that can be traced to one or other of the key ancestors. Therefore the optimum set of key ancestors is the one that maximizes **1'b.** A good approximation to this optimum can be found by the equivalent of forward regression. That is, at each iteration the animal that would increase **1'b** the most is added to the list of key ancestors.

This choice of key ancestors maximises the proportion of genes in the modern population derived from the key ancestors but it ignores the problem of phasing their genotypes. If phasing is to be carried out using linkage analysis, then the list of key ancestors needs to contain related animals. The ability to trace chromosome segments from modern animals to key ancestors will decrease as the number of generations between them increases (Habier *et al*. 2009). This problem can be overcome in two ways. Firstly, one could use denser markers for the sparse SNP panel. Habier *et al.* (2009) used only one SNP per 10 cM so it is not surprising that recombinations eroded the ability to trace chromosome segments. Secondly, one could continually add animals, that occur in the pedigree between the key ancestors and the modern animals, to the list of densely

genotyped animals. Habier *et al.* (2009) found that densely genotyping all parents each generation overcame the problem. However, if denser markers were used to genotype key ancestors it would not be necessary to densely genotype so many animals. For instance in the dairy industry, one might only densely genotype AI sires. Nor is it necessary to genotype these intermediate ancestors for all SNPs in the dense panel used on key ancestors. The intermediate ancestors merely provide a link between the modern animals and the key ancestors. For instance, the key ancestors might be genotyped for all known SNPs whereas the intermediate ancestors are genotyped for 50,000 SNPs.

IN SILICO GENOME SEQUENCING

The ultimate density of markers might be provided by full genome sequencing. As the cost of sequencing drops this will become possible on a list of key ancestors. Then by genotyping modern animals for a panel of sparse markers it would be possible to infer the full genome sequence on each modern animal.

A VISION FOR THE FUTURE

Before this approach to *in silico* sequencing becomes practical, it is likely that the number of SNPs and other polymorphisms available will increase steadily and this will include an increasing number of functional or causal mutations. This could present a considerable problem because SNP panels would rapidly become out of date because they did include recently discovered polymorphisms and the genotypes on older animals would be missing all recent discoveries. The proposal to use a standard sparse SNP panel overcomes this problem. Most animals are genotyped with the sparse panel which stays constant over time. Only the key ancestors need to be genotyped for all newly discovered and useful markers. In this way the cost of genotyping is low but the modern animals receive inferred genotypes for all markers thought to be important enough to genotype on the key ancestors. The same sparse panel could potentially be used for all cattle, all over the world, for many years. This would mean the sparse SNP chip would be manufactured in huge numbers and so should be relatively inexpensive and less costly than SNP chips designed for a limited market even if they genotyped far fewer SNPs.

REFERENCES

De Roos, A. P. W., Hayes, B. J., Spelman, R. and Goddard, M. E. (2008). *Genetics.* **179**:1503.

Goddard, M.E. (2008) In Pinard M-H, Gay C., Pastoret, P-P and Dodet, B. (eds): Animal Genomics for Animal Health. Dev Biol. (Basel), Basel, Karger **132**:383.

Goddard, M.E. and Hayes, B.J. (2008) Artificial selection methods and reagents. Patent application WO/2008/074101.

Habier, D., Fernando, R.L. and Dekkers, J.C.M. (2009) Genomic selection using low-density marker panels. Genetics published ahead of print as 10.1534/genetics.108.100289.

Meuwissen, T. H .E., Hayes, B. J. And Goddard, M. E. (2001) *Genetics* **157**:1819.

Schaeffer, L. R. (2006) *J. Anim. Breed. Genet*. **123**:218.

Scheet, P., and M. A. Stephens (2006) *Am. J. Hum. Genet*. **78**:629.

VanRaden, P. M., Van Tassell, C. P., Wiggans, G. R., Sonstegard, T. S., Schnabel, R. D., Taylor, J. F. and Schenkel, F. S. (2009) *J Dairy Sci* **92**: 16.

INTEGRATION OF DNA MARKERS INTO BREEDPLAN EBVS

D. J. Johnston, B. Tier and H.-U. Graser

Animal Genetics and Breeding Unit*, University of New England, Armidale, NSW 2351

SUMMARY

Genotypes for the four GeneSTAR® tenderness markers and phenotypes for shear force of meat and flight time were used to develop marker assisted estimated breeding values for shear force (SF EBV^M). The partial regression coefficients for the effects of the four GeneSTAR tenderness markers on shear force were estimated and these were used in a prediction equation to calculate a single value for use in a multiple trait animal model to predict SF EBV^M on all animals.

INTRODUCTION

Livestock selection has used estimated breeding values (EBV) based on phenotypic data and pedigree records for more than 40 years. More recently, advances in molecular genetic techniques, in particular DNA sequencing, have led to the discovery of regions of the genome that influence traits in livestock. However utilising both sources of data in genetic evaluation schemes, such as BREEDPLAN, has been a challenge due to the heterogeneity of data sources, the multi-trait nature of the evaluations, and unknown effects of the marker information on all traits in the evaluation. The SmartGene for Beef project identified significant effects of the Catapult Genetics GeneSTAR® tenderness markers on meat tenderness as recorded by the objective measure of shear force (Johnston and Graser 2008; Johnston and Graser 2009). These results have been used to further develop methods for combining EBVs (i.e. phenotypic and pedigree data) and gene marker information into a single marker-assisted EBV called an EBV^M. Flight time is an objective measure of an animal's temperament which has been shown to be heritable and moderately genetically correlated with SF (Kadel *et al.* 2006), thus representing a potential genetic indicator trait for meat tenderness. Therefore the aim of this study was to use phenotypic records and GeneSTAR tenderness markers for the Brahman breed to estimate their combined effects on SF and incorporate them, using a multi-trait framework, into a marker assisted SF EBV^M.

MATERIALS AND METHODS

Data. The computation of the new tenderness EBV^M used three different sources of data. Shear force (SF) records from meat samples from the *M. longissimus thoracis et lumborum* of carcasses were available from Brahman animals measured through the Beef CRC1 (Johnston *et al.* 2003) and CRC2 (Wolcott *et al.* 2009) pedigreed breeding programs (N=1,995). GeneSTAR (commercialized by Catapult Genetics, now Pfizer Animal Genetics) gene marker results for one or more of the four tenderness markers (viz: T1, T2, T3, T4) were available on 7,040 Brahman animals from the Beef CRC1 and CRC2 projects (genotyped through the SmartGene Project) and industry tested animals that had GeneSTAR results submitted to the Brahman breed database. Finally, flight time (FT) records (N=4,737) on Brahmans recorded in the two CRC projects were used, along with a small number of industry recorded animals.

Computation of Cumulative Marker Phenotype (CMP). The first step to enable the inclusion of marker information into an EBV^M, was to compute the effects of each of the markers (i.e. T1, T2, T3, T4) on shear force, specifically for Brahmans using models and datasets developed as part of the SmartGene Project. The model included fixed effects previously defined for the two CRC

* AGBU is a joint venture of NSW Department of Primary Industries and University of New England

experiments (Johnston *et al.* 2003 and Wolcott *et al.* 2009) and the genotypes for the four markers were included as linear covariables in the one model. The number of animals per genotype class and the estimated effects of the four tenderness markers are shown in Table 1.

The estimated partial regression coefficients for the effects of the four markers were used to construct a prediction equation that was then applied to each animal's marker genotypes to give a single prediction of the cumulative effects of the markers (CMP) for all genotyped Brahmans. However, given the significance of the T1 and T3 markers, both had to be genotyped for an animal to have a CMP computed. In total there were 4,729 animals with a computed CMP with a mean of 0.001 and a range of -0.68 to 0.30 kg SF.

Table 1. Number of animals per genotype (0, 1, 2), gene frequency of favourable allele (gf_q) and estimated partial regression coefficients (b) and standard errors of the four GeneSTAR tenderness markers (T1, T2, T3, T4) on SF in Brahman from Beef CRC1 and CRC2

Genotype	T1	T2	T3	T4
0	386	2066	1908	794
1	1322	141	753	1359
2	1081	14	76	627
gf_q	0.62	0.04	0.17	0.47
b (kg SF/allele)*	**-0.139** (0.041)	-0.087 (0.105)	**-0.234** (0.054)	-0.032 (0.040)

* bold significant P<0.05

Multi-trait analysis. A three-trait BLUP model was constructed using SF and FT phenotypes as well as the CMP record for the genotyped animals. Fixed effects of contemporary group (as previously defined) were fitted for FT and SF as well as linear covariates for age and carcass weight, respectively. Variance components were derived using a trivariate model where the residual variance of the CMP was set at 0.001 to assist convergence (i.e. heritability close to 1). The heritabilities of SF and FT were 0.30 and 0.20 respectively and the genetic correlation between SF and FT was 0.25 and between SF and the CMP was 0.28. The estimated genetic correlation between CMP and FT was small (-0.05), as were the residual correlations with SF and FT. Therefore model configuration allowed the information from the markers to contribute to the SF EBV^M through the genetic correlation structure.

Table 2. Genetic correlations (r_g) and percent additive genetic variance ($\%V_A$) explained by the GeneSTAR tenderness MVP on SF in four CRC1 datasets

	MVP	SF records			SF, MVP	SF, MVP
CRC1 Dataset	N	N	V_A	h^2	r_g	$\%V_A$
Straightbred temperate	659	3322	0.433	0.08 (0.04)	0.170 (0.14)	2.9
Straightbred tropical	585	3254	0.612	0.30 (0.06)	0.283 (0.08)	8.0
Crossbred temperate	253	785	0.658	0.26 (0.10)	0.126 (0.14)	1.6
Crossbred tropical	225	762	0.871	0.31 (0.10)	0.547 (0.13)	29.9

Adapted from www.beefcrc.com.au/Assets/473/1/Pfizer2.pdf Table 3

Comparison with new Molecular Value Predictions (MVP™). Whole genome associations studies using tens of thousands of SNP are finding large numbers of SNP, generally with small effects, associated with a range of economically important traits. In March 2009, Pfizer Animal Genetics computed a molecular breeding value (MBV) for meat tenderness (marketed as GeneSTAR® MVP™) using an extended panel of 56 SNP, including the four GeneSTAR

tenderness markers. This development signalled the move away from single SNP markers to the computation of MVPs. The effects of the GeneSTAR tenderness MVP were calibrated in four Beef CRC1 datasets and estimates of the genetic correlation and the percent additive genetic variance explained are presented in Table 2.

RESULTS AND DISCUSSION

Size of effects and gene frequencies. The results showed that the size of effects of the four markers differed (Table 1). Marker T3 had an estimated effect almost as large as the other markers put together, and markers T1 and T2 had intermediate effects. The gene frequencies of the four markers also differed. The frequency of the desirable form of the T2 marker was extremely low whereas that of T1 exceeded 60%. The different estimated size of the marker effects means that animals with the same total number of alleles (i.e. total stars) are likely to have different EBV^M.

Comparison with MVPs. Combining the effects of the four markers into the CMP in this study is a similar technique used to construct MBVs or MVPs. Results from this study showed that the cumulative effects of the four tenderness markers explained 8% of the additive genetic variance of SF (i.e. 0.28^2). This is comparable with variance explained by the new Pfizer tenderness MVP (Table 2) for tropical breeds (8%) and with the estimate of 7.4% in Brahmans from CRC2 (Johnston *et al.* 2009). Therefore the contribution to the SF EBV^M from the new tenderness MVP for Brahmans will be comparable to that from the original four GeneSTAR tenderness markers.

Computed EBV^M. A total of 22,052 animals (those with records and their ancestors) had SF EBV^M computed with a mean of 0.02 and a range of -0.98 to +1.36 kg. The units of the tenderness EBV^M are kg of SF and therefore lower (i.e. more negative) EBV^M have lower SF and more tender meat. The spread in the SF EBV^M of sires with an accuracy greater than 80% was almost 2 kg SF. Sires used in the Beef CRC projects were a random sample of the Brahman breed and the large spread generated was mainly the result of the large number of progeny recorded (N=16 to 68) for actual shear force for each of these sires, and the contribution of the other sources of information (i.e. markers and flight time) was minimal. When marker genotypes were the only available data then the SF EBV^M had a reduced spread equal to the maximum difference in the CMP, with an accuracy of 28%. The accuracy of the SF EBV^M for animals with their own SF record was approximately 55% and 79% when 20 progeny were recorded for SF. Adding marker information when phenotypic records were already available on the trait had little effect on increasing the accuracy.

Contribution of gene markers to EBV^M. The effects of the tenderness markers were included in the EBV^M using the multiple trait model where the relative contribution of the markers to the EBV^M depended on the estimated effect of each marker, their gene frequency and the genetic variation accounted for by the marker. Therefore if a marker has a reasonable size effect but is at high gene frequency (i.e. most animals have 2 copies) in a particular breed then this marker will be explaining very little of the differences between animals and therefore will have little contribution to differences in EBV^M between animals. To generate EBV^M with high accuracies from marker data alone will require finding numerous markers that explain a large amount of the genetic variation of a trait.

Once an animal's genotype has been established there is no benefit for that animal's EBV^M in testing relatives with the same panel of markers. This is different to recording phenotypic information, like flight time, where the records on relatives (i.e. sire, dam, half sibs and progeny) can be of considerable benefit in increasing the accuracy of the EBV and the EBV^M.

Future research. This new trial tenderness EBV^M is the start to a new chapter in the genetic evaluation of beef cattle in Australia. Research is underway to determine if the tenderness markers are genetically correlated with other economically important traits. Early indications, using Brahman BREEDPLAN EBVs, are that the tenderness markers are having no substantial effects on any of the published EBVs. Beef CRC2 research has also shown few antagonisms between shear force and other traits, but research is continuing to assess female lifetime reproductive performance and cow survival. To include tenderness EBV^M into a breeding program will require their inclusion into a selection index, and this will require the determination of the economic value of tenderness and estimates of genetic correlations with all other traits in the breeding objective.

The other important development is the expanding capacity to genotype animals for large numbers of potential markers and therefore the increasing opportunity to explain greater amounts of genetic variation. However with the trend towards panels of SNPs, and then the prediction of MBVs, there will be an ongoing need for the estimation of variances and covariances with phenotypic traits as well as any existing MBVs. To fully utilize this capacity the Australian beef industry needs to record many more phenotypes on animals with a DNA sample, particularly for traits difficult to record in industry. Efforts to do this are currently underway. Although there are theoretical predictions of completely replacing phenotypic records with large panels of markers in the future, this appears to be some way off for the beef industry. The current experience from the dairy industry, where the phenotypes used are the highly accurate EBVs of proven sires, is that the accuracies of EBV^M derived from only DNA marker information are around 70% (e.g. VanRaden *et al.* 2008). However this is unlikely to be achieved in the short-term in beef where unlike the dairy industry, there are relatively few animals with phenotypes or large numbers of sires with highly accurate EBVs for all economically important traits.

CONCLUSIONS

These SF EBV^M represent a first for the Australian beef industry, and a significant advancement in genetic evaluation. The methodology developed can be extended to use molecular breeding values from DNA companies, provided there are accurate estimates from independent datasets of the marker variance and, most importantly, the genetic correlations between the MBV and other traits.

ACKNOWLEDGMENTS

This research was supported by Meat and Livestock Australia. We also thank the contributions of the Australian Brahman breeders, Pfizer Animal Genetics and the Beef CRC.

REFERENCES

Johnston, D.J., Reverter A., Ferguson, D.M., Thompson, J.M. and Burrow, H.M. (2003) *Aust. J. Agric. Res.* **54**: 135.

Johnston, D.J. and Graser, H-U. (2008) http//:www.agbu.une.edu.au/Smartgene% 20Final%20AGBU%20Report.pdf

Johnston, D.J., and Graser, H-U. (2009) *J. Anim. Prod. Sci.* (submitted)

Kadel, M.J., Johnston, D.J., Burrow, H.M., Graser, H-U. and Ferguson, D.M. (2006) *Aust. J. Agric. Res.* **57**:1029.

Wolcott, M.L., Johnston, D.J., Barwick, S.A., Iker, C.L., Thompson J.M. and Burrow H.M. (2009) *J. Anim. Prod. Sci.* **49**:383

VanRadan, P.M., Van Tassell, C.P., Wiggans, G.R., Sonstegard, T.S., Schnabel, R.D., Taylor, J.F. and Schenkel, F.S. (2008) *J. Dairy Sci.* **92**:16.

ACCURACY OF GENOMIC SELECTION: COMPARING THEORY AND RESULTS

B. J. Hayes[1], H. D. Daetwyler[2,3], P. Bowman[1], G. Moser[3], B. Tier[4], R. Crump[4], M. Khatkar[5], H. W. Raadsma[5], and M. E. Goddard[1,6]

[1]Biosciences Research Division, Department of Primary Industries Victoria, 1 Park Drive, Bundoora 3083, Australia, [2]The Roslin Institute and R(D)SVS, University of Edinburgh, Roslin, EH25 9PS, United Kingdom, [3]Animal Breeding and Genomics Centre, Wageningen University, 6700 AH Wageningen, The Netherlands, [5]Centre for Advanced Technologies in Animal Genetics and Reproduction (ReproGen), University of Sydney, Camden NSW 2570, Australia, [4]Animal Genetics and Breeding Unit, University of New England, Armidale NSW 2351, Australia, [6]Faculty of Land and Food Resources, University of Melbourne, Parkville 3010, Australia.

SUMMARY

Deterministic predictions of the accuracy of genomic breeding values in selection candidates with no phenotypes have been derived based on the heritability of the trait, number of phenotyped and genotyped animals in the reference population where the marker effects are estimated, the effective population size and the length of the genome. We assessed the value of these deterministic predictions given the results that have been achieved in Holstein and Jersey dairy cattle. We conclude that the deterministic predictions are useful guide for establishing the size of the reference populations which must be assembled in order to predict genomic breeding values at a desired level of accuracy in selection candidates.

INTRODUCTION

Genomic selection refers to the selection of animals for breeding based on genomic breeding values. Meuwissen . (2001) demonstrated using simulation that the accuracy of genomic breeding values can be very high if they are predicted from a large number of DNA markers. Provided the markers are dense enough, the accuracy of genomic breeding values will depend on the number of individuals genotyped and phenotyped in the reference population where the effect of the markers are predicted, the heritability of the trait, and the number of independent loci or chromosome segments in the population (Goddard 2008; Daetwyler *et al.* 2008). Goddard (2008) and Hayes *et al.* (2009) further derived deterministic predictions of the number of independent chromosome segments based on the effective population size and the length of the genome of the species in question. These deterministic predictions would have great value in guiding the design of experiments to implement genomic selection if the accuracy they predicted agreed with that observed in real data. Such data is now available; recently, Van Raden *et al.* (2009) reported accuracies of genomic breeding values as high as 0.75 for total merit index in Holstein Friesian dairy cattle using 38416 single nucleotide polymorphism (SNP) markers genotyped in 3576 progeny tested bulls. Accuracies of genomic selection are also available for Australian Holstein Friesian and Jersey cattle, using a similar number of SNPs.

The aim of this paper was to assess the value of the deterministic predictions of accuracy of genomic breeding values given results that have been achieved in Holstein and Jersey dairy cattle.

MATERIALS AND METHODS

In Daetwyler *et al.* (2008) the accuracy of genomic breeding values was predicted as $r = \sqrt{Nh^2/(Nh^2 + q)}$ where N = number of individuals genotyped and phenotyped in the reference population, h^2 = heritability of trait or reliability of breeding values in the reference population, q = number of independent chromosome segments in the population. Daetwyler et al.

(2008) also proposed a corrected for their prediction when $N \geq q$. The correction was to add $r^4 q/(2N)$ to the above prediction to get the final accuracy. As $N \geq q$ for most of the situations we will investigate, we will use the accuracy from the above equation with the correction.

In Goddard (2008), the accuracy of genomic breeding values was predicted as $r = \sqrt{\left[1 - \lambda/(2N\sqrt{a}) * \ln((1 + a + 2\sqrt{a})/(1 + a - 2\sqrt{a})\right]}$ where a= 1+2 λ/N , and $\lambda = qk/h^2$, with $k = 1/log(2N_e)$, where N_e is the effective population size. Note that this derivation assumes that σ_e^2 is close to the phenotypic variance. For both predictions, the value of q used was the number of independent chromosome segments, $2N_eL$, where L is the length of the genome in Morgans (Hayes *et al.* 2009). The difference between the formula of Daetwyler *et al.* (2008) and Goddard (2008) potentially arises because Goddard (2008) assumed that that the effect estimate for common QTL is more accurate for QTL with intermediate allele frequency, because they explain more of the genetic variance than QTL with extreme allele frequency In contrast, Daetwyler *et al.* (2008) assumed the accuracy of estimating QTL effects was equal regardless of their frequency. The accuracy of genomic breeding values for the two deterministic predictions were compared for a range of heritabilities, N=5000 and N_e=100.

Accuracy of breeding values from the two predictions were also compared to accuracies of genomic breeding values reported by VanRaden *et al.* (2009) and United States Department of Agriculture results (http://aipl.arsusda.gov/reference/genomic_comparison_yng_0901.htm) for total net merit in Holstein Friesian cattle and Jersey cattle, and our own results in Australian data for these two breeds. The phenotypic records in the reference population were daughter yield deviations (DYD) for total merit index for the US data or de-regressed breeding values for Australian Profit Ranking (APR) in the Australian data. The average reliability of the DYD in the reference population was 0.9. In order to deterministically predict the accuracy that these experiments could have achieved, an assumption of the N_e in each breed was required. Young and Seykora (1996) gave an estimate of 100 for the N_e of US Holsteins. The N_e in Australian Holsteins is similar (deRoos *et al.* 2008). For US Jerseys, the effective population size is smaller, with a recent estimate of 30 (Weigel *et al.* 2008). The N_e of Jersey's in Australia is likely to be similar given the large contribution of US Jersey bulls to the Australian population. Given these estimates of N_e in the two breeds, we used N_e=100 in the predictions for Holsteins and 30 in Jersey's. A genome length of 30 Morgans was assumed.

RESULTS AND DISCUSSION

The accuracies of genomic breeding value predicted by Goddard (2008) and Daetwyler *et al.* (2008) are similar, though Daetwyler *et al*. (2008) would predict a lower accuracy of breeding value at low to moderate heritabilities given the same number of independent chromosome segments and number of phenotypic records, Figure 1. Both deterministic predictions agreed fairly well with the accuracies of genomic breeding value reported for US and Australian Holstein Friesian and Jersey dairy cattle, Figure 2. The % error was low for the Goddard prediction vs the US Holstein data at 3%.. However in the Australian Holstein data the observed accuracies were somewhat higher than the predictions. This may just reflect a small validation sample used in the Australian data leading to a large standard error for the estimate of reliability.

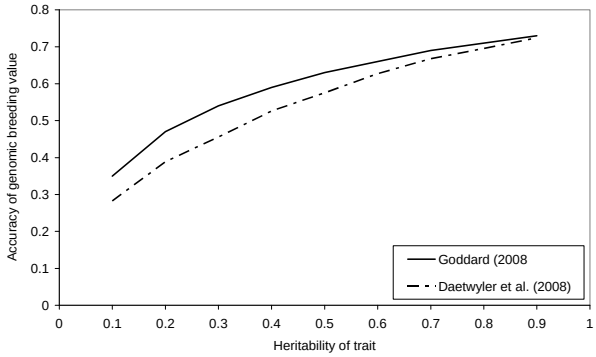
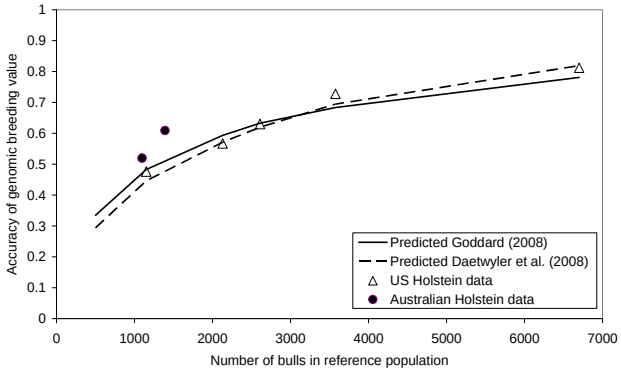


Figure 1. Accuracy of genomic breeding values with 5000 phenotypic records, effective population size of 100 and increasing heritability, predicted by the deterministic formula of Goddard (2008) or Daetwyler et al. (2008).

A



B

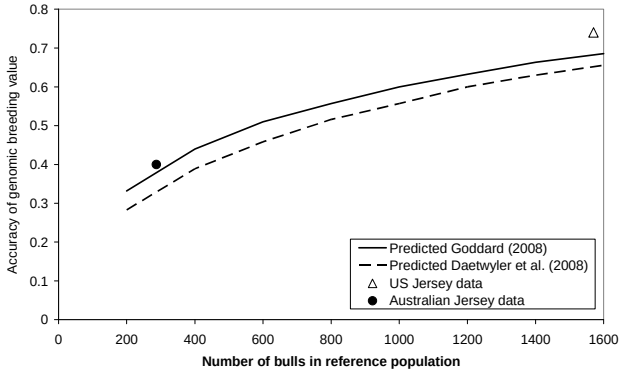


Figure 2. A. Accuracy of genomic breeding values from the deterministic prediction of Goddard (2008) and Daetwyler *et al.* (2008) with N_e=100, and accuracy of genomic breeding value for total merit index or Australian Profit ranking in US or Australian Holstein Friesian cattle. B. Accuracy of genomic breeding values from deterministic predictions with N_e=30, and accuracy of genomic breeding value for total merit index or Australian Profit ranking in US or Australian Jersey cattle respectively.

Another contributing factor may be that the deterministic predictions assume that the accuracy of breeding values is a result of the SNPs capturing the effect of QTL, whereas some of the accuracy of genomic breeding values in livestock populations may be a result of the SNPs capturing the effect of relationship, particularly if there are large half sib families in the population (eg. Habier *et al.* 2007). For comparison, the accuracy of parent average breeding values for net merit available for young bulls in the US data was 0.37 (VanRaden *et al.* 2009).

The deterministic method of Goddard (2008) used here assumes a normal distribution of QTL effects. For the majority of traits studied by Van Raden *et al.* (2008), methods for predicting genomic breeding values which assumed a normal distribution of quantitative trait loci (QTL) effects performed almost as well as methods assuming a exponential distribution of QTL effects. The exception was traits with a QTL of known large effect, eg. fat percentage (Grisart *et al*. 2003). For such traits, the deterministic prediction of Goddard (2008) would under-predict accuracy of genomic selection. The accuracies of prediction also depend on N_e. The values of N_e used here are estimates of N_e in the recent past, however N_e in cattle has been much larger historically. It is not clear how the change in historical N_e should affect accuracy of genomic breeding values. Nevertheless, using current N_e gave good agreement between predictions and observed results.

CONCLUSIONS

The deterministic predictions of accuracy of genomic selection presented by Goddard (2008) extended by Hayes *et al.* (2009), and that of Daetwyler *et al.* (2008) agree well with observed accuracies of genomic selection in US and Australian Holstein Friesians and Jerseys. We can conclude that these deterministic predictions are a useful tool to guide design of genomic selection experiments, for example how large should the reference population be to achieve a desired level of accuracy. It must be noted we have compared predicted and observed accuracies of genomic breeding value for a situation where phenotypes were very accurate predictors of breeding value. The performance of the deterministic predictions of both Daetwyler *et al.* (2008) and Goddard (2008) should be also evaluated in other situations where the heritability of the trait is lower, as the difference predicted accuracy of genomic selection is greater at lower heritabilities.

ACKNOWLEDGMENTS

The authors are grateful to Curt van Tassell and Tad Sonstegard from the USDA for providing genotypes under a collaborative agreement between USDA and Department of Primary Industries Victoria. We are also grateful to Paul Van Raden and George Wiggans of USDA for providing some of the information required to assemble Figure 2.

REFERENCES

Daetwyler, H.D, Villanueva, B. and Woolliams, J.A. 2008. *PLoS ONE* **3**: e3395.
de Roos, A.P., Hayes, B.J., Spelman, R.J. and Goddard, M.E. 2008 *Genetics* **179:**1503
Goddard, M.E. 2008. *Genetica*. doi: 10.1007/s10709-008-9308-0.
Habier D, Fernando RL, Dekkers JC. 2007. *Genetics.* **177**:2389
Hayes, B.J., Visscher, P.M., and Goddard, M.E. 2009. *Genet. Res.* **91**:47
Meuwissen, T.H.E., Hayes, B.J, and Goddard, M.E. 2001. *Genetics* **157**: 1819
VanRaden, P.M. *et al.* 2009 *J. Dairy Sci.* **92**:16
Weigel, K. A. *J.Dairy.Sci.***84**(E. Suppl.):E177
Young, C.W. and Seykora, A.J. 1996.. *J. Dairy Sci.* **79:**502

POTENTIAL BENEFIT OF GENOMIC SELECTION IN SHEEP

J.H.J. van der Werf [1,2]

[1]Cooperative Research Centre for Sheep Industry Innovation, Armidale, NSW 2351
[2]School of Environmental and Rural Science, University of New England, Armidale, NSW, 2351

SUMMARY

The short term benefits of applying whole genome selection to sheep breeding programs are estimated. If breeding values of all objective traits can be predicted with accuracy equal to the square root of heritability, genomic selection could increase overall response for a terminal sire index by about 30%, and a fine wool merino index by about 40%. Response is relatively more increased for those traits that are normally not measured on breeding animals.

INTRODUCTION

Major research efforts are undertaken worldwide to attempt prediction of genetic merit from genomic information. This renewed effort is motivated by the idea of genomic selection (GS), i.e. selection based on the joint effects of very many genes linked to markers densely covering the whole genome. The first paper to propose the idea of predicting genetic merit from dense gene markers was published by Meuwissen *et al.* (2001), showing theoretically a potential accuracy of up to 80%. Developments in genotyping technology, along with the publication of genome sequences for human in 2001, and bovine in 2006, have allowed rapid detection of Single Nucleotide Polymorphisms (SNPs) and development of SNP-chips for rapid genotyping. A bovine SNP chip was initially developed for 10,000 (10k) markers and in 2007 a 56k chip was released by Illumina. Similarly, a 57k ovine chip was developed and released in August 2008, partly based on a *virtual ovine genome sequence* (Dalrymple *et al.* 2007).

First results of whole genome association studies in dairy cattle have shown some convincing results, with the ability to predict breeding values of young dairy bulls with an accuracy of around 60%. This is the correlation between a predicted breeding value based on genomic data and one based on the mean performance of a large number (~100) of progeny, i.e. genomic information can predict around 36% of the variation in true breeding value (additive genetic variance). The benefit of this to dairy breeding programs is large. As genetic change is largely based on selection of bulls, and since milk production is sex limited, breeders have to accept long generation intervals and wait for a progeny test, or select at a younger age based on less accurate EBVs. Genomic selection allows selection of bulls at a young age based on a much improved accuracy. Schaeffer (2006) suggested that dairy programs could possibly double their rate of genetic gain by using GS.

The benefit of GS might be smaller in sheep as 1) many traits under selection in sheep can be measured on both sexes an before selection of animals for first mating, 2) several important traits have a high heritability, and 3) there maybe less information available to derive accurate prediction equations as there are fewer progeny tested sires. However, some traits in sheep are also difficult to measure on breeding animals, e.g. female fertility, slaughter traits, wool traits when measured on adults and parasite resistance. Furthermore, the Australian sheep industry invests in an *Information Nucleus* (Fogarty *et al.* 2007) that allows measurements of those objective traits that are normally not measured in commercial stud flocks. The overall benefit of GS will depend on measurability of the various breeding objective traits, and their relative economic importance. The purpose of this paper is to give a ballpark figure of potential benefit of GS in sheep breeding programs.

MATERIALS AND METHODS

Selection index theory was used to predict selection accuracy of male and female selection candidates in different age classes. The genomic information is modelled as an additional information source known at birth, explaining a percentage of the additive genetic variation (V_{QTL}). This percentage will be larger if more animals are used in deriving prediction equations and when the heritability is higher (Goddard 2006). Selection index theory was used to predict the accuracy of selection on a breeding objective, assuming the usual sheep traits measured. BLUP selection was mimicked, with information available on 29 half sibs, and 30 progeny on sires, when the appropriate age is reached. The index accuracy increases with age, and can vary between sexes. The age structure of a breeding program was optimized by truncation selection across age classes for each sex. Therefore, if genomic information was available, it was more likely that younger animals had high enough accuracies to be selected, and the optimal generation interval was generally lower. Increased selection accuracy and decreased generation intervals were therefore modelled as contributing to the benefit of GS.

The annual response to selection was compared for scenarios with and without genomic selection. First, this comparison was made for single trait selection, where the trait heritability, the trait measurability and V_{QTL} were varied. Secondly, the benefit of GS was compared for indexes relevant to the industry and derived by *SHEEP OBJECT* (Andrew Swan, personal communication), both for fine wool merinos and for terminal sire breeds. The terminal sire index, which is not (yet) used by the industry, contains carcass traits in the objective and growth and ultrasound scan measures for fat and muscle as selection criteria traits (see Table 2 for detail on traits). Three carcass traits have all a similar dollar value per genetic standard deviation, except that the value for carcass fat is negative and about one third as large. The main objective traits in the fine wool index were adult weight, adult clean fleece weight, adult mean fibre diameter, adult staple strength and number of lambs weaned. Young breeding animals have measurements for none of these traits at first selection.

Two GS scenarios were compared: under GS1 it was assumed that V_{QTL} is equal to the trait heritability (h^2). This is not a functional relationship, but a reasonable approximation of expected predicted accuracy of GS when phenotypic information on about 2500 animals is used for deriving GS prediction equations. Alternatively, $V_{QTL} = ½ h^2$, which would be achievable in an experiment with about 1000 phenotypic measurements

RESULTS AND DISCUSSION

Results of single trait selection with and without GS are shown in Table 1. The benefit of GS is clearly highest for traits that are more difficult to measure and have low heritability. The maximum benefit is a doubling of genetic gain. It should be noted that in practice it is more difficult to achieve a high GS accuracy for traits with low h^2 as much more phenotypic data are needed to derive GS prediction equations of such accuracy. The benefit is obviously highest for traits that cannot be measured at all, unless such traits have very high correlations with measured traits. Table 1 gives a benchmark figure for the approach used here. The potential benefit for sheep will have to be evaluated for a combination of traits, each with benefits varying according to conditions outlined in Table 1.

Results in Table 2 show for terminal sire breeds a 32% increase in overall response with accurate genomic selection (GS2) and a 16% increase with less accurate genomic selection (GS1). Obviously, individual carcass traits benefit greatly from GS as these traits are not measured on breeding animals. As a consequence, the GS response for traits measured post weaning is lower than with no GS. Results for a selection index for fine wool merino show a 38% increase in overall response with accurate genomic selection (GS2) and a 19% increase with GS1. Although fine wool traits are highly heritable and can be measured, the adult expression of these traits is usually

not recorded. Moreover, number of lambs weaned has limited measurability and is much more improved whereas staple strength, being an unmeasured trait, declines much less under genomic selection.

Table 1 Increase of rate of genetic gain (%) for single trait genomic selection for various degrees of variance explained by genotype (V_{QTL}), heritability (h^2) and trait measurability conditions.

Trait Measurability	V_{QTL} = 10%			V_{QTL} = 30%		
	h^2=0.1	h^2=0.3	h^2=0.5	h^2=0.1	h^2=0.3	h^2=0.5
Measured < 1 year, males and female	13	4	2	37	13	6
Measured > 1 year, males and females	19	9	6	64	29	18
Measured >1 year, females only	38	17	13	109	54	39
Measured on Corr. Trait, r_g = 0.9	17	7	4	48	20	11
Measured on Corr. Trait, r_g = 0.5	61	33	24	143	83	62

Listed heritability values refer to trait under selection, i.e. to correlated trait if applicable

Table 2. Accuracy and annual response to selection for meat sheep and fine wool merino indexes in scenarios with no (no GS), less accurate (GS1) and more accurate (GS2) genomic selection.

			Accuracy			Response (trait units)		
Meat Sheep Index	Phen	h^2	no GS	GS1	GS2	no GS	GS1	GS2
Muscle conformation (mm)	0	0.25	0.58	0.63	0.68	0.36	0.34	0.34
Dressing %	0	0.30	0.37	0.47	0.59	0.16	0.27	0.37
Saleable meat yield %	0	0.30	0.35	0.46	0.59	0.36	0.46	0.55
Carcass fat depth (mm)	0	0.30	0.41	0.51	0.62	-0.05	-0.06	-0.07
Post weaning weight	1	0.30	0.68	0.72	0.76	0.39	0.36	0.34
PW fat depth US	1	0.20	0.58	0.62	0.66	0.01	0.01	0.01
PW eye muscle depth US	1	0.30	0.66	0.70	0.75	0.34	0.31	0.29
Overall Merit ($Index)			0.46	0.58	0.66	1.30	1.51	1.71

			Accuracy			Response (trait units)		
Wool Sheep Index	Phen	h^2	no GS	GS1	GS2	no GS	GS1	GS2
Adult body weight	0	0.40	0.57	0.67	0.75	0.17	0.10	0.03
Adult clean fleece weight	0	0.44	0.48	0.62	0.74	0.03	0.06	0.07
Adult mean fibre diameter	0	0.60	0.66	0.76	0.85	-0.26	-0.25	-0.26
Adult staple strength	0	0.30	0.41	0.53	0.63	-0.37	-0.16	-0.01
Number of lambs weaned	fem	0.06	0.25	0.30	0.34	0.001	0.004	0.006
Overall Merit ($Index)			0.43	0.52	0.60	3.30	3.94	4.54

Variance explained by genotypes as proportion of additive genetic variance is equal to the heritability (h^2) for GS2 and ½h^2 for GS1, Phen indicates whether or not a trait is measured.

DISCUSSION

The approach followed in this study gives an approximate estimate of the potential value of GS. The results point out the main findings, i.e. the magnitude of the relative additional benefit is similar for meat and wool sheep, being between 15% and 40 %, and the potential change in trait

emphasis, away from easy to measure traits towards economically important but hard to measure traits. This shift in trait response is an important outcome of GS, besides the overall additional gain in total merit. The selection index approach works from the basis of information sources and their explained variance, and in that sense should be a reasonably robust approach in predicting GS outcomes. The main shortcoming of the method used here is that the long term effect of selection is not modelled. In an ongoing population under directional selection, the genetic variance generally reduces to about 75% of the variance in unselected populations. More importantly, it is mainly the between family variance that decreases whereas the within family (WF) variance is much less affected by selection. It is not fully clear whether GS exploits between or WF selection. If GS is based on estimated QTL effects, it would be more likely to exploit WF variance. A simulation study by Daetwyler *et al.* (2007) showed that GS limits the decrease in effective population size, suggesting that it works more at the WF variance level. However, we currently know little about the true genetic model underlying GS. There could be a number of factors that cause GS to be overestimated with the current theoretical approach, e.g. in reality, the prediction equations might not be equally accurate in all environmental and genetic backgrounds. Experimental evidence is needed to support these theoretical predictions. In addition, more work is needed to find ways of practical implementation of GS in the sheep industry.

ACKNOWLEDGEMENT

Andrew Swan from AGBU is acknowledged for assisting with genetic and economic parameters.

REFERENCE

Dalrymple, B.P., Kirkness, E.F., Nefedov, M., McWilliam, S., Ratnakumar, A., Barris, W., Zhao, S., Shetty, J., Maddox, J.F., O'Grady, M., Nicholas, F., Crawford, A.M., Smith,T., de Jong, P.J., McEwan, J., Oddy, V.H. and Cockett, N.E. (2007) The International Sheep Genomics Consortium. *Genome Biology* **8**:R152

H.D. Daetwyler, H.D., B.Villanueva, P.Bijma andJ.A. Woolliams (2007). *J. Animal Breed. Genet.* **124**:369

Goddard, M.E. (2007). *Genetica* 14 Aug 2008 (doi: 10.1007/s10709-008-9308-0).

Fogarty, N.M., Banks, R.G., van der Werf, J.H.J., Ball, A.J. and Gibson, J.P. (2007) *Proc.Assoc. Advmt. Anim. Breed. Genet.* **17:**29.

Meuwissen, T.H.E., B.J. Hayes and M.E. Goddard. (2001). *Genetics* **157**:1819.

Schaeffer, L.R. (2006) *J. Animal Breed. Genet.* **124**:218

GENOTYPE x ENVIRONMENT INTERACTIONS AND MERINO BREEDING PROGRAMMES FOR WOOL PRODUCTION

B.J.McGuirk

Strowan Lodge, 1c Albert Terrace, Edinburgh EH10 5EA, UK

SUMMARY

This paper and earlier reviews on genotype*environments interactions, many conducted on research stations and often restricted to fleece weights, indicate that whether the "genotypes" are breeds, Merinos bloodlines, selection flocks or, but rarely, sires, statistically significant genotype*environment interactions are of the scale type, and do not indicate changes of rank. These findings are discussed in the context of national across-flock genetic evaluations, and this paper suggests that greater use should be made of planned sire and bloodline evaluations if we are to better both quantify and explain such interactions, at least to the Merino industry.

INTRODUCTION

The term "genotype x environment interaction" (G x E) implies that the difference between genotypes is not equal in a set of environments (Woolaston, 1987). "Genotypes" might mean breeds, or, in the context of Merinos, strains, bloodlines, or even smaller genetic differences, such as between sire progeny groups. However, even this rather simple description hides some perhaps unexpected difficulties, such as the advisability of data transformations, or whether genotypes should be treated as fixed or random effects. But to my mind the critical issue is that our definition of interactions deals in outcomes, and not necessarily with causes, and that to "understand" such interactions, we feel the need to go beyond outcomes, and deal with causes. It seems to me fairly self-evident that as we move away from definable "causes", we are less likely to find consistent results across "experiments".

The possible existence of G x E interactions for fleece traits has long been discussed among practical Merino breeders and research workers. The focus here is on studies that were conducted in what might I term the "research station era" (see Morley, 1980, for a brief history), when comparisons were made of breeds, strains of Merino and of groups of smaller genetic differences (eg selection flocks), under the relatively well-defined environmental challenges possible on research stations. In other words, "causes" were usually well-controlled, as when samples of genetic groups were fed different diets (intakes) under pen conditions, with all other factors kept constant. As this topic was thoroughly reviewed by Woolaston (1987), the focus here will generally be on more recent studies.

The overriding impression from the older studies, aside from this control over the environmental factors generating the production differences, was the wish to show statistically significant G x E interactions for fleece weight. In consequence, populations of very different genetic merit were compared under environmental (generally nutritional) treatments that were also very different in degree. Thus these trials fall into what Dunlop (1962) termed Type 4 interactions. While this might be a good design to demonstrate a statistically significant G x E interaction, with an ensuing paper almost guaranteed, the relevance of this work to present-day industry concerns is sometimes hard to see. This review will attempt to bridge that gap.

COMPARISON OF WIDELY DIFFERENT GENETIC POPULATIONS - STRAINS AND SELECTION FLOCKS

This topic has had a fairly long history. For example, Dunlop (1962) makes reference to unpublished work of Marston, Pierce and Carter, who compared Strong and Fine-wool sheep fed

at two nutritional levels. As Dunlop (op. cit.) states: "The figures quoted make it almost certain that there were appreciable interactions of strain and plane of nutrition, in that the strong-wool strain had a greater capacity to respond in body size and wool production to good nutrition then the fine wool strain".

Saville and Robards (1972) published a notable paper on the same theme, when they compared samples of various Merino strains at Trangie, with ewes being fed to maintain weight, and then *ad libitum*. Again one might conclude that differences in wool growth between the strains are more apparent on unlimited feed (see Table 1) than at maintenance. But from the evidence (the strain means), the differences between the populations are in scale, not in rank.

Table 1. Wool growth per head (g/day) for Merino Strains at different intake levels (from Saville and Robards, 1972).

	Peppin Merinos (Trangie Random Flock)	Bungarees	Collinsvilles
Maintenance Feed	7.1	8.1	8.4
Unlimited Feed	11.4	14.5	15.9

BIG GENETIC DIFFERENCES, SMALL ENVIRONMENTAL DIFFERENCES

In his classic study, Atkins (1980) compared hogget wool production of five breeds (Peppin and South Australian Merinos, Corriedale, Polwarth and a fixed Border Leicester x Merino halfbred) at Temora in different years, in which average production varied between years, but in a manner similar to that on any commercial property.

Using the joint regression method of Freeman (1973) and Hill (1975), Atkins (1980) showed that the breed x year of measurement interaction was significant, and that the "advantage of the breeds with the highest (average) fleece weights", which in this case were the South Australian Merino and the Corriedale, "increased as the mean yearly fleece weight increased".

In a small scale study, Williams (unpublished data) compared the mid-side greasy wool production of Fleece Plus, Random and Fleece Minus rams at pasture over a period in 1962-1963, when the rams were generally grazed together. This was after about five generations of selection. If the "environments" (in this case six week periods of wool growth) are ranked in ascending order, the wool production (measured as g wool produced on a mid-side patch) in the Fleece Plus rams increases by 1.27 g/g increase in average production, compared with only 0.75 g/g increase in the Fleece Minus flock. Again differences between the flocks were greatest at the highest levels of production, which occurred in the Spring.

My own introduction to G x E interactions started with the Fleece Plus Relaxed flock at Trangie (see McGuirk 1980 for a brief mention). The flock was formed in the late 1960s, in response to concerns that response in the Fleece Plus flock had "plateaued" (see Pattie and Barlow 1974 – but see also Hatcher and Atkins 1998). I thought that this suggestion could be tested by splitting the Fleece Plus ewe flock, using extreme selected rams in the Fleece Plus flock, while average rams from the Fleece Plus flock were used to create a Relaxed line. After it was created, all replacements were selected at random from within the Relaxed flock. Unfortunately no formal comparison of the Fleece Plus and Relaxed flocks has yet been made.

Here the Fleece Plus Relaxed flock has been compared with its unselected Random control flock over five groups of hoggets, when we might expect that the relative genetic merit of the flocks would be the same. However if we compare the hoggets in five different years, it is clear that the difference between the flocks in average clean fleece weight increases as the average clean fleece weight of the Random flock increases (Table 2). In short, the better the environment, the bigger the response.

Table 2. Average Clean Fleece Weights (kg) for Fleece Plus Relaxed and Random hoggets.

Drop	Fleece Plus Relaxed	Random	Difference
1970	2.39	1.97	0.42
1971	2.56	2.03	0.53
1973	2.96	2.42	0.54
1968	3.57	2.98	0.59
1969	3.99	3.11	0.88

This is not a radical or even a new suggestion. Turner, Dolling and Kennedy (1968) discussed the possibility that drought conditions affected responses in the latter years of the CSIRO flock selected primarily for clean fleece weight. When summarising 41 years of selection for increased fleece weight at Trangie, Hatcher and Atkins (1998) concluded that the "pattern of response suggests that the expression of genetic superiority may be sensitive to the level of nutrition", and went on to illustrate this point with reference to responses in drought years. The Fleece Plus Relaxed flock at Trangie simply provides another example of this phenomenon.

SMALL GENETIC DIFFERENCES, LARGE ENVIRONMENTAL CHALLENGE

Recently I re-examined the data used by Morley (1956), in what was the first study of G x E interactions with Merino sheep. While it is often cited, the results are rarely discussed. In my re-presentation of this material, I came to very different conclusions from Morley, although my conclusions are apparent in the raw data (McGuirk 2009).

In general terms, Morley (1956) described two groups of daughters of 23 sires, drawn from various selection flocks in 1953. This puts the study at the very beginning of the selection phase at Trangie. All sires were in fact selectively drawn from a common line of rams, and these in turn were mated to ewes which were selected from previously unselected groups of base ewes. But while there was some degree of assortative mating, the genetic differences between the "flocks" were still quite small, and hence similar to that which might be observed in "the industry". Sire groups were split, and, between weaning and 12 months of age, young ewes were managed at pasture to achieve two very different levels of production. In all there were 93 ewes in the Low Plane group, and 80 in the High Plane, so that the number of daughters per sire was low, on average 4 or fewer in each treatment. This obviously has implications for the "power" of any sire x nutrition effect.

Production differences between the two nutritional treatments were considerable. At 17 months of age, average bodyweights were 29.6 kg (Low Plane) and 38.0kg (High Plane), with corresponding clean fleece weights of 2.64 and 3.67 kg. While differences between the treatments for these traits were thus respectively 25 and 35 percent, the effects on clean scoured yield and crimp frequency were small.

When the untransformed data are analysed using a model which includes the effects of Flocks, Nutritional Treatment and their interaction, the interaction effect was significant for clean fleece weight. However if we apply Tukey's test for non-additivity (see Tukey 1949), the Flock x Nutrition interaction for clean fleece weight was reduced to non-significance, as it was if we used log-transformed data. In other words, the statistical significance of the Flock x nutrition interaction for clean fleece weight was simply a scale effect.

If we look at the results of this trial in detail, we find that:

- The flock rankings were similar on the two Nutritional treatments.
- The variation between flocks was less on the Low Nutritional treatment, and

- The difference in performance between the two Nutritional treatments, the responsiveness of the flocks, was positively related to their performance or rankings on the Low treatment.

When the raw data were analysed using a model which includes the effects of Sires, Nutritional Treatment and their interaction, the interaction effect was not significant for any trait. This is hardly surprising for, with very small numbers of offspring per sire, the power of any test for statistical significance is very low. However, the variation (standard deviation) was greater among sires for offspring on the High Nutritional treatment, and the difference between the two nutritional treatments was positively related to sire means on the Low Nutritional treatment.

GENERAL ASSESSMENT OF RESEARCH STATION STUDIES

If we look at fleece weight comparisons involving breeds, strains, selection flocks and the limited information on sires (McGuirk 2009), we consistently find statistically significant G x E interactions, with the differences in performance greater as average production increases This holds both in studies where differences in production were solely due to intake or dietary differences (ie where the "causes" of the differences in production were determined experimentally) or where we are dealing with time effects associated with differences in production. But we see interactions of the scale type, not of rank (see James 2008). However these studies focussed almost exclusively on fleece weight, and we really know little if anything about other economically important traits that are of concern in national evaluations.

But to "understand" the nature of sire x environment interactions, I suggest we need to look at evidence from planned experiments, which is not the same as imposing an experimental "design" on a national dataset, where the performance defines the environmental treatments. We should look at various levels of environmental effects, and whether differences in performance can adequately be explained by scale differences. Ideally, our "experiment", should be seen as a vehicle to "explain" G x Es to a sceptical audience, as well as establishing statistical significance. Ideally then, our experiment should:

- Include a sufficient (and probably large) number of industry rams.
- Record as many traits as are economically important, including disease traits.
- Environmental differences should wherever possible be imposed, and not simply observed, with three or more "environments" represented.
- Sufficient offspring should be represented in each environment, so that all sire "proofs" have an agreed and minimum level of reliability.
- The results should be presented in a way that they are understood by producers, which will probably mean using regression methods, plotted with a pencil and ruler, rather than the sole reliance on estimates of genetic correlations.

As such an experiment is probably going to be seen as excessively large and expensive, we probably need to look at pre-existing data sets that meet at least some of these conditions. Thus the data set described by Dominik, Crook and Kinghorn (1999) should be revisited, especially as it deals with causes, and not simply with consequences, and includes data on a wide range of economically important traits. The data set used in "Merinos to Match" also needs a more rigorous assessment, even though the genetic groups included there are bloodlines, not sires. And data generated from Central Test Sire Evaluation programmes warrants continued monitoring. None of these studies meets all of the suggested requirements set out above, but I believe that the results might add to our overall understanding of G x E interactions.

My own view is that any G x Es observed for individual traits will be primarily of the scale type. However we might find rank changes if we look at indexes, as traits might show differential responses to changing environmental conditions.

IN WHICH ENVIRONMENT SHOULD WE SELECT?

As James (2008) pointed out, interest in the importance of G x E interactions was ignited by Hammond's (1947) proposal that selection should be in the best conditions available. Falconer (1952) responded that there was no general answer to this issue, which should be addressed in the general framework of correlated traits.

What then of the situation for fleece weight? My reading of the research station information, with only scale type interactions evident, is that the genetic correlation between "environments" is one or very close to that value. Hatcher and Atkins (1987) suggested that the heritability of clean fleece weight is essentially the same in Good and Poor years at Trangie, a conclusion supported by the work of Dominik ,Crook and Kinghorn (1999) in good and poor environments at Katanning. So if we want to improve clean fleece weights in "commercial" situations, unexpected differences in selection differentials aside, I suggest it does not matter whether the business of ram breeding is undertaken in "Stud" or "Commercial" conditions.

GENOTYPE x ENVIRONMENT INTERACTIONS IN THE MERINO INDUSTRY AND THE IMPLICATIONS FOR NATIONAL GENETIC EVALUATIONS

Are G x E interactions likely to be important in national genetic evaluations, such as those published by Sheep Genetics (SG)? Strictly speaking, from the evidence so far to hand, we do not know. Does it matter if sire x environment interactions were important in the Merino industry? The answer to that question is clearly "yes", as such interactions would make across flock sire evaluations rather meaningless, and complicate sire evaluations, even if they were only of the scale type. Yet the industry view is even more scathing. Carrick (2005) reported that a "recent survey among wool producers by the Victorian DPI" found "that over 90% of them believed that sires would **rank differently in different environments**" (my emphasis). Sadly I am unaware of how the producers were sampled for this survey, the precise wording of the questions posed, or if there was any empirical evidence for their views.

After numerous conversations, I am still not sure whether

- Producers do not accept that performance records for rams from the same flock/drop accurately reflect the ranking of these rams when they have offspring in other environments
- Whether the concern rather is to do with the suitability of various bloodlines in different environments
- Whether the concern is with the joint effect of these factors, and hence with cross-flock evaluations.

Each of these questions suggests analyses that might be undertaken using performance and progeny records.

I endorse the current SG practice of publishing information on the statistical importance of sire x flock/year effects (see for example, Brown, Swan, Johnston and Grasser 2009), as is commonly done with other species (see for example Mrode and Swanson 1994). However, perhaps SG should also routinely publish a more detailed description of such data sets. Specifically, for a group of sires, what numbers are used across flocks, and what proportion are sires used in different years? According to data collected Sheep Genetics, 105 sires had offspring born in different flocks in 2007, out of the 58,846 Merinos born in that year, for whom sire identities (1168 rams altogether) were also known. I suggest we need to know more about these 105 rams, whether they were used in other years, or other flocks, as part of Central Test Sire Evaluations, and whether they themselves had performance data. And what of the 228 rams used in flocks other than the one in which they were born?

I would advocate a more direct challenge to voices that are sceptical of cross-flock evaluations. But rather then doing this in a rather generalised manner, phrased in terms of genetic parameters, I would approach the matter rather more forensically, and be led by the sort of data presentations outlined above. This could be done "in-house" by Sheep Genetics, as a routine but integral part of its educational programme, accumulating information around an agreed set of questions.

ACKNOWLEDGEMENTS

I am grateful to the Director, Agricultural Research Centre, Trangie, for continuing access to historical data at Trangie, to Professor John James for statistical advice and general counselling, to numerous other people for stimulating discussions and access to their data, and to Sheep Genetics, especially Dan Brown and Sam Gill, for the data provided by that organisation.

REFERENCES

Atkins, K.D. (1980). *Aust. J. Exp. Agric.Anim.Husb*, **20**:280.

Brown, D.J., Swan, A.A., Johnston, D.J. and Graser, H-U. (2009) *Proc. Aust. Assoc. Anim. Breed. Genet.* **18:** these proceedings.

Carrick, M. (2005). Final report on Genotype by Environment Interaction in Wool Sheep. Meat and Livestock Australia, 58 pages.

Dominik, S., Crook, B.J, and Kinghorn, B.P.(1999). *Proc. Aust. Assoc. Anim. Breed. Genet.* **13:**98.

Dunlop, A.A., 1962. *Aust. J. Agric.Res.* **13**:503.

Dunlop, A.A. and Young, S.S.Y., 1966. *Aust. J. Agric. Res.* **17**: 227.

Falconer, D.S., 1952. *Amer. Nat.* **86**:293.

Freeman, G.H. (1973). *Heredity*, **31**:339.

Hammond, J. 1947. *Biol. Rev.* **22**: 195.

Hatcher, S. and Atkins, K.D. (1998). *Proc. Sixth Wld. Congr. Genet. Appl. Livest. Prod*. **24**:47.

Hill, J. (1975). *J. Agric. Sci.* **85**:477.

James, J.W. (2008). In "Adaptation and fitness in animal populations : evolutionary and breeding perspectives on genetic resource management". Eds J van der Werf, R. Frankham, H-U Graser and A.Gondro. Springer Verlag Berlin. Pp. 151 – 167.

McGuirk, B.J. (1980). *Proc Aust Soc Anim Prod* **13**:171.

McGuirk, B.J. (2009). *Anim. Prod. Sci*. (in press).

Morley, F.H.W. (1953). *Sci. Bull. 73, Dep. Agric. N.S.W.*

Morley, F.H.W. (1956*). Aust. J. Agric Res*., **7**:140.

Morley, F.H.W. (1980). *Proc Aust Soc Anim Prod* **13**:168.

Mrode, R. and Swanson, G.J.T. (1994) *Proc. Fifth Wld. Congr. Genet. Appl. Livest. Prod.* **17**:19.

Pattie, W.A. and Barlow, R. (1974) *Aust. J. Agric Res*., **29**:1077.

Saville, D.G. and Robards, G.E. (1972) *Aust. J. Agric Res*., **23**:117.

Tukey, J.W. (1949) *Biometrics*, **5**: 232.

Turner, Helen Newton, Dolling, C.H.S. and Kennedy, (1968) *Aust. J. Agric Res*., **19**: 79.

Williams, A.J. (1979). In "Physiological and Environmental Limitations to Wool Growth", ed. J. L. Black and P. J. Reis (The University of New England Publishing Unit, Armidale), pp. 337-354.

Woolaston, R.R. (1987) In "Merino Improvement Programs in Australia". (B.J.McGuirk, ed.), Australian Wool Corporation, Melbourne, Australia. pp. 421 – 435.

SIRE BY FLOCK-YEAR INTERACTIONS FOR BODY WEIGHT IN POLL DORSET SHEEP

D. J. Brown, A. A. Swan, D. J. Johnston and H.-U. Graser

Animal Genetics and Breeding Unit*, University of New England, Armidale, NSW 2351

SUMMARY

The performance of sires across flocks and years are likely to vary due to factors such as interactions with the environment, differences in ewe genotypes, errors in recording, non-random mating, and preferential treatment of progeny. Research in other breeds and species has shown that these sire by flock-year interactions typically account for less than 5% of the phenotypic variance. This paper examines the significance of sire by flock-year effects for weight traits in Poll Dorset sheep. The results demonstrate that while the sire by flock-year effects explained between 2 and 4% of the phenotypic variance, they significantly improved the fit of the model and resulted in a direct-maternal genetic correlation closer to zero. However, heritabilities were reduced significantly by up to 50%, indicating that sire by flock-year effects may be removing too much genetic variation for traits with maternal effects. On balance however, it seems advisable to include sire by flock-year interaction effects in the across flock evaluation.

INTRODUCTION

The sheep genetic evaluation system in Australia (Brown *et al.* 2007) contains performance data from flocks located across Australia. As the use of performance recording and Estimated Breeding Values (EBVs) increases, many sires are being used across many different environments and possibly very different ewe genotypes. It is possible that sires may perform differently across these flocks and years, requiring inclusion of sire by flock-year effects in evaluation models. Significant heterosis between Merino bloodlines has also been observed (Mortimer and Atkins 1987). It is possible that including sire by flock-year effects will remove some of the non-additive variance associated with heterosis or genotype by genotype interactions, to improve the reliability of EBVs.

Including a term for sire by herd interaction has been shown to significantly improve genetic analysis models for many traits in beef cattle (Robinson 1996; Lee and Pollak 1997; Meyer 1997; Meyer 2003; Dodenhoff *et al.* 1999) and to reduce the bias caused by unrecorded or inadequately recorded management and mating groups (Meyer 2003). Inclusion of sire by herd effects also resulted in reductions in the correlation between direct and maternal genetic effects (Meyer 1997; Berweger Baschnagel *et al.* 1999; Dodenhoff *et al.* 1999; Maniatis and Pollott 2002).

Accurate estimation of sire by flock variance requires sires, or related sires, to be used in multiple flocks. Confounding of sires or groups of sires within flocks (typically through using home-bred sires) makes it difficult to partition sire variance between the sires' main effects and interaction effects (Meyer 2003). However, in the terminal sire genetic evaluation program (LAMBPLAN) there is increasing information from sires used across flocks, allowing the effect to be investigated. Therefore the aim of this study was to examine the significance of sire by flock-year interactions on the body weight traits in industry recorded Poll Dorset sheep data.

MATERIALS AND METHODS

Data. Pedigree and performance data were extracted from the Sheep Genetics LAMBPLAN

* AGBU is a joint venture of NSW Department of Primary Industries and University of New England

database (Brown *et al.* 2007). This database consists of pedigree and performance records submitted by Australian ram breeders which are used for genetic evaluation purposes.

As the complete Poll Dorset database was too large for parameter estimation analyses, a subset of 39 of the best linked flocks was selected. Within these flocks, all straightbred Poll Dorset animals with full pedigree and born since 1995 were included. Data were extracted for weaning (Wwt), post weaning (Pwt) and yearling body weight (Ywt). All contemporary groups were transformed to a common mean of 35, 50 and 60kg for weaning, post weaning and yearling weight respectively, as is done routinely for Sheep Genetic analyses (Brown *et al.* 2007). A summary of the data used for each trait is shown in Table 1. The pedigree was built using 1 generation of ancestral information. Depending on the trait, this resulted in pedigree files comprising between 52,065 and 166,582 animals and data files comprising between 35,858 and 146,547 animals. The data originated from 39 flocks across 14 years. Only 7% of the animals studied had records on all three traits however 65% on the animals had weights record at 2 different ages.

Table 1. Summary of the data used in this study

Trait	Number of animals	Number of records	Number of sires	Number of dams	Sire Flock Yrs	Number of CGs	Mean	SD	Min	Max
Wwt	166,582	146,547	2,406	49,069	5,610	6,039	35.0	6.1	10.4	72.6
Pwt	159,083	138,514	2,463	49,226	5,726	4,320	50.0	6.6	14.9	89.1
Ywt	52,065	35,858	1,269	20,133	2,372	1,275	60.0	6.5	27.7	103.4

Models of analysis. Parameters were estimated in univariate analyses for each trait, fitting an animal model in ASReml (Gilmour *et al.* 2006). The base model (model 1) fitted the fixed effects of contemporary group, birth type, rearing type, age of dam, and animal's age at measurement. Contemporary group was defined as flock, year of birth, sex, date of measurement, management group subclass. Random effects fitted included correlated direct and maternal genetic effects and a maternal permanent environment effect. This is the model currently used in LAMBPLAN routine genetic evaluations for these traits. Model 2 included the additional random environmental effect of sire by flock-year (no relationships), with the two models compared using likelihood ratio tests.

To examine the impact of data structure, additional analyses were conducted using the same data sets with restrictions on which records were fitted with a sire by flock-year effect. To avoid the confounding caused by sires used within single flock-years, the additional analyses were conducted using the same data set and fixed effect models however the sire by flock-year effects were only fitted for records from progeny of sires used across multiple flock-years. Heterogeneous residuals were also fitted so that records without sire by flock-year fitted had a different residual variance to those with sire by flock-year effects fitted. The genetic parameters are presented for both records with and without the sire by flock-year fitted.

RESULTS AND DISCUSSION

Genetic parameters. The genetic parameters estimated for each model are summarised in Table 2. For all three traits there is a significant improvement in the log likelihood with the inclusion of sire by flock-year, although the effect only explained 2 to 4% of the phenotypic variance. The inclusion of the sire by flock-year effect also resulted in a significant reduction in the direct heritability and magnitude of the direct-maternal genetic correlation. For post weaning weight and yearling weight the direct-maternal correlation was not significantly different to zero. These observations agree with the studies of Meyer (1997), Berweger Baschnagel *et al.* (1997),

Dodenhoff *et al.* (1999) and Maniatis and Pollott (2002).

The reduction in the direct heritability is quite extreme for these traits, and appears to be related to the difficulty in partitioning variances in the presence of correlated maternal genetic effects. Results not presented for the same early weight traits in other LAMBPLAN breeds show very similar trends. By contrast, analyses for traits unaffected by maternal effects and measured on the Poll Dorset animals used in this paper (fat depth and eye muscle depth) show significant sire by flock-year interactions, but only marginal reduction in heritability. Anecdotally, it seems unlikely that the heritabilities for early weight traits are as low as shown in Table 2 given the improvement made in growth in this breed since the early 1990's. Meyer (2003) presented an alternative partitioning of variance for this problem which may give better results, but we are yet to implement the method in these data.

Table 2. Phenotypic variance (σ^2_p), direct (h^2_d) and maternal (h^2_m) heritabilities, direct-maternal genetic correlation (r_{am}), maternal permanent environmental (c^2) and sire by flock-year (s^2) effects as a proportion of phenotypic variance with and without sire by flock-year interactions included in the model (se in parentheses). LLR = difference in log likelihood relative to model 1.

Trait	Model	σ^2_p	h^2_d	h^2_m	c^2	r_{am}	s^2	LLR
Wwt	1	25.80 (0.11)	0.10 (0.01)	0.08 (0.01)	0.11 (0.00)	-0.36 (0.04)		0
Wwt	2	25.80 (0.11)	0.05 (0.01)	0.07 (0.01)	0.11 (0.00)	-0.18 (0.07)	0.02 (0.00)	163
Pwt	1	35.26 (0.16)	0.15 (0.01)	0.06 (0.01)	0.07 (0.00)	-0.27 (0.04)		0
Pwt	2	35.17 (0.16)	0.08 (0.01)	0.05 (0.01)	0.07 (0.00)	0.10 (0.08)	0.03 (0.00)	273
Ywt	1	37.86 (0.34)	0.23 (0.02)	0.08 (0.01)	0.06 (0.01)	-0.44 (0.06)		0
Ywt	2	37.69 (0.33)	0.12 (0.02)	0.05 (0.01)	0.06 (0.01)	-0.07 (0.14)	0.04 (0.00)	53

Data structure. In the data used for this study, 65 to 79% of the progeny with records are from sires used in more than one flock and year. On average sires have progeny in 2.3 flock years and 20% of progeny come from sires used across 10 or more flock years. The progeny that come from sires used only in 1 flock and year provide no information to estimate their sire by flock-year effects.

Table 3. Phenotypic variance (σ^2_p), direct (h^2_d) and maternal (h^2_m) heritabilities, direct-maternal genetic correlation (r_{am}), and sire by flock-year (s^2) effects with sire by flock-year effects only fitted for progeny of sires used across at least 2, 5 or 10 flock-years (FYs). Parameters reported for both records with and without sire by flock-year effects.

Trait	FYs	No sire by flock-year			With sire by flock-year					FYs[#]
		σ^2_p	h^2_d	h^2_m	σ^2_p	h^2_d	h^2_m	r_{am}	s^2	%
Wwt	2	25.20 (0.20)	0.07 (0.01)	0.08 (0.01)	26.08 (0.13)	0.07 (0.01)	0.07 (0.01)	-0.30 (0.06)	0.02 (0.00)	79%
Wwt	5	25.18 (0.13)	0.09 (0.01)	0.08 (0.01)	27.07 (0.17)	0.08 (0.01)	0.07 (0.01)	-0.36 (0.05)	0.03 (0.00)	39%
Wwt	10	25.34 (0.12)	0.10 (0.01)	0.08 (0.01)	28.59 (0.26)	0.09 (0.01)	0.07 (0.01)	-0.36 (0.05)	0.03 (0.00)	16%
Pwt	2	35.27 (0.29)	0.11 (0.01)	0.05 (0.01)	35.53 (0.18)	0.11 (0.01)	0.05 (0.01)	-0.09 (0.06)	0.03 (0.00)	79%
Pwt	5	34.36 (0.18)	0.14 (0.01)	0.06 (0.01)	37.37 (0.25)	0.13 (0.01)	0.06 (0.01)	-0.22 (0.05)	0.04 (0.00)	39%
Pwt	10	34.84 (0.17)	0.15 (0.01)	0.06 (0.01)	38.53 (0.36)	0.13 (0.01)	0.06 (0.01)	-0.26 (0.04)	0.03 (0.00)	18%
Ywt	2	37.61 (0.49)	0.21 (0.02)	0.07 (0.01)	38.11 (0.41)	0.20 (0.02)	0.07 (0.01)	-0.37 (0.08)	0.03 (0.00)	65%
Ywt	5	37.08 (0.35)	0.24 (0.02)	0.08 (0.01)	41.99 (0.78)	0.21 (0.02)	0.07 (0.01)	-0.43 (0.07)	0.02 (0.01)	17%
Ywt	10	37.61 (0.34)	0.24 (0.02)	0.08 (0.01)	41.54 (1.15)	0.21 (0.02)	0.07 (0.01)	-0.44 (0.06)	0.01 (0.01)	7%

[#] Proportion of records with the sire by flock-year effects fitted

To separate the influences of these within-flock-year sires additional analyses of the same data were conducted in which sire by flock-year effects were only fitted for animals from sires used across multiple flock-years. The genetic parameters from these analyses are show in Table 3. Fitting effects only for progeny of sires used across flock-years resulted in less reduction in the heritability and direct-maternal genetic correlation. The remaining genetic parameters remained relatively constant between analyses.

With the exception of the phenotypic variance the genetic parameters from the models with sire by flock-year effects only fitted to progeny from sires used across 10 or more flock-years are the same as the model with no sire by flock-year effects included (Model 1 in Table 2). The phenotypic variance for animals with sire by flock-year effects in Table 3 increases as the number of flock-years increases from 2 to 10. As fewer animals receive the sire by flock-year effect the direct-maternal correlation becomes more negative in Table 3, approaching the estimate for Model 1 in Table 2.

In fitting sire by flock-year interaction effects we are implicitly estimating an average breeding value across flocks, with deviations from this average within individual flock-years removed by the interaction term. This is desirable when the deviations are caused by factors such as recording errors, non-random mating, or preferential treatment of progeny, but it may not be desirable in other situations in which useful genetic variation is removed by fitting the interaction term. The difficulty is that it is not possible to separate these two cases. On balance however, given that the goal of LAMBPLAN is to evaluate the across flock merit of animals, it seems appropriate to include sire by flock-year effects in the evaluation model.

CONCLUSIONS

Sire by flock-year effects are in a sense a "necessary evil" in LAMBPLAN genetic evaluations. On the one hand they are useful in dealing with data quality problems across flocks and years, but on the other, they may be removing too much genetic variation. This seems to be particularly the case for traits which are modelled with correlated direct and maternal genetic effects. More research is needed in this area, possibly using estimation models which partition the variation differently, but ideally using better structured data. One possibility will be to base studies on the Sheep CRC Information Nucleus flock, where all sires involved are being used at all sites.

ACKNOWLEDGMENTS

This research is funded by Meat and Livestock Australia and Australian Wool Innovation, and Sheep Genetics is made possible through the support of the Australian sheep industry.

REFERENCES

Berweger Baschnagel, M., Moll, J. and Kunzi, H. (1999) *Liv. Prod. Sci* **60:**203.

Brown, D.J., Huisman, A.E., Swan, A.A., Graser, H-U., Woolaston, R.R., Ball, A.J., Atkins, K.D. and Banks, R.B. (2007) *Proc. Ass. Adv. Anim. Breed. Genet.* **17**:187.

Dodenhoff, J., Van Vleck, L.D. and Wilson, D.E. (1999) *Anim. Sci.* **77(12):**3176.

Gilmour, A.R., Gogel, B.J., Cullis, B.R., and Thompson, R. (2006) ASReml User Guide Release 2.0 VSN International Ltd, Hemel Hempstead, HP1 1ES, UK.

Maniatis, N. and Pollott, G.E. (2002) *Anim. Sci.* **75:**3.

Meyer, K. (1997) *Liv. Prod. Sci* **52(3):**187.

Meyer, K. (2003) *Proc. Aust. Ass. Anim. Breed. Genet.* **15**:131.

Mortimer, S.I. and Atkins, K.D. (1987) *Wool Tech. Sheep Breed.* **35(4):**208.

Lee, C. and Pollak, E.J. (1997) *J. Anim. Sc.* **75:**68.

Robinson, D.L. (1996) *Liv. Prod. Sci.* **45:**111.

CLASSIFYING SHEEP GRAZING ENVIRONMENTS USING SATELLITE DATA TO QUANTIFY GENOTYPE BY ENVIRONMENT INTERACTIONS

M. B. Whelan[1], D. J. Cottle[2], K. G. Geenty[3] and D. J. Brown[4]

[1]Southern Cross University, Lismore, NSW 2480
[2]University of New England, Armidale, NSW 2351
[3]CRC for Sheep Industry Innovation, University of New England, Armidale, NSW 2351
[4]Animal Genetics and Breeding Unit*, University of New England, Armidale, NSW 2351

SUMMARY

Australian sheep grazing environments are currently classified into 3 very broad zones (High Rainfall, Wheat/Sheep and Pastoral) that do not differentiate sheep grazing environments to a level allowing sheep producers to assess the impact grazing environments may have on sire progeny performance. If a genotype by environment interaction (GEI) is expressed more as environments diverge then a finer classification of environments may help breeders when selecting stud rams. A sheep grazing environment classification system has been developed in this study using readily-obtainable monthly Normalised Difference Vegetation Index (NDVI), from satellite data, and monthly maximum temperature for a 10 year period.

Cluster analysis was used on the NVDI and temperature data to create 25 sheep grazing environment classes (SGEclass) around Australia. Two-way analysis of variance revealed a significant interaction between sire progeny performance and SGEclass for hogget weight, fibre diameter and greasy fleece weight. Further ASReml analysis of Merino data from Sheep Genetics illustrated that sire by SGEclass explained similar amounts of variation as sire by flock. Recording the geographic location of the flock would improve the ability to account for environmental differences between flocks.

INTRODUCTION

The Sheep Genetics (SG) database contains records for approximately 2 million sheep and is used by ram breeders and commercial producers throughout Australia (Sheep Genetics 2008). SG offers breeders and commercial growers information about the breeding value of sheep from all sheep growing regions. However, selecting sires bred and evaluated in environments that differ greatly from where their progeny will be farmed increases the chances of GEI being expressed for some traits (Amores *et al.* 1999; Brown *et al.* 1999; Carrick 2005). On the other hand, ignoring superior sires because of an expectation of GEI may be reducing productivity on some farms (Kaine *et al.* 2002).

Many researchers consider that GEI is not uniform for all traits (Atkins *et al.* 1998; Carrick 2005). SG accommodates a large number of traits and using many of these in a breeding objective increases the likelihood of GEI. Therefore, classifying environments for a national breeding system will improve the effectiveness of breeding values because environment can be included as a fixed effect in the model. Providing a mechanism that enables breeders and commercial growers to quantify GEI will enable them to be more objective in selecting sires.

Classification of farming land has been used to help quantify GEI in beef cattle (Bertrand *et al.* 1985) and dairy cattle (Weigel and Rekaya 2000). Bertrand *et al.* (1985) considered that predicted breeding values of beef sires would be less accurate if a significant sire x environment interaction existed. The present ABARE division of grazing environments into High Rainfall, Wheat/Sheep

* AGBU is a joint venture of NSW Department of Primary Industries and the University of New England

and Pastoral zones is probably too broad to be useful in describing environments in the context of GEI. The aim of this study was to evaluate a sheep grazing environment classification derived from satellite data to determine if it could be successfully used as a fixed factor representing grazing environments in the SG database.

MATERIALS AND METHODS

The Sheep Grazing Environment classes (SGEclasses) were derived by combining temperature data, as a surrogate for pasture quality, (Wilson and Ford 1972) and the normalised difference vegetation index (NDVI) derived from National Oceanic and Atmospheric Administration, Advanced Very High Resolution Radiometer (NOOA AVHRR) data as a measure of pasture quantity (Hill *et al.* 1998). These data were recorded monthly for each cell of a 10km by 10km grid of Australia for a 10 year period (1996 to 2005). Cluster analysis of the resulting image (120 layers of NDVI/temperature data) was used to generate 25 SGEclasses. The SGEclasses formed a mosaic of grazing environments across Australia.

The SG data were collected without any geographic information. The postcodes recorded as part of the property managers' addresses were used to locate SG sites. A SGEclass was allocated to each postcode by overlaying the postcode coverage on a map of SGEclasses.

Preliminary analysis. Preliminary analysis was conducted by filtering the data to select sires whose progeny were represented in 5 SGEclasses. The same sires could not be evaluated for all traits. The performance of progeny was based on hogget weight (4 sires, 1265 progeny), fibre diameter (5 sires, 1123 progeny) and fleece weight (2 sires, 1383 progeny) and were compared in 5 SGEclasses. Two-way analysis of variance, with sire and SGEclass as main effects and Sire x SGEclass as an interaction term, was used to determine if an interaction between sire progeny performance and SGEclasses existed. The significance ($P<0.05$) of the interaction was used to determine the need for further analysis.

Analysis of Sheep Genetics data. Data for yearling body weight, greasy fleece weight and fibre diameter were extracted from the MERINOSELECT database (Brown *et al.* 2007) for animals born in 2000 and later to conduct a more detailed analysis. The dataset contained 168,685 animals with pedigree information and 77,744, 70,461 and 80,337 animals with records for body weight (WT), greasy fleece weight (GFW) and fibre diameter (FD) respectively. The 95 flocks with data were grouped into 12 SGEclasses based on the post code of the flock. Three models were analysed for each trait; 1) normal Sheep Genetics model; 2) a model with sire by flock interactions fitted as an additional random effect; and 3) a model with sire by SGEclass interactions fitted as an additional random effect. A more detailed description of the models for analyses can be found in Brown *et al.* (2009).

RESULTS AND DISCUSSION

Preliminary results. Two-way analysis of variance revealed a significant interaction ($P<0.001$) between sire and SGEclasses for all traits. Figure 1 illustrates the differences in progeny performance (hogget weight) in the different environments. Similar differences were found for GFW and FD. SGEclass 25 represents the best grazing environment while class 16 represents the worst grazing environment illustrated with the other classes falling between these extremes. However, the gradation between the classes is not scaled.

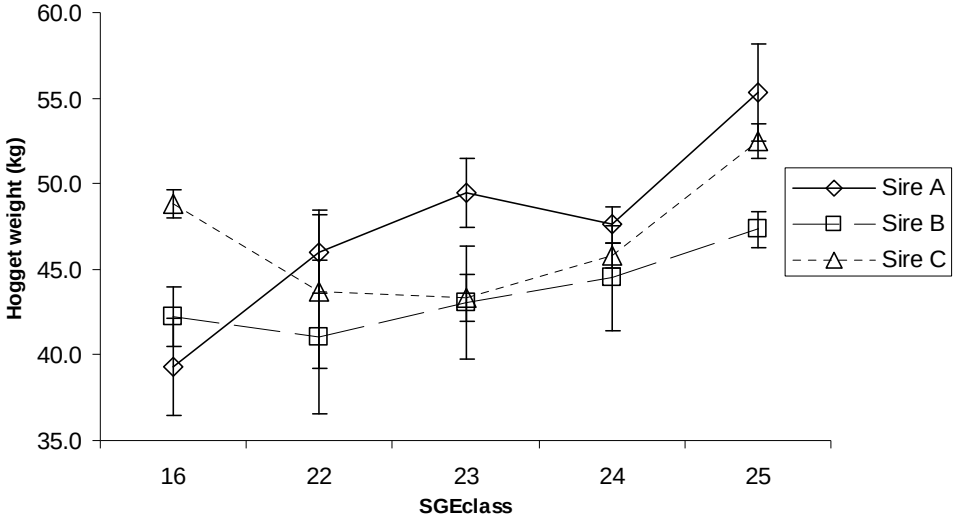


Figure 1. Hogget weight of sire progeny in 5 SGEclasses with 95% confidence interval.

Sheep Genetics data. The preliminary analysis was restricted to a small number of sires represented in 5 SGEclasses. The ASReml analysis included a much larger number of observations of progeny performance and took into account genetic and maternal effects and included a larger number of sires from the SG database. The significant interaction between sire and SGEclass meant that the more robust analysis using ASReml was warranted.

Table 1. Phenotypic variance (σ^2_p), direct (h^2_d) and maternal (h^2_m) heritabilities, maternal permanent environmental (c^2) and sire by flock-year (Model 2) or sire by SGEclass (Model 3) (s^2) effects as a proportion of phenotypic variance with and without sire by flock-year or SGEclass interactions included in the model (se in parentheses)

Trait	Model	σ^2_p	h^2_d	h^2_m	c^2	s^2	LogL
Wt	1	34.68 (0.23)	0.44 (0.01)	0.03 (0.01)	0.02 (0.01)		-995.3
Wt	2	35.29 (0.26)	0.38 (0.02)	0.03 (0.01)	0.03 (0.01)	0.05 (0.00)	-859.3
Wt	3	35.32 (0.26)	0.39 (0.02)	0.02 (0.01)	0.03 (0.01)	0.04 (0.00)	-891.5
Gfw	1	0.41 (0.00)	0.36 (0.01)	0.03 (0.01)	0.03 (0.01)		-1841.9
Gfw	2	0.42 (0.00)	0.25 (0.02)	0.02 (0.01)	0.04 (0.01)	0.08 (0.01)	-1574.5
Gfw	3	0.42 (0.00)	0.25 (0.02)	0.02 (0.01)	0.04 (0.01)	0.08 (0.01)	-1627.0
FD	1	1.34 (0.01)	0.61 (0.01)				-5530.6
FD	2	1.35 (0.01)	0.58 (0.01)			0.03 (0.00)	-5470.4
FD	3	1.35 (0.01)	0.58 (0.01)			0.03 (0.00)	-5476.0

Sire by flock (Model 2) had the greatest improvement in Log likelihood of the 3 models, however sire by SGEclass (Model 3) explained similar amounts of phenotypic variation as sire by flock (Table 1). This suggests that the SGEclass category may provide a simpler way of accounting for environmental variance and GEI in SG analyses than fitting flock effects, as there are much fewer SGEclasses than there are flocks.

To accurately estimate interactions of sire across flocks or SGEclasses adequate representation of sire across different flocks or SGEclasses is essential however in these data 25% of all progeny were from sires used across more than 1 flock, while 21% of all progeny came from sires used across SGEclasses.

CONCLUSIONS

When breeders choose a sire to improve performance, the flock-year term in SG analyses represents the grazing/management environment of the sire's progeny, the effect of ewes on production traits and other factors (e.g. incorrect data recording, preferential treatment of link sires and heterosis). Recent research has indicated that fitting an additional random interaction term for sires used across flocks and years improved the accuracy of breeding values. The results presented here demonstrate that it may be possible to replace the flock-year term with SGEclasses. Given that the SGEclasses are based on 10 years of pasture availability (NDVI) and quality (temperature) it may provide a more usable measure of the grazing environment than the many flock-year effects. The flock in the SG database represents a single geographic point in sheep grazing areas of Australia. The SGE classification provides a method of clustering flocks into groups that have similar grazing environments. These groupings may give breeders more confidence when selecting sires from other flocks within the same SGEclass or the ability to determine the level of risk of selecting a desirable sire with performance records from a different SGEclass.

In this study the geographic location of a property was based on the postcode of the contact address and not the location of the property. Many flocks had to be removed from the analysis because the postcodes were not in sheep grazing areas or the area covered by a single postcode was represented by a number of SGEclasses. Collecting GPS coordinates of the flocks would enhance the reliability of the allocation of flocks to the SGEclasses.

REFERENCES

Amores, B.G., Hinch, G.N., Mortimer, S.I. and Sivarajasingam, S. (1999) *Proc. Ass. Adv. Anim. Breed. Genet.* **13**:110.

Atkins, K.D., Gilmour, A.R., Thompson, R., Coelli, K.A. and Casey, A.E. (1998) *Wool Tech. Sheep Breed.* **46**:1.

Bertrand, J.K., Berger, P.J. and Willham, R.L. (1985) *J. Anim. Sci.* **60**:1396.

Brown, D.J., Crook, B.J. and Purvis, I.W. (1999) *Proc. Ass. Adv. An. Breed. Gen.* **13**:274.

Brown, D.J., Huisman, A.E., Swan, A.A., Graser, H-U., Woolaston, R.R., Ball, A.J., Atkins, K.D. and Banks, R.B. (2007) *Proc. Ass. Adv. Anim. Breed. Genet.* **17**:187.

Brown, D.J., Swan, A.A. and Graser, H-U. (2009) *Proc. Ass. Adv. Anim. Breed. Genet.* **18**, (these proceedings)

Carrick, M. (2005) "Genotype by environment interaction in wool sheep". Sydney: Meat and Livestock Australia.

Hill, M.J., Donald, G.E., Wheaton, G.A.B., Hyder, M.W., Plaisted, T.W. and Smith, R.G.C. (1998) *9th Aus. Rem. Sens. Photo. Conf. Sydney.*

Kaine, G., Court, J. and Niall, E. (2002) *Wool Tech. Sheep Breed.* **50**:423.

Sheep Genetics (2008) *Sheep Genetics,* from http://www.sheepgenetics.org.au/index.html, Date Accessed: 13/2/2008

Weigel, K.A. and Rekaya, R. (2000) *J. Dairy Sci.* **83**:815.

Wilson, J.R. and Ford, C.W. (1972) *Aust. J Agric. Res*. **24**:187.

GENOTYPE BY ENVIRONMENT INTERACTION BETWEEN REGISTERED AND COMMERCIAL HERDS FOR DAIRY TRAITS IN AUSTRALIA

M. Haile-Mariam[1] and M.E. Goddard[1,2]

[1]Biosciences Research Division, DPI Victoria, 1 Park Drive, Bundoora 3083, Australia
[2]Faculty of Land and Food Resources, University of Melbourne, Parkville, Vic 3052 Australia.

SUMMARY

The presence of genotype by environment interaction (GxE) between registered and commercial herds for dairy traits was investigated using Australian Holstein-Friesian data. The traits studied were milk yield, fertility, some type and workability traits and survival. To determine the importance of GxE the same traits recorded in registered and commercial herds were treated as different traits and genetic correlations between them were estimated using a bi-variate sire model. For most traits, genetic correlations between registry statuses were above 0.92 suggesting limited bull re-ranking as a result of having daughters in registered or commercial herds. Genetic correlations were slightly lower between commercial and registered herds for calving interval (0.92), likeability (0.88) and temperament (0.87). Heritability of likeability in registered herds was about 3 times that in commercial herds. The reason for the lower genetic correlations between registered and commercial herds for likeability and temperament may be that owners of registered herds considered the ancestry of cows and other attributes such as milk yield when scoring cows for these traits. For most traits current genetic evaluations using data from registered and commercial herds are adequate and better definition of the workability traits should help to further minimise GxE.

INTRODUCTION

The goal of a national genetic evaluation system is to identify sires that are consistently superior across all environments (type of herds) for economic traits. A recent analysis of milk production, survival and fertility data showed that bull re-ranking was significant but small when environments were defined by environmental descriptors such as herd size, average production level, regions and calving system in Australia (Hayes *et al*. 2003; Haile-Mariam *et al*. 2008). Another reason for possible genotype by environment interaction (GxE) could be the registry status of herds. The herd management and production objective of registered herds could be different from those of commercial herds. Culling and selection decisions could vary between registered and commercial herds (eg Dekkers *et al*. 1994). In registered herds cows with good type and ancestry may be valued more than those with good milk yield or fertility. These differences could lead to re-ranking of bulls for some traits. Currently the genetic evaluation for all dairy traits does not consider registry status of animals. In the Australian Holstein-Friesian (HF) breed the proportion of registered animals in the total population, which is about 29% (Monro 2004) may not be large, but their contribution to genetic progress could be higher than expected from their population size. The effect of GxE on the accuracy of selection and overall efficiency of selection schemes is documented (Mulder and Bijma 2005). Uncertainties about the possible impact of GxE on bull ranking can reduce the creditability of the genetic evaluation systems and the use of the results. Therefore, testing for the presence of GxE and informing farmers about its effect could help to increase farmers' confidence in the evaluation system. Knowing the extent of GxE could also lead to appropriate decisions to calculate ABVs. The objective of this study was therefore to test for possible GxE between registered and commercial herds by estimating genetic parameters.

MATERIALS AND METHODS

Calving and survival data of HF cows were extracted from the Australian Dairy Herd Improvement Scheme (ADHIS) database. Details of the milk yield and calving data used for this study are given by Haile-Mariam *et al*. (2008). These data were merged with either type or workability data. Workability traits (temperament, likeability and milking speed) are scored by farmers from 1 to 5 with the most docile, most liked and fastest milking cows given a score of 1. Type traits are scored by classifiers of the HF breed. Of the type traits, overall type and mammary system are composite traits scored from 1 to 15 but the rest are linear traits scored 1 to 9. Classifications before 18 months of age and after 45 months were excluded. Stage of lactation at classification between day 6 and day 365 was included. Lists of herds that are reported as registered or commercial were also provided by ADHIS. The commercial herds that are included are those that had data on either workability or type traits. Table 1 shows the structure of the data.

Table 1. Number of herds, number of cows and mean (standard deviation) of the traits for registered and commercial herds in Australia

Traits	Registered			Commercial		
	No. herds	No. cows	Mean(SD)	No. herds	No. cows	Mean(SD)
Milk yield, kg	853	63000	5946(1452)	3786	188441	4990(1284)
Fat yield, kg	853	63000	222(53)	3786	188642	192(47)
Protein yield, kg	853	63000	187(47)	3786	188565	156(41)
Survival, %	874	74306	84(36)	3808	213530	84(36)
Calving interval, days	865	58691	416(87)	3672	164781	402(79)
Likeability	853	63088	2.39(0.90)	3787	188642	2.45(0.85)
Temperament	853	63088	2.40(0.89)	3787	188642	2.45(0.86)
Milking speed	853	63088	2.49(0.87)	3787	188642	2.54(0.83)
Overall type	906	65953	9.98(1.66)	2386	59197	8.93(1.84)
Mammary system	906	65954	10.28(1.66)	2386	59193	9.31(1.82)
Udder depth	906	65888	5.87(1.13)	2385	59111	5.87(1.21)
Pin set	906	65952	3.90(1.33)	2386	59194	3.81(1.38)
Foot angle	906	65888	4.96(1.09)	2385	59114	4.77(1.09)
Angularity	906	65952	5.62(1.17)	2385	59193	5.44(1.18)
Body depth	906	65887	6.34(1.10)	2385	59110	6.10(1.16)
Udder texture	906	65952	6.19(1.15)	2385	59189	5.88(1.20)

All the available pedigree data of sires of the cows with records and their ancestors as far back as 1950 were included in the pedigree. The number of sires with progeny varied from 4320 for workability traits to 7611 for survival. The proportion of bulls commonly used in both type of herds varied from 54% for type traits to 75% for survival (of all bulls). Traits analysed were milk yield traits, survival, calving interval (CI), workability traits, and a selected number of type traits. All data were analysed fitting fixed effects such as herd-year-season of calving, age at calving, month of calving and the random effect of sires. In the case of type traits, the fixed effects fitted were age at classification, stage of lactation (days in milk) at classification, month of calving and herd-classifier-round. Genotype by registry status (registered vs. commercial herds) interaction was examined by treating performance recorded in registered herds as trait 1 and that recorded in commercial herds as trait 2. To test the significance of GxE, the log likelihood of a full model was compared to log likelihood of a model in which the genetic correlation between registry statuses was fixed at unity. A χ^2 test with 1 degree of freedom was used to test for the significance of GxE. Data analyses performed using ASReml (Gilmour *et al.* 2006).

RESULTS AND DISCUSSION

Mean and standard deviation for milk yield traits and CI were higher in registered herds than in commercial herds (Table 1). The means for most type traits were slightly higher in registered herds than in commercial herds (Table 1). On the other hand means for workability traits were slightly higher in commercial herds than in registered herds (Table 1). Differences in mean and variance in milk yield traits and CI between registered and commercial herds were similar to the difference between year-round and seasonal calving herds observed in Australian cows (Haile-Mariam *et al*. 2008). Based on means, standard deviation and phenotypic variance for milk yield traits and CI registered herds are more similar to year-round calving herds and commercial herds are similar to seasonal calving herds. Differences in mean milk yield traits between registered and commercial herds appear to be higher in the current data than reported by others (Powell and Norman 1986).

Phenotypic variances were higher in registered than in commercial herds for milk yield traits, survival, CI and workability traits (Table 2). The difference in phenotypic variance between registered and commercial herds observed in the current study is of the same magnitude as that observed between year-round and seasonal calving herds for milk yield traits, survival and CI (Haile-Mariam *et al*. 2008). Differences in heritability estimates were small but were generally higher in registered herds than in commercial herds for milk yield and most type traits (Table 2). Heritabilities of workability traits in registered herds were higher than in commercial herds. The biggest difference was for likeability which is a catch-all trait followed by temperament which is also closely associated with likeability because the most docile cows are also the most liked.

Table 2. Heritabilities and genetic correlations for milk yield, fat yield, protein yield, survival, calving interval and for workability and some type traits of registered and commercial herds

Traits	Registered	Commercial	Genetic correlation
	Heritability	Heritability	
Milk yield, kg	0.31 ± 0.02 (823900)[A]	0.29 ± 0.01 (639000)[A]	0.96 ± 0.01
Fat yield, kg	0.24 ± 0.02 (1212)	0.22 ± 0.01 (968)	0.98 ± 0.01
Protein yield, kg	0.25 ± 0.02 (778)	0.22 ± 0.01 (599)	0.96 ± 0.01
Survival, %	0.03 ± 0.01 (1237)	0.03 ± 0.0[B] (1190)	0.98 ± 0.04
Calving interval, days	0.03 ± 0.01 (6049)	0.03 ± 0.0[B] (4740)	0.92 ± 0.06
Likeability	0.29 ± 0.02 (0.66)	0.11 ± 0.01 (0.56)	0.88 ± 0.03[C]
Temperament	0.20 ± 0.02 (0.63)	0.09 ± 0.01 (0.56)	0.87 ± 0.03[C]
Milking speed	0.24 ± 0.02 (0.60)	0.14 ± 0.01 (0.50)	0.96 ± 0.02
Overall type	0.19 ± 0.02 (2.14)	0.17 ± 0.01 (2.47)	0.96 ± 0.02
Mammary system	0.22 ± 0.02 (2.22)	0.18 ± 0.01 (2.52)	0.99 ± 0.01
Udder depth	0.33 ± 0.02 (0.99)	0.33 ± 0.02 (1.11)	0.97 ± 0.01
Pin set	0.31 ± 0.02 (1.52)	0.32 ± 0.02 (1.60)	0.99 ± 0.01
Foot angle	0.15 ± 0.01 (0.94)	0.11 ± 0.01 (1.0)	0.99 ± 0.01
Angularity	0.21 ± 0.02 (1.03)	0.20 ± 0.02 (1.13)	0.98 ± 0.02
Body depth	0.32 ± 0.02 (0.88)	0.30 ± 0.01 (1.02)	0.98 ± 0.01
Udder texture	0.19 ± 0.02 (1.11)	0.15 ± 0.01 (1.21)	0.99 ± 0.01

[A]Phenotypic variance; [B]Standard error rounded to zero; [C]Significantly different from unity ($P < 0.05$).

Genetic correlations between commercial and registered herds for all traits were close to unity except for likeability, temperament and CI. Compared to results in Table 2 Haile-Mariam *et al.* (2008) estimated a genetic correlation of 0.83 for CI between split and year-round calving herds. For milk yield traits the lowest genetic correlation of 0.9 was estimated for fat yield between region 1 (New South Wales, Queensland, South and West Australia) and region 3 (Gippsland and

Tasmania). Genetic correlations for milk yield traits between calving systems were all higher than 0.9 (Haile-Mariam *et al.* 2008).

For subjectively scored traits such as likeability and temperament differences in heritability as well as genetic correlations below unity could be related to difference in perceptions and preferences of herd owners of registered and commercial herds (Dekkers *et al*. 1994). The way these two traits are perceived may vary between the two groups of herd owners. According to Beard and Jones (1991) the score for likeability was the response to the question: "all things being equal would you like more cows like this one in your herd? ". This could be perceived to include ancestry of the cow by breeders of registered cows and could mean that cows from a certain group of sires may be liked more than those from other groups of sires whereas scoring in commercial herds may not be associated with ancestry. The reason for the difference in heritability between registered and commercial herds and the low genetic correlation (thus possible GxE) for temperament may be related to its high correlation (~0.85) with likeability. The higher heritabilities for workability traits in registered than in commercial herds could also be because cows in registered herds are more consistently and accurately scored. Another additional reason could be that likeability in one group of herds is more closely associated with other highly heritable traits than in the other group of herds. To test this further we estimated genetic and residual correlations of likeability with milk yield. The result showed that genetic (-0.62 vs. -0.58) and residual (-0.30 vs. -0.28) correlations in registered herds between likeability and milk yield were slightly stronger than in commercial herds (results not tabulated elsewhere). This suggests that perhaps milk yield is considered as an additional criterion when scoring for likeability to a greater extent in registered herds than in commercial herds.

According to Robertson (1959) the impact of GxE is economically important if the genetic correlation of a trait expressed in different environments falls below 0.8. In this study genetic correlations were higher than 0.86 showing limited amount of sire re-ranking. These genetic correlations could be increased or GxE could be reduced further if both likeability and temperament are better defined. Genetic correlations of most traits recorded in registered and commercial herds were closer to unity than observed between Australian regions or calving systems (Haile-Mariam *et al.* 2008).

In conclusion, the genetic correlations between registered and commercial herds even for likeability and temperament are high enough and should not be of concern to the dairy industry. This means owners of both registered and commercial herds should have confidence in the appropriateness of the current genetic evaluation systems. However, better definition of the traits and helping producers to standardising the scoring should be useful.

ACKNOWLEDGEMENTS

This research is funded by Dairy Australia and Victorian Department of Primary Industries. We thank ADHIS for providing the data and Dr Kevin Beard for data extraction.

REFERENCES

Beard, K. and Jones, L. (1991) *Proc. Aust. Assoc. Anim. Breed. Genet.* **7:**425.
Dekkers, J.C.M., Jairath, L. K, and Lawrence, B. H. (1994*) J. Dairy Sci*. **77**:844.
Gilmour, A.R., Gogel, B.J., Cullis, B.R. and Thompson, R. (2006) ASReml User Guide Release 2.0 VSN International Ltd, Hemel Hempstead, HP1 1ES, UK.
Haile-Mariam, M., Carrick, M. and Goddard, M.E. (2008*) J. Dairy Sci*. **91**: 4840.
Hayes, B. J., Carrick, M., Bowman, P. and Goddard, M.E. (2003*) J. Dairy Sci*. **86**:3736.
Monro, G. (2004) In **11**th World Holstein Conference, Paris, France.
Mulder, H. A., and Bijma. P. (2005) *J. Anim. Sci*. **83**:49.
Robertson, A. (1959) *Biometrics* **15**:469.

GENOTYPE BY ENVIRONMENTAL INTERACTION FOR LIVE WEIGHT BETWEEN TWO PRODUCTION ENVIRONMENTS IN THE GIFT STRAIN (NILE TILAPIA, *OREOCHROMIS NILOTICUS*)

H.L. Khaw[1], R.W. Ponzoni[1], A. Hamzah[2] and N. Kamaruzzaman[1]

[1] The WorldFish Center, Jalan Batu Maung, 11960 Bayan Lepas, Penang, Malaysia
[2] National Prawn Fry Production and Research Center, Kg. P. Sayak, 08500 Kedah, Malaysia

SUMMARY

A genotype by environmental interaction study was conducted using the live weight data collected from three discrete spawning seasons of Genetically Improved Farmed Tilapia selective breeding program in Malaysia. Two production environments were used to grow-out the progeny, namely, cages and ponds. The analysis was carried by using animal mixed model and treating live weight in cages and in ponds as two different traits to determine the genetic correlation, which was used to quantify the genotype by environmental interaction for these two environments. The heritabilities estimated from the animal variance component were 0.34±0.061 and 0.40±0.067, for cages and ponds, respectively. The genetic correlation between live weight in cages and ponds was 0.75±0.091. Responses to selection were separately estimated for live weight in these two environments, and were 18.6% in cages and 15.3% in ponds after two generations of selection. Based on these results, we conclude that selection response was being achieved in both environments and that, despite the presence of a non-unity genetic correlation between live weight in cages and ponds, there was no significant evidence for genotype by environmental interaction for these two main aquaculture systems in Malaysia.

INTRODUCTION

The two main culture systems in Malaysia for tilapia farming are cage and pond (Hanafi and Chua 2008). Due to the rich natural resources in Malaysia, cage culture system is considered more economic in terms of land used and management compared to the pond system (Hanafi and Chua 2008). In Asia most of the selective breeding programs for Nile tilapia were conducted under intensive pond culture system, including the GIFT (Genetically Improved Farmed Tilapia) breeding program by the WorldFish Center in Malaysia (Eknath *et al.* 1993; Bolivar 1998; Eknath and Acosta 1998; Tayamen 2004; Ponzoni *et al.* 2005). This situation raises the issue of genotype by environment interaction (GxE) between pond and cage culture systems. It is important to know if genetic gain is being realised even if fish are grown in an environment that is different from that in which selection is taking place.

In the context of animal breeding, GxE describes the situation where different genotypes do not respond in the same way to different environments, so that the genetic and environmental effects are not additive. In order to examine the issue of genotype by environment interaction, we treated live weight at harvest in each culture environment as two different traits (Falconer 1952). The objectives of this study were to estimate the genetic parameters for live weight expressed in cage and pond environments, to evaluate the response to selection in both environments, and to determine the genotype by environment interaction between cage and pond culture systems.

MATERIALS AND METHODS

The data used in this study consisted of 10,065 fish with phenotype from three discrete generations (spawning season 2002 to 2004) of the GIFT selective breeding program in Malaysia. The breeding program had two lines: a selection line (selected for high breeding value for live weight) and a control line (selected on average breeding values for live weight). There were a

total of 177 sires and 244 dams involved in both lines. The details of base population, mating scheme and selection method are described in Ponzoni *et al.* (2005).

For the GxE study, the individually tagged fingerlings from each full-sib family were separated into two groups of equal size and grown out either in cages or earthen ponds. For the cage environment, the fingerlings were reared at cages of size 3m x 3m, with initial stocking density of 55 fish per square meter of water surface. For the pond environment, earthen ponds of 0.1 hectare were used and the initial stocking density was three to four fish per meter square. At both environments, the fish were fed twice a day with commercial dry pellet feed that contained 32 percent of protein.

The fish in both environments were harvested after approximately 120 days of grow-out. At harvest, the fish were recorded for live weight (grams), standard length (cm), width (cm), depth (cm) and sex. Based on the spawning date and harvesting date, the age (in days) of each fish was computed. Only the results corresponding to the GxE for live weight are presented in this paper.

The phenotypic and genetic parameters for live weight (square root transformed) were estimated using ASReml (Gilmour *et al.* 2002). In order to quantify the GxE between cage and pond environments, the genetic correlation was estimated by treating live weight in cage and pond as two different traits in a bivariate analysis. A mixed model was fitted to the data, with spawning season, line and sex as fixed effects, and animal and dam (solely accounting for the maternal and common environmental effect on the progeny, without a genetic structure) were fitted as random effects. Age was fitted as a covariate with the spline function available in ASReml.

The selection responses for both environments were calculated based on the average estimated breeding values by line and by generation, and it was expressed as a percentage of the least square mean on control line.

RESULTS AND DISCUSSION

The number of observations, simple means, minimum and maximum, standard deviation and coefficient of variation values for body weight and age at harvest in the cage and pond environments are presented in Table 1. The mean weight for the fish grown in ponds was greater than that for fish grown in cages.

Table 1. Descriptive statistics for live weight (g) and age (days) at harvest in cage and pond

Variable	Environment	N	Mean	Minimum	Maximum	Standard deviation	Coefficient of variation (%)
Live weight	Cage	5086	146.5	13	591	77.8	53
	Pond	4979	223.0	7	682	104.4	47
Age at harvest	Cage	5086	240	151	289	27.5	11
	Pond	4979	230	125	302	32.7	14

Table 2. Analysis of variance of live weight ($g^{0.5}$) **in cage and pond: Tests of fixed effects using PROC MIXED (SAS Institute Inc. 1997)**

Effect	Cage		Pond	
	F value	Prob. > F	F value	Prob. > F
Spawning season (SS)	65.22	< 0.0001	56.21	< 0.0001
Line (L)	13.93	0.0002	24.52	< 0.0001
Sex (S)	498.16	< 0.0001	440.65	< 0.0001
SS x S x L	3.91	0.0015	20.84	< 0.0001
Age at harvest	72.41	< 0.0001	409.67	< 0.0001
Residual variance	4.0277		3.4832	

Table 2 shows the statistical significance for the fixed effects and the linear covariate (age at harvest) for cage and pond, respectively. All main effects and the covariate were statistically significant ($P<0.05$).

The REML estimates of variance components, heritability, maternal common environmental effect and genetic correlation are shown in Table 3. The heritability estimates for cage and pond were slightly higher compared to other reported estimates in tilapia (0.24 by Gall and Bakar 1999; 0.20 by Gall and Bakar 2002; 0.34 by Ponzoni *et al.* 2005; 0.32 by Maluwa *et al.* 2006). The estimated maternal common environmental effects for cage and pond are in agreement with the estimates in literature (0.15 by Ponzoni *et al.* 2005; 0.21 by Rutten *et al.* 2005)

Table 3. Phenotypic and genetic parameters, and selection response for live weight (LW, $g^{0.5}$) in cage and pond

Parameter	REML estimate	
	LW in Cage	LW in Pond
Additive genetic variance (σ^2_A)	2.406	2.804
Maternal common environmental variance ($\sigma^2_D = \sigma_{M\_Ec}{}^2$)	1.282	1.666
Phenotypic variance (σ^2_P)	6.995	6.951
Heritability [h^2(s.e.)]	0.34 (0.06)	0.40 (0.07)
Maternal common environmental component [c^2(s.e.)]	0.18 (0.03)	0.24 (0.03)
Genetic correlation [r_g(s.e.)]	0.75 (0.09)	

The magnitude of the genetic correlation between cage and pond estimates the degree to which the same genes are involved in the expression of weight in these two environments. The genetic correlation estimated was 0.75±0.09 (Table 3). This result indicates that if selection were conducted in one environment (say, cages), but progeny were to perform in another environment (say, ponds), assuming equal heritability in both environments, selection in cages would capture 75 percent of the gain that could be achieved if it were carried out in ponds.

The estimates of genetic gain were encouraging, 18.6% gain in cages and 15.3% gain in ponds, after two rounds of selection. The response was large enough to indicate that genetic change was being achieved in both the cage and pond environments, and in the intended direction. Furthermore, the gains in cages and ponds, resulting from the bivariate analysis used in this study, were in good agreement with those resulting from a univariate analysis (treating the expression in both environments as a single trait) earlier reported by Ponzoni *et al.* (2005).

CONCLUSIONS

Falconer's (1952) approach of treating the expression of the trait in different environments as if they were different traits is useful in understanding and drawing practical conclusions from GxE studies. In the present case, the genetic correlation between live weigh in cages and in ponds indicates that if selection were conducted in one of the environments, 75 percent of the gain achieved in that environment would be captured in the other environment. The 95 percent confident interval for the estimated genetic correlation ranged from 0.66 to 0.84, which indicated moderate to low GxE. Coupled with the high heritability and selection responses obtained, we conclude that there was no evidence of GxE between cage and pond culture environments for tilapia farming in Malaysia that was large enough to warrant separate breeding programs. However, having or not a single breeding program should not be solely based on the genetic correlation, but also the economic importance of each culture environment and on the feasibility of implanting an additional program under the specific circumstances in question.

REFERENCES

Bolivar, R.B. (1998) PhD thesis, Dalhousie University, Canada.

Eknath, A.E. and Acosta, B.O. (1998) "Genetic improvement of farmed tilapias (GIFT) project: Final report, March 1988 to December 1997". International Center for Living Aquatic Resources Management, Makati City, Philippines.

Eknath, A.E., Tayamen, M.M., Palada-de Vera, M.S., Danting, J.C., Reyes, R.A., Dionisio, E.E., Capili, J.B., Bolivar, H.L., Abella, T.A., Circa, A.V., Bentsen, H.B., Gjerde, B., Gjedrem, T. and Pullin R.S.V. (1993) *Aquaculture* **111**:171

Falconer, D.S. (1952) *Amer. Nat.* **86**:293

Gall, G.A.E. and Bakar, Y. (1999) *Aquaculture* **173**:197

Gall, G.A.E. and Bakar, Y. (2002) *Aquaculture* **212**:93

Gilmour, A.R., Gogel, B.J., Cullis, B.R., Welham, S.J. and Thompson, R. (2002) ASReml User Guide Release 1.0, VSN International Ltd., Hemel Hempstead, HP1 1ES, UK.

Hanafi, H.H. and Chua, H.P. (2008) http://www.fri.gov.my/pppat/culture_protoc.htm [Consulted on December 4th, 2008].

Maluwa, A.O., Gjerde, B. and Ponzoni, R.W. (2006) *Aquaculture* **259**:47

Ponzoni, R.W., Hamzah, A., Tan, S. and Kamaruzzaman, N. (2005) *Aquaculture* **247**:203

Rutten, M.J.M., Komen, H. and Bovenhuis, H. (2005) *Aquaculture* **246**:101

SAS Institute Inc. (1997) "SAS/STAT® Software: Changes and Enhancements through Release 6.12", Cary, NC, USA.

Tayamen, M.M. (2004) *Proc. of 6th ISTA*, Manila, Philippines.

VARIABILITY IN THE DISTRIBUTIONS OF SINGLE NUCLEOTIDE POLYMORPHISM EFFECTS IN LIVESTOCK POPULATIONS

E.J. Smith and J.M. Henshall

F.D. McMaster Laboratory Chiswick, CSIRO Livestock Industries, Armidale NSW 2350

SUMMARY

Variability in the distributions of single nucleotide polymorphisms (SNP) effects were investigated by simulation. Realised distributions after thousands of generations of selection were found to be sensitive to parameters relating to the genome size, SNP density, population size and the distribution of sampled SNP effects. In particular, the distributions were not generally exponential, and in some cases SNPs of smaller effect were less likely to segregate than SNPs of larger effect.

INTRODUCTION

In association studies with dense SNP and phenotype data, Bayesian methods require assumptions regarding the distribution of SNP effects. However, data available to gain an understanding of the true but unknown distribution are limited. Firstly, the minimum sized quantitative trait locus (QTL) that can be detected will vary with the experimental design. Also, the properties of the trait and closely linked QTL can result in misleading estimates of size and effect. A meta-analysis of QTL effect distributions on pig and dairy data indicated effects were skewed with a few QTL of large effect (Hayes and Goddard 2001), and a review of QTL analyses in *Drosophila* concluded that the distribution of homozygous QTL is exponential with the larger effects contributing to most of the variation between parental lines (Mackay 2004); however, there has been no study of any statistical precision on the distribution of QTL effects (Roff 2007).

In this paper we present results of a simulation study that explored the sensitivity of the distribution of SNP effects to factors such as population size, genome size, SNP density and sampling distributions. Although the simulated population structures are modelled on sheep or cattle populations, we make no claims that the realised distributions are applicable to real livestock populations. Rather, the purpose of the study is to identify factors that might influence the distribution of SNP effects.

MATERIALS AND METHODS

Five population structures loosely based on a ruminant animal with annual joining were simulated for 105000 years (after equilibration) and data collected at years 1, 50000 and 100000 for a 5000 year period. All base animals were assigned as homozygous with allele "0" and equilibration was deemed to be the first instance when the frequency, q, of SNP's with allele "0" was equal to the frequency, p, of SNP's with allele "1", where p and q sum to one. Each year, 100 or 50 males were joined to 500 or 250 females, respectively, with animals eligible to enter the breeding herd between the ages of 2 and 6. Females had only one progeny per year. Selection was on a trait with a heritability of 25%. Animals were ranked according to their distance from a selection target.

Genomes with varying number of SNPs distributed over chromosomes of differing lengths (see Table 1 for details) were simulated with a SNP mutation rate of 3.1×10^{-4} per gamete. SNP effects were drawn from a uniform distribution, U, over (-5,5) or (-10,10); however, new effects were only sampled if the SNP was fixed (i.e., p = 0.0 or p = 1.0) in the current simulated population, reducing the realised mutation rate. Where the SNP was fixed with p = 1.0 the simulated population mean was adjusted to account for the change in the SNP effect. No polygenic variance

was simulated; all of the genetic variance was due to the SNPs. The desired heritability was achieved by holding the environmental variance constant at 60.0 and tuning the simulation as it ran to achieve the appropriate genetic variance of 20.0. To create genetic variance the selection target was moved a small amount each year, with the step-size dynamically tuned according to the size of the current realised genetic variance in relation to the desired genetic variance of 20.0. A second trait was also simulated, and after animals had been selected for breeding each year, mate allocations were assortative based on the second trait. Such an assortative mating system increases genetic diversity, and can be justified as accounting for a spatial component in wild or domestic populations. SNP transmission between animals as well as mutation and recombination events were recorded and written to a file after the simulation. The differences in parameters for the five simulated populations are outlined in Table1.

Table 1 – Population parameters used in each simulation (cM = centimorgans)

Simulation	# SNPs	SNP / cM	Dams	Sires	U
1	1600	16	500	100	(-5,5)
2	800	16	500	100	(-5,5)
3	1600	32	500	100	(-5,5)
4	1600	16	250	50	(-5,5)
5	1600	16	500	100	(-10,10)

RESULTS AND DISCUSSIONS

The distributions for the contribution of each individual SNP to the additive variance, $2pq\alpha^2$, and the absolute value of SNP effect sizes, α, are shown in Figure 1. The data is collated for $0.1 \leq p \leq 0.9$. Results for the three 5000 year periods of data collection (first two not shown) were similar, indicating the simulations had stabilized.

SNP effect distributions used as priors in Bayesian approaches are often assumed to be a decaying exponential function, which is a function whose second derivative is always greater than zero. Whilst the SNP effect distributions obtained for simulations 1, 3, 4 and 5 generally indicate that there are a greater number of SNPs with smaller effects than SNPs with large effect, the distributions contain inflexion points (second derivative is zero) and sections where the second derivative is less then zero (concave "down" shape) and therefore are functionally different from a decaying exponential distribution. In contrast, simulation 2 displays a uniform distribution. These results indicate that the distribution of α may differ depending on the simulation parameters and may not necessarily be exponential. In simulation 5 the distribution of SNP effects also shows that mutations on SNPs with effects sampled with a value of less than -6 or greater than 7 never survive beyond a frequency of 0.1. Further analysis also indicated that in simulation 5, mutations on SNP's with effects greater than the absolute value of 6 never became fixed in the population, indicating an upper limit to the effect size for mutations that can survive in a population.

The distribution of $2pq\alpha^2$ values were exponential in behaviour and do not vary significantly between the different populations. The possible exception is simulation 2 where the distribution appears more normal. This is due to the behaviour of the SNP effect distribution as opposed to the frequencies, as these are almost identical for simulations 1 and 2 (not shown). It should be noted that the frequency distributions were not identical for all 5 simulations. Increasing the SNP density increased the frequency of mutations segregating in the population, and as discussed below, increasing the range of the sampling distribution of SNP effects decreased the frequency of mutations segregating.

Figure 2 shows the distribution of $2pq\alpha^2$ as a function of p for simulations 1 and 5. Qualitative inspection indicates that mutations move through the population with varying "velocities" and that

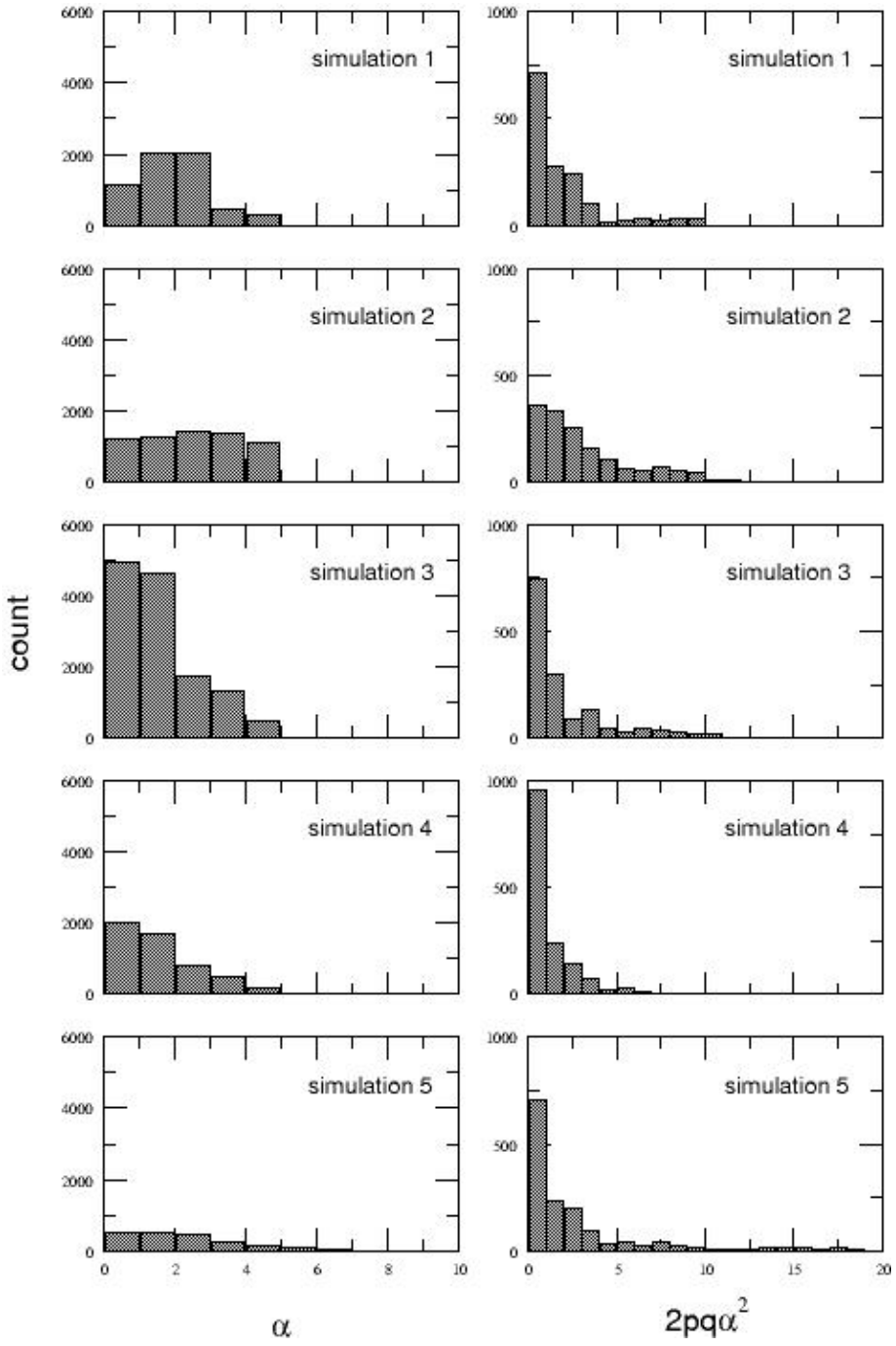


Figure 1. Distribution of the absolute value of the effect size, α, for $0.1 \leq p \leq 0.9$, and the contributions to the additive variance from individual SNPs, $2pq\alpha^2$, collected during the last 5000 years of the simulated populations.

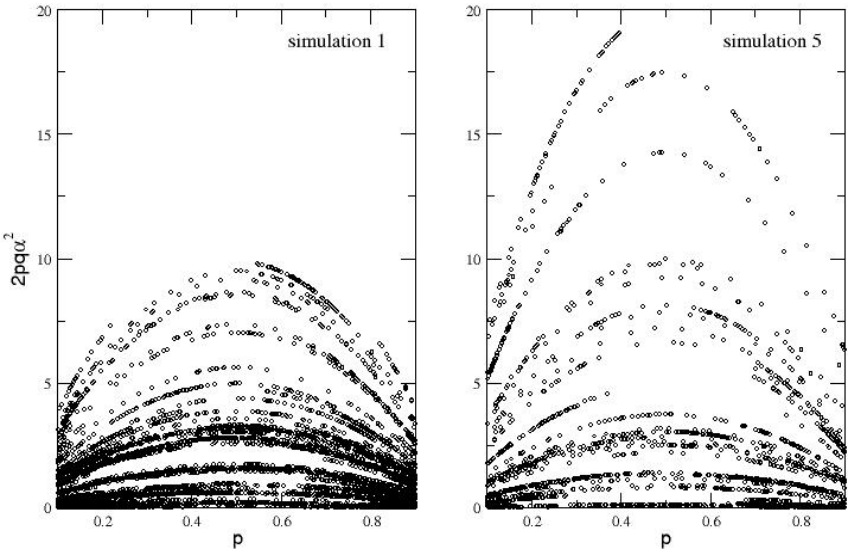


Figure 2. Data from simulations 1 and 5 showing the contribution of each SNP, $2pq\alpha^2$, to the additive variance as a function of the frequency, p, of SNP's with allele "1", collected during the last 5000 years of the simulated populations.

this observation is similar for both simulations. Quantitative analysis of the data showed that the average age to fixation and deletion was approximately the same for all simulations, and was also reasonably independent of the size of the SNP effects. However, Figure 2 also indicates that there are fewer mutations moving through the population for simulation 5 than simulation 1, and this was confirmed by extracting a histogram of the frequencies. The explanation may lie with the sampling distribution. As SNP effects are sampled from a uniform distribution, they are all initially equally as likely. Effects in simulation 5 are sampled from a distribution twice as broad as simulation 1, but the mutations with effects greater than those from the sampling distribution used for simulation 1 mostly do not segregate beyond p = 0.1. In other words, while the same number of mutations are occurring in each simulation, less of those in simulation 5 actually exist in the population for any significant period of time.

In this study we considered only one approach to producing a population with a SNP based genetic variance consistent with heritabilities observed in livestock populations. We recognise that this may also have a large influence on the distribution of SNP effects.

CONCLUSIONS

The distributions of SNP effects and their contributions to additive variances in livestock populations were investigated for populations that were simulated with different parameters. Results indicate not only a variability in the SNP effect distributions, but that the distributions do not consistently follow an exponential decay as effect size increases. This study supports that there is little to justify using particular distributions as priors for Bayesian analyses of SNP effects.

REFERENCES

Hayes, B. and Goddard, M.E., (2001) *Genet. Sel. Evol.* **33**:209
Mackay, T.F.C. (2004) *Cur. Opin. Genet. Dev.* **35**:253
Roff, D.A. (2007) *Evolution* **61**:1017

USE OF GENOTYPE PROBABILITIES AND SELECTIVE GENOTYPING FOR ESTIMATION OF MARKER EFFECTS

M.L.A.N.R. Deepani and B.P. Kinghorn

The Institute for Genetics and Bioinformatics, University of New England, Armidale, NSW

SUMMARY

Cost of genotyping associated with marker assisted selection is reduced by partial genotyping. Genotype probabilities (GPs) and selective genotyping can both be used to alleviate the reduction in power that partial genotyping can bring. This paper tests the effect of the combined strategies on the estimation of a single marker effect**.** The results suggest that with selective genotyping the marker estimates become less biased when genotype probabilities are used and when animal relationships are included in the selection function.

INTRODUCTION

Marker assisted selection (MAS) aims to give higher response to selection by increasing accuracy of genetic evaluations (Hayes *et al*. 2007). The cost of genotyping associated with MAS is reduced with partial genotyping and use of genotype probabilities for non-genotyped animals. Baruch *et al*. (2008) used the segregation analysis method of Kerr and Kinghorn (1996) to help estimate the effects of two QTL for a population in which only the bulls had been genotyped. Selecting animals for genotyping on the basis of their phenotypes ("Selective Genotyping", Lander *et al.* (1989); Darvasi and Soller (1992)) also reduces genotyping cost while managing the loss of precision in detecting a QTL effect. This study combines these two strategies, and extends this by considering animal relationships as well as phenotypes when selecting animals to genotype.

MATERIALS AND METHODS

A population was simulated starting with 30 males and 150 females in different age classes. For the foundation animals a bi-allelic SNP marker was simulated with allele frequency 0.5 and allele substitution effect of 6. Alleles were sampled from a uniform distribution, and then propagated in a Mendelian pattern. The mean, phenotypic standard deviation and polygenic heritability for the single trait of interest were assumed to be 100, 30 and 0.25 respectively. At the end of ten years of random selection and mating, the marker effect was estimated using BLUP in which the marker genotype or its probabilities were fitted as a covariable.

There were two major experiments for treatments. Experiment1 was partial genotyping with random selection of animals for genotyping (Rnd). Under Experiment 1 there were 5 treatments. Treatment 1 used 100% genotype information and this was used as a control. Treatments 2 to 5 used 50%, 35%, 20% and 5% genotyping with random selection of animals for genotyping, and genotype probabilities from Kerr and Kinghorn (1996) for the ungenotyped animals. Treatments 2 to 5 all included sub-treatments with different thresholds of genotype probability index (GPI, Kinghorn, 1997) as minimum levels for accepting animals into the BLUP analysis. A threshold of 100% means that only animals with 100% GPI were included. Low GPI animals tend to be poorly connected to genotyped individuals and have poor quality genotype probabilities.

Experiment 2 involved selective genotyping, using the method introduced by Kinghorn *et al.* (2006). This uses an evolutionary algorithm to select animals on an objective function of two components: "Distance" - a measure of phenotypic dispersal (following Davarsi and Soller 1992) across traits, and "Relationship" – a metric of both decreased relationship between phenotypically similar animals (the contrast between phenotypically divergent groups is weakened by a small

number of parents contributing to each group) and increased relationship between phenotypically dissimilar animals (which helps contrast QTL effects, as in some sib- pair designs). "Relationship" is aimed at cleaner designs that reduce the number of false positive calls.

Three treatments were: full emphasis on "Distance" denoted by "D"; full emphasis on "Relationship" denoted by "R", and equal emphasis on each, denoted "DR", which used a function to target the averages of the D and the R results achieved. As for Experiment 1, the marker effects were estimated with the same five levels of genotyping and with or without GPs, with sub treatments using different GPI thresholds when GPs are used. Twenty replicates were run for each treatment.

RESULTS AND DISCUSSION

Experiment 1. *Random selection and use of genotype probabilities*. It was observed that estimated marker effects were close to the true value when GPI thresholds up to 30% were used compared with using information on genotyped animals only ("No GP", Figure 1). Standard deviations (SDs) between replicates reflect the standard error of one replicate and hence accuracy of estimation. SDs were higher for estimates obtained without using GPs (No GP) than with GPs. The minimum SD was obtained when all records were used (All GPI), even those with very low GPIs. At 5% genotyping, SD was ±3.775 with "No GP" while it was ±3.398 with "All GPI".

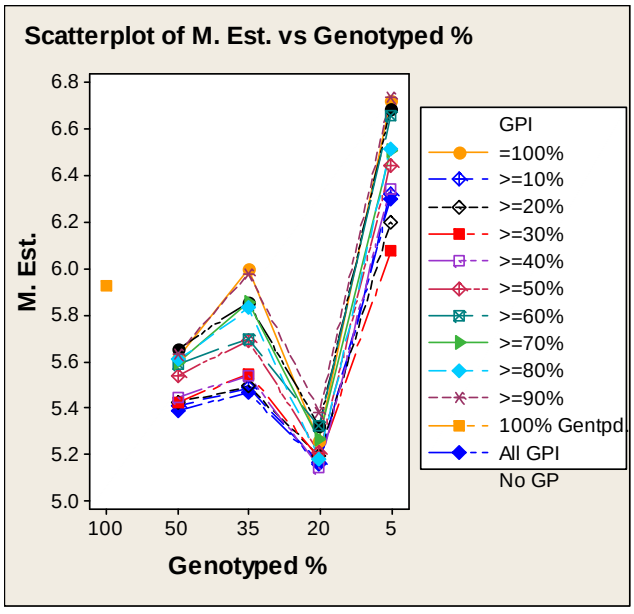


Figure 1. Marker effect estimates (M. Est.) with different proportions of genotyping.

Experiment 2. *Selective Genotyping*. Under selective genotyping, the estimates obtained were biased upwards in comparison to random selection, as predicted by Darvasi *et al.* (1992). This overestimation was reduced with the inclusion of relationship together with distance as a selection criterion (DR and R, Figure 2). SDs also reduced when the relationship was taken into account, while minimum SD observed when equal emphasis was given to Distance and relationship (DR). At 5% genotyping SD of estimated marker effect under DR, with "All GPIs" is 3.667 which is very close to the SD under random selection (3.398, Figure 2) although the marker effect is overestimated.

The results for selective genotyping were corrected according to the following equation;

$$\delta = D_T / \gamma_p \qquad \textbf{Equation 1}$$

Where "D_T" is the raw estimated marker effect by selective genotyping, "δ" is the corrected estimate of the marker effect and $\gamma_p = 1+Z_{1-p/2} .i_{p/2}$. Where "Z" denotes the truncation point, "p" is

the proportion of the animals that have been genotyped and "i" is the selection intensity (Darvasi *et al*. 1992).

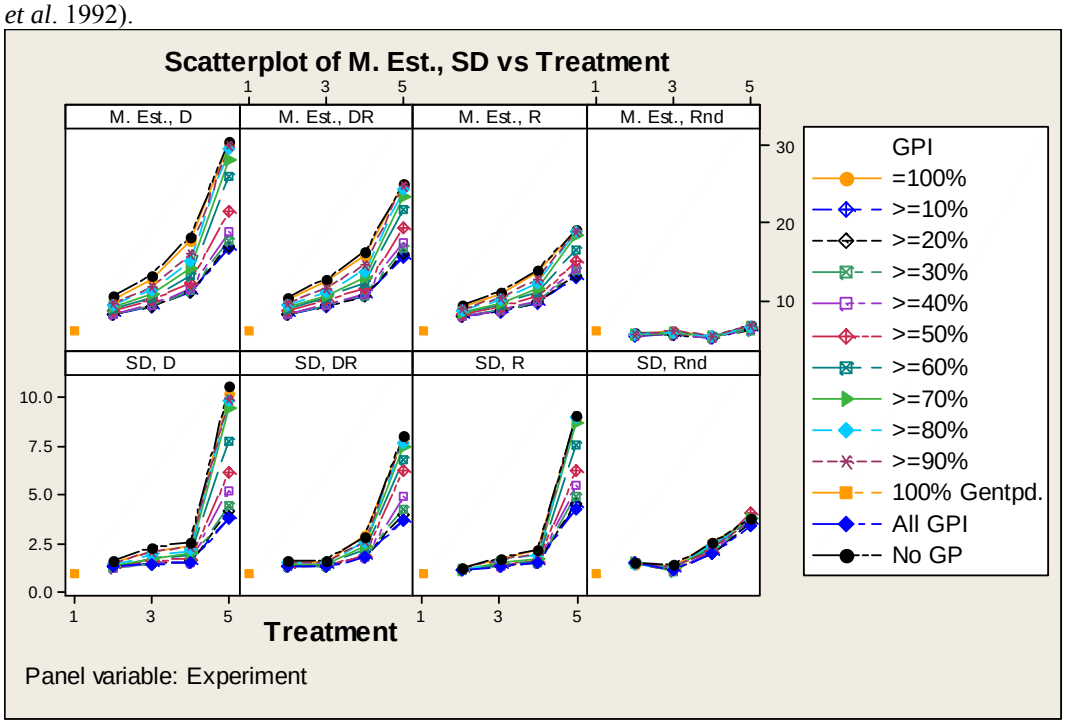


Figure 2. Comparison of marker effect estimates(row 1) and standard deviations (row 2) obtained from Experiment 1(Rnd) and Experiment 2; D, DR and R.

When the correction described above was made to the estimates, they were overcorrected especially at low selection proportions. This was true for "No GP" (Figure 3) as seen by (Muranty *et al.* 1997). Under selective genotyping the magnitude of overestimation is reduced by using GPs, which partly fill the missing information, making the strategy more similar to genotyping all animals, whence no is correction needed. This is why equation 1 results in the over correction seen in Figure 3.

Under D, the corrected results for "No GP" were less biased compared to the "No GP" under DR and R (Figure 3, upper row). This may be because for D, the selection is purely based on the phenotypic distance and this is where the formula for correction more precisely applied. For DR and R relationship was also taken into account and the formula is not designed to accommodate that criterion. The SDs for corrected marker effects were smaller (< 2) in all 3 selective genotyping methods compared to the standard deviation of the marker estimates obtained from random selection of animals (Figure 3), indicating higher accuracy.

Under random selection of animals for genotyping, marker effect estimates are more accurate if genotype probabilities are used; with increased accuracy related to the population average GPI. Under selective genotyping, including relationships in the selection function gives less bias results compared to phenotypic dispersal alone, and bias is further reduced if GPs are also used. Although Darvasi and Soller's equation tends to overcorrect this reduced bias, it more substantially reduces the SD of the estimate compared to random selection. Therefore accommodating relationships and

GPs increases the accuracy of estimating a positive marker effect even though it is underestimated after simple correction using equation 1.

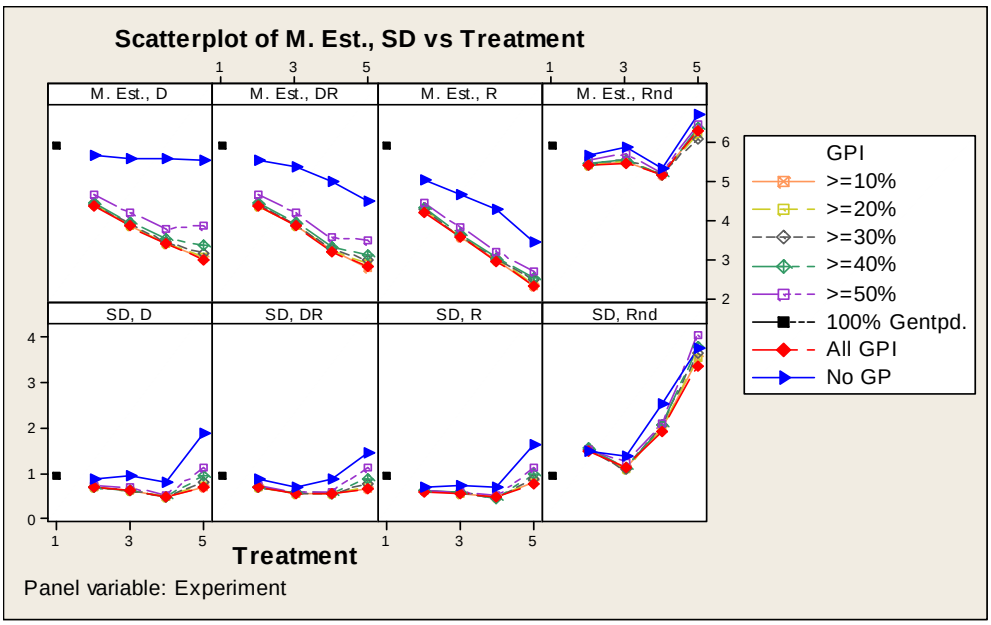


Figure 3. Corrected marker effect estimates (row 1) and their standard deviations (row 2) in selective genotyping (D, DR and R) compared with random selection (Rnd).

CONCLUSIONS

Accommodation of relationships as carried out in this paper reduces the bias in marker estimates that are caused by correlations with polygenic effects. Use of genotype probabilities gives further improvement, by including animals that would not otherwise be used in the analysis. Bias due to relationships is more important than bias due to selective genotyping *per se*, because the latter affects all markers in a similar manner, such that ranking of markers is less affected. Therefore, this combined approach should help when choosing discovery project markers for validation. In such cases, general trend in estimation bias is less critical than capturing markers that in fact have true effects.

ACKNOWLEDGEMENT

To the IRQUE project of the Faculty of Veterinary Medicine and Animal Science University of Peradeniya, Sri Lanka for funding this study.

REFERENCES

Baruch, E., J. I. Weller (2008) J. *Dairy Sci.* **91**: 4365.
Darvasi, A., Soller, M. (1992) *Theoritical and Applied Genetics* **85**: 353.
Hayes, B. J., Chamberlain, A. J. *et al.* (2007) *Genetics Research* **89**: 215.
Kerr,R. D., Kinghorn, B.P. (1996) *J. Anim. Breed. Genet.* **113**: 457.
Kinghorn, B.P. (1997) *Genetics* **145**:479.
Kinghorn, B.P., Bastiaansen, J.W.M., van der Steen, H.A.M., Deeb, N., Yu, N., and Mileham, A.J. (2006) *Proc. 8th Wrld. Cong. on Gen. Appld. Livst. Prod.,* Paper 20-05. 4 pages.
Lander, E. S., Botstein, D. (1989) *Genetics* **121**: 185.
Muranty, H., Goffinet,B. (1997) *Biometrics* **53**(2): 629.

PHASING OF SNP DATA BY COMBINED RECURSIVE LONG RANGE PHASING AND LONG RANGE HAPLOTYPE IMPUTATION

J.M. Hickey[1,2], B.P. Kinghorn[2], B. Tier[3], J.H.J. van der Werf [1,2]

[1]Cooperative Research Centre for Sheep Industry Innovation, Armidale, NSW 2351
[2]School of Environmental and Rural Science, University of New England, Armidale, NSW, 2351
[3]Animal Breeding and Genetics Unit, University of New England, Armidale, NSW, 2351

SUMMARY

A Long Range Haplotype Imputation algorithm was developed and combined with a Recursive Long Range Phasing algorithm. These were tested on simulated SNP data sets which had structures possibly similar to those found in sheep and dairy cattle. Performance, measured by accuracy of phasing (>97.6%) and computing time, was competitive in comparison to industry standard software.

INTRODUCTION

High density SNP arrays can provide detailed information on aspects of the genetic make-up of individuals, enabling whole genome SNP association studies, with possible applications in prediction of genetic value/predisposition in medicine, forensics, and animal and plant breeding. Data from SNP arrays are unphased i.e. paternal and maternal alleles are not determined. However, knowledge of the phase could be advantageous. The high density of SNP arrays allows inference about phase, and relatively efficient algorithms can be found to determine phase.

Local phasing methods like fastPHASE (Scheet and Stephens 2006) exploit strong correlations within LD blocks but they are computationally intensive, and SNPs separated by many LD blocks are not reliably phased (Kong *et al.* 2008). Long range phasing (Kong *et al.* 2008), which is faster and more powerful, was formulated in a recursive algorithm (RLRP) by Kinghorn *et al.* (2009). RLRP gives high computational speed and can lead to inference of phase with little error, for large proportions SNPs. However like LRP it may be unable to phase proportions of SNPs in a data set.

The first objective of this research was to develop a long range haplotype imputation (LRHI) algorithm to impute phase for the proportions of SNP alleles unphased by an RLRP algorithm. The second objective was to test the performance of a combined RLRP-LRHI algorithm for accuracy and computational efficiency under different scenarios possibly existing in some livestock populations.

MATERIALS AND METHODS

The RLRP-LRHI algorithm outlined here uses a modified version of the RLRP algorithm outlined by Kinghorn *et al.* (2009). It uses cores and spans to move along a chromosome. A core is the string of consecutive SNPs being phased. A span is a string of consecutive SNPs which include the core but may extend outside it. Spans are used to identify surrogates (defined below) at cores. Animals that have phase imputed by LRHI have their cores aligned at the end by re-running the algorithm with the core shifted by *n* SNPs to create overlapping cores for each location. Core and span length as well as tuning parameters, *threshold % of SNPs* and *n SNPs* (these feature below), which help manage genotyping errors present in real data, can be optimised for a given population and data structure.

Recursive Long Range Phasing Algorithm.

Step 1 – Identification and partitioning of surrogates. Surrogates of each animal are defined as animals for which >threshold % of consecutive SNP loci are not of opposite homozygote

genotypes, and are partitioned into surrogates of the maternal and paternal haplotypes using a dummy dam and sire. A dummy dam (or sire) is identified by searching elements of the numerator relationship matrix pertaining to the dam (sire) and the sire (dam) of the animal for a surrogate that is related to its dam (sire) but not its sire (dam). Surrogates of the maternal (paternal) haplotype are then identified as the surrogates that are surrogates of the animal and its dummy dam (sire).

Step 2 – Initial phasing of each SNP. Each SNP in the maternal (paternal) gamete is attempted to be phased (see Kinghorn et al., 2009 for details). The stopping criteria are (a) a number of surrogates found suggesting that phase is one of the categories and zero conflicting surrogates, (b) if *a* is not satisfied the recursion continues until the number of surrogates suggesting that phase is one of the directions is statistically significantly, and (c) if neither *a* or *b* are satisfied before a maximum Erdös number is reached phase is undeclared. These criteria can be tuned to suit the data.

Step 3 – Pruning surrogates and re-phasing of each SNP. The lists of surrogates for each animal are then pruned using the pruning step of Kong et al. (2008). Step 2 is then repeated using the pruned lists of surrogates with slight modification. If phase of the paternal (maternal) allele is unambiguous, phase of the maternal (paternal) allele is imputed as the genotype minus the paternal (maternal) allele.

Long range haplotype imputation algorithm

Step 1 – Building of haplotype library. Upon completion of the RLRP step a library of long haplotypes for the region being phased is created from completely phased haplotypes of animals.

Step 2 – Identification of candidate haplotypes. Candidate haplotypes for an animal with unphased SNPs are identified by comparing its phased SNPs to haplotypes in the library. A candidate haplotype for an animal's gamete is allocated if > threshold % of its SNPs match the gamete's phased SNPs.

Step 3 – Resolving phase. If just one candidate haplotype agrees with > threshold % of the animals SNP genotypes, phase is taken to be that haplotype. The alternative gamete has its unphased SNPs imputed as the genotype minus the candidate haplotype and this new haplotype is added to the library. If only two candidate haplotypes are identified and their sum agrees with > threshold % of the animal's SNP genotypes, phase is taken to be these haplotypes. If more than two candidate haplotypes are identified they are paired and if a sum of a single pair agrees with > threshold % of the animal's SNP genotypes, phase is taken to be these haplotypes.

Step 4 – Incompatibility check. Each SNP genotype for each animal is compared to the imputed phase. If these did not agree phase is undeclared at this location.

Step 5 – Iteration. Steps 1 to 4 were repeated until no new haplotype was added to the library.

Simulations. Sheep and dairy cattle populations in equilibrium between mutation, drift, and recombination were simulated using a gene-drop involving a burn in of 9,000 discrete generations of random mating of 500 (50) males and 500 (50) females for sheep (dairy cattle). Each individual's genome consisted of a pair of chromosomes, each 0.1 Morgan long. Mating involved sampling a parent twice without replacement, in such a manner that each animal in the parental generation had two offspring and parents were randomly paired at each mating. For each parent, a gamete was formed by sampling its chromosomes using a recombination rate (0.1 per gamete) and

crossover points sampled from a uniform distribution respectively. Base animals had all alleles set to 0. Each gamete had one locus randomly selected as a candidate mutation locus. Mutational events occurred at candidate loci if no segregation at these loci occurred in the population in that generation. 400,000 loci were simulated. In both the sheep and cattle populations two data structures were simulated each with parental and offspring generations of genotyped animals. The offspring generation comprised 1,000 animals. The parental generation comprised 500 dams while one structure had 10 sires and the other had 100 sires. Genotypes of animals were created by continuing the gene drop into these pedigrees. At the end of each simulation 200 and 1,000 SNPs that had a minor allele frequency of >0.1 were randomly chosen for placement on two SNP arrays, reflective of 50,000, and 300,000 SNP array density, and possibly the uneven distribution of SNPs, with regard to physical location, on these arrays.

Validation methods. Three replicates were simulated and analysed for each scenario. There were two groups of animals to be phased, parents, who themselves had unknown and ungenotyped parents, and offspring, who had known and genotyped parents. Performance was measured for these groups as % of SNPs correctly phased, % of SNPs incorrectly phased, and % of SNPs not attempted to be phased. For comparison the widely used phasing software, fastPHASE, was used to phase one replicate of the simulated data.

RESULTS

Mean performance, across the three replicates, for sheep and cattle with sire family size of 10 is given in Table 1. Results for sire family size of 100 were very similar. Variation in performance across replicates was small. Performance for RLRP ranged from 86.88% to 99.73% correctly phased for all SNPs in the offspring. The higher performance was only obtained for dense SNPs in the sheep population. With the exception of where RLRP performed very well, the SNPs that were not correctly phased were mostly not attempted to be phased rather than incorrectly phased. The highest incorrectly phased % was for the sheep population when low SNP density was used (1.40% - 1.92%). For the other scenarios the incorrectly phased % for RLRP was less than 0.24%. The performance of the combined RLRP-LRHI algorithm was encouraging (>97.6%) resulting in major reductions in the not phased % with, at worst, only tiny increases in the incorrectly phased %. A strong relationship between mistakes generated by the RLRP and the LRHI for the offspring animals existed. Performance of RLRP and LRHI increased with increasing SNP density, slightly increased with reducing population size used during the burn in, but was unaffected by family size. Computational time was many times greater for fastPHASE (circa 18 hours) than for RLRP-LRHI (circa 0.33 hours). In the sheep scenario phasing performance of RLRP-LRHI was at worst only very slightly poorer than fastPHASE but in some cases was slightly or considerably better (Table 2). Differences were smaller for cattle as inbreeding was higher.

DISCUSSION

For this first test of an RLRP-LRHI algorithm performance were encouraging both in terms of computational speed and phasing accuracy. Given the wide range of options implemented in the algorithm greater performance is likely. Any reduction in the incorrectly phased % in RLRP would also improve LRHI. More stringent stopping criteria during the first recursion of RLRP could improve performance. Running the algorithm multiple times using different start and end points for cores and spans, or randomly deleting large percentages of surrogates in a quasi Bootstrap manner could identify SNPs that are prone to being incorrectly phased by RLRP. Currently RLRP and LRHI use the same core length, which is probably suboptimal. Longer cores may be more appropriate for LRHI. An implementation of the RLRP which does not use pedigree information is possible and should improve the performance of the LRHI. The current algorithm is rule based

except for a small step where a statistical test is invoked. Incorporation of statistical inference may improve performance especially when genotyping errors are present in the data, particularly for LRHI. A large proportion of the errors from the RLRP-LRHI algorithm occurred where strings of consecutive heterozygous SNPs existed. These result from pairs of very common haplotypes that dominate SNP selection on allele frequency to the extent that loci that are IBS for these haplotypes are not chosen. When these haplotype form a zygote, this lack of IBS yields strings of heterozygotes. This appears to be uncommon in real datasets, but that may partly be due to incorrect map order. Performance differences between RLRP-LRHI and fastPHASE are expected to increase with increasing effective population size. The RLRP-LRHI algorithm outlined here could be used for imputing dense SNP or full sequence data through a pedigree, possibly increasing power of genome wide association/selection studies by reducing allelic heterogeneity and by increasing the numbers of phenotyped animals with genotypes imputed. Performance in human populations, which generally have smaller family sizes, larger effective population sizes, and different relationship structures compared to livestock populations, may be different and will be the subject of a further study.

Table 1. Percentage of SNPs correctly (Cor.), incorrectly (N cor.), and not (N pha.) phased for parents and offspring in sheep and dairy cattle populations with sire family sizes of 10 offspring using SNP arrays of 50,000 and 300,000 density

			SNP Density					
			50,000			300,000		
			Cor.	N cor.	N pha.	Cor.	N cor.	N pha.
Sheep	RLRP alone	Parents	73.07	0.00	26.93	73.09	0.00	26.91
		Offspring	90.54	1.92	7.55	96.67	0.20	3.13
	RLRP + LRHI	Parents	98.08	0.09	1.83	98.91	0.22	0.87
		Offspring	98.34	0.97	0.68	99.50	0.32	0.18
Cattle	RLRP alone	Parents	67.93	0.00	32.07	67.85	0.00	32.15
		Offspring	87.00	0.10	12.90	88.93	0.07	11.01
	RLRP + LRHI	Parents	97.59	1.35	1.06	98.29	0.93	0.78
		Offspring	99.35	0.09	0.57	99.45	0.09	0.46

Table 2. Performance of RLRP and LRHI and fastPHASE (Individual error minimised/Switch error minimised) for offspring in replicate 1 of the sheep (family size = 10) scenario.

	SNP Density	
	50,000	300,000
fastPHASE	98.59 / 97.69	93.03 / 96.11
RLRP+LRHI	98.03	99.36

REFERENCES

Kinghorn, B.P., Hickey, J.M., and van der Werf, J.H.J. (2009) *Proc. Assoc. Advmt. Anim. Breed. Genet.* **18**:76.

Scheet, P., and Stephens, M. (2006) *American Journal of Human Genetics.* **78***:*629

Kong et al. (2008) *Nature Genetics* **40:**1068.

A RECURSIVE ALGORITHM FOR LONG RANGE PHASING OF SNP GENOTYPES

B. P. Kinghorn, J.M. Hickey and J.H.J. van der Werf

The Institute of Genetics and Bioinformatics,
University of New England, Armidale, Australia, NSW 2351

SUMMARY

High density genotyping of individuals does not by itself yield information on phase: linkage of alleles between loci in the contributing gametes. This additional information is important, because it helps to determine which segments of chromosome are identical by descent across the population, and thereby improve inference about segregation of quantitative trait loci (QTL). This paper presents a recursive algorithm to give efficient implementation of long range phasing, a recently-published strategy that infers IBD between distant relatives from long strings of SNP genotypes that show no opposing homozygotes (i.e. no evidence of lack of IBS). These relatives can be both distant and unknown to the analysis. Successful implementation of this strategy gives opportunity to use haplotype and/or combined linkage-linkage disequilibrium analysis for QTL mapping or genomic selection.

INTRODUCTION

Information on phase gives added power to map and/or exploit QTL, and to make inference about causal mutations. Without full linkage disequilibrium, alleles at a QTL will vary within the group of gametes that contain a given allele at the most informative marker locus (eg. for all gametes carrying marker allele M, some will carry QTL allele q, and some will carry Q). This reduces power to detect or exploit that QTL. Full phasing can act to place a notional subscript to each marker allele, denoting the identity by descent (IBD) segment of chromosome to which it belongs, and giving a tight relationship with the QTL allele carried.

Phasing at an individual locus can often be carried out with knowledge of the genotypes of the proband (the individual of interest) and its two parents. Only when all three are heterozygous at a biallelic marker locus is it not possible to declare which progeny allele has been inherited from which parent. In such cases, we can replace at least one of the heterozygous parents with a "surrogate parent" that is homozygous, and is known with high probability to carry the parental allele that the real parent transmitted to the proband. This is a novel approach presented by Kong *et al.* (2008). Its key feature is recognition that each parent of a proband can be represented by surrogate parents – individuals that carry the same parental haplotype at the region concerned, and can thus act in place of the parent when deducing which allele has been transmitted from parent to offspring. The method is termed "Long Range Phasing", which reflects the fact that information used for phasing the proband can come from individuals that are distantly connected by pedigree, at long range, with many degrees of separation involved.

If neither the parent nor any of its surrogates are homozygous, then there is still a chance inferring the solution: if any one of the heterozygous surrogates can have its genotype phased, then its homozygous parent or surrogate must be IBD at the opposite haplotype, and, being a homozygote, cannot be a surrogate of the original proband, which must thus have inherited the opposite allele. This second layer of surrogates are at an Erdös number of 2 from the proband (See Kong *et al.*, 2008 for more detail). There is no prior limit to the Erdös number of surrogates that can be used, such that phasing can be carried out using information from distantly related individuals, and non-relatives such as "in-laws" that are connected in the pedigree.

Surrogacy is bidirectional: if A is a surrogate of B, then B is a surrogate of A. Surrogacy is declared upon sufficient evidence that two individuals share a haplotype at the prevailing region. This in turn is inferred by detecting that no opposing homozygotes exist between the individuals over a long distance involving many genetic markers. With sufficient distance and marker density, there is a sufficiently small chance that no opposing homozygotes would occur without IBD due to a common haplotype.

This paper presents algorithms for long range phasing of biallelic markers in a population, given that surrogacy relationships have already been inferred. These algorithms use surrogates available or required at all levels of Erdös number for each individual/SNP combination that is phased. For simple presentation, incompatibilities and errors are not dealt with in this paper, but are addressed by Hickey *et al.* (2009). Kong *et al.* (2008) used an iterative algorithm with steps to deal with incompatibilities and errors. The recursive algorithm presented here has the apparent advantage of using more information, because all surrogates beyond Erdös number 2 can be used, not just those that can be assigned to the paternal or maternal side of the pedigree of the individuals for which they are direct surrogates.

METHODS

The simplest component of the overall task is to infer the source of inheritance (paternal or maternal) of each allele in a heterozygous individual. If genotyping and pedigree errors are absent, we only need to do this for one of the two alleles. For simple presentation we will refer to this process as phasing, even though it is done for a single locus at a time. The individual to be phased is referred to as the proband, and it is taken to be heterozygous, as the case for homozygosity is trivial: the same allele is inherited from each parent.

Each marker locus is handled in turn. Because of this, there is a need to separate maternal and paternal surrogates for the proband itself, so that direction of inheritance for each phased marker can be aligned with those for the linked markers.

The recursive algorithm. The first call is to a non-recursive function (steps 3, 4, 5 below) to give special treatment to the proband. If phasing is not made using the direct maternal and paternal surrogates of the proband, this function makes calls to the recursive function (steps 6, 7, 8 below) which does not distinguish between maternal and parental surrogates.

The dummy argument for the recursive function is referred to as the "current individual" (CI), and the function result is either an allele (as described at step 6), or a code denoting no success in phasing the CI. As no CI is the proband, it is sufficient to identify a homozygous surrogate from either side of its pedigree, as this defines the allele that is in the opposite haplotype to the one that is IBD with that in the calling individual, one Erdös level below. The recursive function has a generally bigger overall pool of surrogates, because it includes those surrogates that cannot be allocated to one (or occasionally both) side(s) of the CI's pedigree.

1. Erdös number, stored globally, is set to one.
2. The non-recursive function (steps 3, 4, 5) is called for the proband.
3. A pass is made across all the individuals that are a surrogate of both the proband and its father, looking for a surrogate that is homozygous. If a homozygote is found, the allele of the homozygote is the allele inherited by the proband from its father, and the non-recursive function is exited to Step 9 with a value denoting the paternal allele.
4. As for step 3, but replace 'father' with 'mother'. However, the 'paternal' allele is still the allele reported to Step 9.
5. If no homozygote is found, the non-recursive function continues, the Erdös number is incremented by one, and ***the recursive function is called*** (starting at Step 6) in turn for

each of the surrogates described in step 3, and then in step 4 with an allele swap to report the paternal allele. If a function call returns a positive paternal allele result, the non-recursive function is exited to Step 9 with a value denoting the paternal allele. If no positive paternal allele result arises, Erdös number is decreased by one and the non-recursive function is exited to Step 9 with a value denoting lack of a phasing call.

6. ***The recursive function is entered*** with the CI as the dummy argument and the alternative allele status of the CI as the result. A pass is made across all the individuals that are a surrogate of the CI, looking for a surrogate that is homozygous. If a homozygote is found, the allele of the homozygote is the allele inherited by the CI at the alternative site to that which the CI has transmitted to the calling individual (for which the CI is a surrogate), and the current instance of the function is exited into Step 5 or Step 7, from where it was called, with a value denoting the alternative allele.
7. If no homozygote is found, the prevailing Erdös number is incremented by one, and ***the recursive function is called*** in turn for each of the surrogates described in step 6. As soon as a function call returns a positive alternative allele result, the current instance of the function is exited into Step 5 or Step 7, from where it was called, with a value denoting the alternative allele.
8. If the current instance of the recursive function has not been exited, the Erdös number is decreased by one and the function is exited into Step 5 or 7, from where it was called, with a value denoting lack of success in phasing the CI.
9. From Step 5 alone. If a positive result returns for the proband, the Erdös number in which homozygosity was found has to be taken into account: If mod(ErdösNumber,2)=0 then the allele of the result is swapped so that it pertains to the proband. [EG. when ErdösNumber =2, a heterozygous surrogate of the proband has had one of its alleles phased, and it is the other allele that is IDB with the proband.] After this manipulation, the reported allele is the paternally inherited allele.

The algorithm can be seen to be acting as a ratchet, increasing Erdös number when no homozygote has been found in the current Erdös 'layer', for the current part of the pedigree. In the end, if no homozygote is found, the ratchet slips back through shells of the recursive function to return a negative result for the proband. However, given the recursive nature of the algorithm, multiple pathways are searched.

An iterative algorithm.

1. For each unphased individual:
    a. Cycle through all paternal surrogates. If a homozygote is found, record that allele as the paternal allele, flag the individual as phased and loop to the next individual.
    b. Cycle through all maternal surrogates. If a homozygote is found, record the alternate allele as the paternal allele, flag the individual as phased and loop to the next individual.
    c. Cycle through all paternal surrogates. If a surrogate is found to have been phased, record the alternative allele as the paternal allele in the individual, flag the individual as phased and loop to the next individual.
    d. Cycle through all maternal surrogates. If a surrogate is found to have been phased, record that allele as the paternal allele in the individual, flag the individual as phased and loop to the next individual.
2. If Step 1 resulted in at least one new phasing, go to Step 1, else stop

This iterative algorithm is simpler, but it is less powerful because at any stage it uses surrogates to make inference about phase in a current individual. In contrast, the recursive algorithm, at all levels above Erdös 1, only needs to make inference about which allele in the surrogate belongs to the haplotype that is not IBD with the proband. It does not matter whether it was paternally or maternally inherited. This permits inclusion of all surrogates, including those that are not connected by pedigree and those that are surrogates to individuals with unknown parents.

Notes on implementation. Improvements can be made to both of these algorithms as presented here. Conditions are deemed satisfactory when the first homozygous surrogate is found, and this is risky because of pedigree errors, genotyping errors, and errors in inference of IDB. In practice the algorithms should be allowed to continue to collect further information about homozygous surrogates, and to handle ambiguous situations. Both algorithms as presented require at least some surrogates to be shared with at least one parent, to align direction of inheritance across linked loci. However, strategies to group surrogates according to homozygosity status across linked loci provides a route to releasing the need for pedigree to carry out long range phasing. In both cases, individuals that are surrogate on both sides of a proband's pedigree should be detectable by identical genotypes across the region concerned, and these should be eliminated as uninformative.

Hickey *et al.* (2009) provide preliminary results from use of a developed version of the recursive algorithm, based on simulated and real data. They report promising results, both for phasing and for diagnosing map and genotyping errors in high-density SNP data.

REFERENCES

Hickey, J.M., Kinghorn, B. P. and van der Werf, J.H.J. (2009) *Proc. Assoc. Advmt. Anim. Breed. Genet.* **18**:72.

Kong, A., *et al.* (2008) *Nature Genetics* **40**:1068

GENOMIC SELECTION USING A FAST EM ALGORITHM
1. UNDERSTANDING THE METHODOLOGY

R.K. Shepherd[1] and J.A. Woolliams[2]

[1]FABIE, CQUniversity, Rockhampton, QLD 4702
[2]The Roslin Institute & R(D)SVS, University of Edinburgh, Roslin, Midlothian, EH25 9PS UK

SUMMARY

The paper uses the technical computer software Mathematica® to explain the features inherent in the procedure emBayesB which is a fast EM algorithm for implementing genomic selection by mapping QTL in genome-wide dense SNP marker data. The prior mixture for a SNP effect and the bimodal shape of the posterior distribution of a SNP effect are displayed graphically, along with visualisations and calculations of how emBayesB estimates genomic breeding values. The companion paper (Shepherd *et al.* 2009) uses emBayesB to analyse simulated data.

INTRODUCTION

Genomic selection is a new tool for genetic improvement in animal breeding which uses genome-wide dense SNP markers to ensure all QTL are in linkage disequilibrium (LD) with at least one marker. The first step in genomic selection is the estimation of SNP effects using phenotype and genotype data in a reference population (training data), followed by calculation of genomic breeding values (GEBV) using only marker genotypes (and previously estimated SNP effects) in the population for selection (validation data). Mixed model methods and Bayesian MCMC (Markov Chain Monte Carlo) methods have been recommended for genomic selection. Bayesian MCMC methods generally have the highest accuracy of predicting GEBV but are slow computationally (Lund *et al.* 2009). An Expectation Maximisation (EM) algorithm can use valuable information in a prior distribution as in a Bayesian approach and is usually much faster. This paper describes an EM algorithm called emBayesB for genomic selection and explains visually the features inherent in emBayesB using the technical computer software Mathematica®.

EM APPROACH FOR 1 SNP

It is instructive to first visualise the estimation of the effect of one SNP. Then the algorithm is extended to the estimation for *m* SNP where *m* is usually much larger than *n*, the number of individuals.

Data model for 1 SNP. The linear model $\mathbf{y} = \mathbf{b}g + \mathbf{e}$ is used to relate record y_i of individual *i* to the SNP effect g where element b_i of vector $\mathbf{b}$ is the number (0, 1 or 2) of reference SNP alleles for individual *i*. We standardise $\mathbf{b}$ so that $\mathbf{1}'\mathbf{b} = 0$ and $\mathbf{b}'\mathbf{b} = n$. The errors are assumed normal and independent so that $\mathbf{y} \mid g \sim N\left(\mathbf{b}g, \mathbf{I}\sigma_e^2\right)$. Using maximum likelihood (ML) we find that the likelihood distribution of the estimate of the SNP effect g given the data is normal ie. $g \mid \mathbf{y} \sim N\left(g_L, \sigma^2\right)$ where the *best* estimate of the SNP effect is the mean g_L $(= \frac{1}{n}\mathbf{b}'\mathbf{y})$ which is the weighted data average and $\sigma^2 = \frac{1}{n}\sigma_e^2$. The two likelihoods displayed in Figure 1A (for $g_L = 0.6$ and $\sigma_e^2 = 1$) illustrate the finding that, as *n* increases, the likelihood becomes narrower ie. we are more confident about the ML (or *best*) estimate of g with more information.

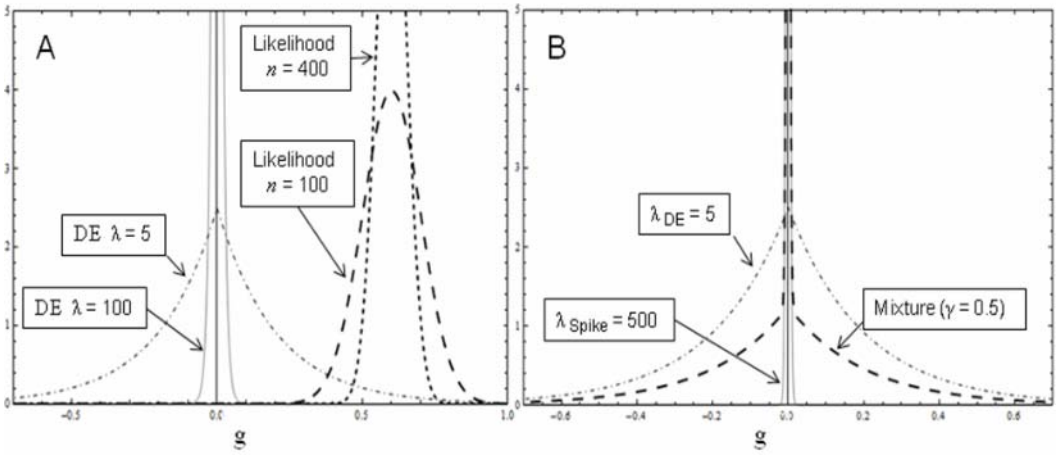


Figure 1. A. Normal likelihoods and DE distributions. B. Prior mixture for g when $\gamma = 0.5$

Prior distribution for 1 SNP. It is assumed *a priori* that SNP effect g has a probability $1-\gamma$ of being 0 and a probability γ **(**due to LD with QTL) of being distributed as a Double Exponential (DE) with parameter λ ie. $p(g) = 0.5\lambda \exp(-\lambda \mid g \mid) = PDF_{DE}$. Figure 1A shows the DE shape for two values of λ. The prior for g can be written as a mixture $\pi(g) = \gamma PDF_{DE} + (1-\gamma)\,\delta(g)$ where $\delta(g)$ is the Dirac Delta function which has all its probability mass at 0. Figure 1B shows the prior mixture for g using a Spike (or DE with $\lambda_{Spike} = 500$) for the Dirac Delta function (as a Dirac Delta function cannot be easily graphed) and a DE with $\lambda_{DE} = 5$ for the 50% ($\gamma = 0.5$**)** chance of the SNP being in LD with QTL. As the Spike's λ gets larger and γ gets smaller, the prior mixture is often described as a 'spike and slab' prior (see prior mixture in Figure 2A).

Posterior distribution for 1 SNP. The posterior for g is illustrated in Figure 2. When the likelihood estimate (g_L) is distant from 0 the posterior distribution resembles the likelihood distribution, but is slightly displaced (or regressed) toward 0 as shown in Figure 2A. When g_L is much greater than 0, the mode of the regressed likelihood is $g_L - \lambda_{DE}\sigma^2 \left(= DE_{mode}\right)$ which is the posterior mode for a DE only prior, ie. the LASSO estimate (Yi and Xu 2008) of a SNP effect, as the Spike has no influence if g_L is distant from 0. As g_L gets closer to 0 the posterior becomes bimodal, with the height at 0 increasing the closer g_L is to 0 (Figure 2). This reflects the fact that the true g is more probably 0, the closer g_L is to 0. Using Mathematica® it can be shown that the area under the DE part of the posterior is 0.99, 0.60 and 0.14 for g_L values of 0.2, 0.15 and 0.1 respectively, assuming the parameter values given in Figure 2. These DE areas are basically the posterior probabilities of g being non-zero given the assumed or current parameter estimates. The

posterior probability γ_{post} (see Figure 2) of a SNP effect being non-zero (ie. in LD with at least one QTL) form the E-step of the EM algorithm for genomic selection called emBayesB.

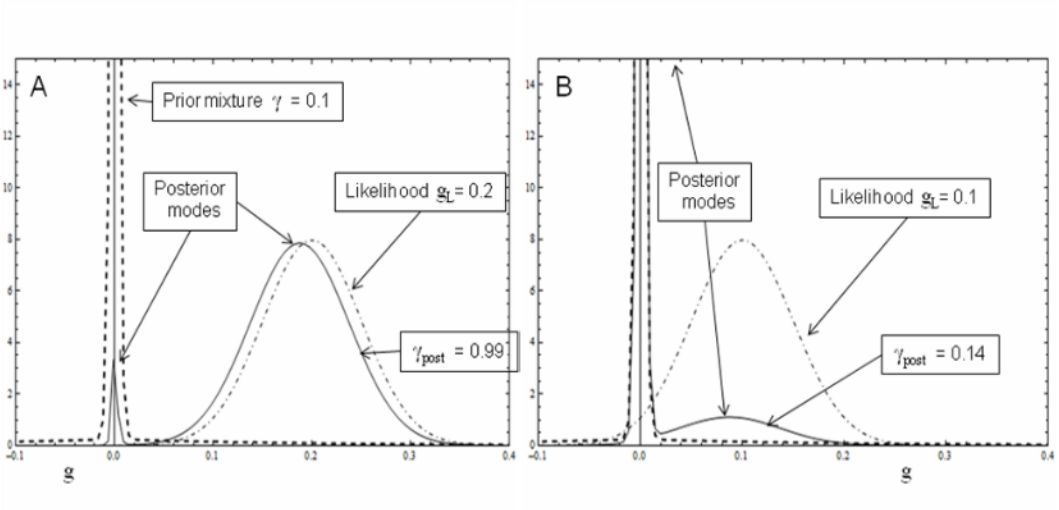


Figure 2. Bimodal posterior distribution of g as g_L approaches 0 for a 'spike and slab' prior where γ, λ_{DE}, λ_{Spike}, σ_e^2 and n are 0.1, 5, 500, 1 and 400 respectively.

EM ALGORITHM FOR MANY SNP

The data model for *m* SNP is $\mathbf{y} = \mathbf{Bg} + \mathbf{e}$ which linearly relates record y_i of individual *i* to the j^{th} SNP effect g_j where element b_{ij} of matrix $\mathbf{B}$ is the number (0, 1 or 2) of reference alleles of SNP *j* for individual *i* (usually standardised). The errors are assumed normal and independent such that $\mathbf{y} \mid \mathbf{g} \sim N\left(\mathbf{Bg}, \mathbf{I}\sigma_e^2\right)$. If we knew precisely which SNP were in LD with QTL (maybe only 100 SNP), then the problem would be much easier. This missing information is crucial in formulating an EM algorithm. We define an unknown variable z_j which indicates if the j^{th} SNP is in LD with QTL ($z_j = 1$) or not ($z_j = 0$). If $z_j = 1$, the SNP effect g_j is assumed to be a Double Exponential random variable with parameter λ; while if $z_j = 0$, the SNP effect is assumed to be distributed as a Dirac Delta (DD) function which has all its probability mass at 0. We assume *a priori* that a fraction γ of the SNPs are in LD with QTL.

Using EM theory we are able to develop an iterative sequence of E and M-steps which converge to maximum *a posteriori* (MAP) parameter estimates. At iteration *k* the E-step involves calculating $\gamma_j^k \left(= E\left[z_j \mid \mathbf{y} \,\&\, \textit{all current estimates}\right]\right)$, the posterior probability that SNP *j* is in LD with QTL given the data and all current parameter estimates (eg. like calculating γ_{post} in Figure 2). This is done analytically and fast. Then the M-step uses derived formulae (not shown here) to calculate updated estimates of $g_j, \gamma, \lambda, \sigma_e^2$ given the data and the current values of γ_j^k. This step is also done very quickly using Gauss-Seidel iteration for the many estimates of g_j. Iterating

between the E and M-steps the algorithm converges quickly to produce MAP estimates of g_j, ML estimates of $\gamma, \lambda, \sigma_e^2$ and posterior probabilities γ_j^k (which are useful for mapping QTL).

DISCUSSION

The derived formula $\hat{g}_j = \gamma_j^k DE_{mode} + \left(1-\gamma_j^k\right) DD_{mode} = \gamma_j^k DE_{mode}$ gives the MAP estimate of g_j. This formula shows that the best estimate of a SNP effect is a weighted average of the two posterior modes. However the mode (DD_{mode}) for a Dirac Delta only prior is always 0. So we have that the best estimate of the SNP effect with emBayesB is a proportion of the DE mode. But genetic gain is greatest if the posterior mean is used to estimate each QTL effect (Goddard and Hayes 2007). Using Mathematica®, we can show for Figure 2 that the posterior mean, for g_L values of 0.2, 0.15 and 0.1, is 0.1848, 0.0829 and 0.0124 respectively, while the weighted average of the two posterior modes is 0.1847, 0.0827 and 0.0120 respectively. Hence the weighted average of the two posterior modes is an accurate estimate of the posterior mean of a SNP effect. Bayesian MCMC methods use the estimated posterior mean of each SNP effect in the prediction equation $\mathbf{GEBV = B\hat{g}}$, whereas emBayesB uses the weighted average of the two posterior modes. Hence it is no surprise to find that the accuracy of 0.85 between GEBV and true breeding value for emBayesB is similar to the accuracies of 0.84 to 0.87 for Bayesian MCMC methods when analysing the validation data of the QTLMAS XII dataset (Shepherd *et al.* 2009).

emBayesB works by shrinking the ML estimates of the SNP effects. If the ML estimate is distant from 0 the shrinkage is mainly due to the DE prior and the shrunken estimate is called the LASSO (Yi and Xu 2008). The closer the ML estimate is to 0, the greater is the shrinkage. This is due to the Dirac Delta prior kicking in and reflects the fact that only a small proportion (ie. γ) are believed non-zero *a priori*. The algorithm combines this prior belief with the data to iteratively derive a probability for each SNP of being non-zero. Then it further regresses the DE_{mode} (or LASSO estimate) for this SNP by its probability of being non-zero. It is this double shrinkage which makes emBayesB so accurate and able to handle all the noise in the data from having so many SNP most of which aren't in LD with QTL. Basically it removes the effects of lots of SNP from the genomic breeding value as these SNP are most probably not in LD with QTL.

Not only is emBayesB accurate when predicting breeding values but it is also very fast. emBayesB uses Gauss Seidel iteration to quickly calculate an analytical *posterior-like* mean for for each SNP effect and iterates until the SNP estimates converge. Bayesian MCMC methods sample sequentially from distributions which eventually converge to the true posterior distribution for each SNP. As there are thousands of SNP, the SNP distributions take a long time to converge.

ACKNOWLEDGMENTS

This paper reports collaborative research instigated while RKS was on sabbatical at the Roslin Institute with support from CQUniversity and the Roslin Institute.

REFERENCES

Goddard, M.E. and Hayes, B.J. (2007) *J. Anim. Breed. Genet.* **124**:323

Lund, M.S., Sahana, G., de Koning, D-J., Su G. and Carlborg, O. (2009) *BMC Proc.* **3**(Suppl 1):S1

Meuwissen, T.H.E., Solberg T.R., Shepherd R., Woolliams J.A. (2009) *Genet. Sel. Evol.* **41:**2

Shepherd R.K., Meuwissen, T.H.E. and Woolliams J.A. (2009) *Proc. Assoc. Advmt. Anim. Breed. Genet.* **18**: 84.

Yi, N. and Xu, S. (2008) *Genetics* **179**:1045

GENOMIC SELECTION USING A FAST EM ALGORITHM 2. ANALYSIS OF SIMULATED DATA

R.K. Shepherd[1], T.H.E. Meuwissen[2] and J.A. Woolliams[3]

[1]FABIE, CQUniversity, Rockhampton, QLD 4702
[2]IAAS, Norwegian University of Life Sciences, Box 5003, N1432 As. NORWAY
[3]The Roslin Institute & R(D)SVS, University of Edinburgh, Roslin, Midlothian, EH25 9PS UK

SUMMARY

The paper reports on a fast EM algorithm for genomic selection by mapping QTL in genome-wide dense SNP marker data. The algorithm called emBayesB was used to analyse a 6000 SNP dataset simulated for the QTLMAS XII workshop. True breeding value was accurately predicted by GEBV with a correlation of 0.85 in the validation data, while the regression coefficient of 0.97 indicated unbiased predictions of breeding value. The results were similar to Bayesian MCMC estimates but were calculated in a fraction of the time. emBayesB was also able to accurately map the location of individual QTL which explain more than 1% of the total genetic variation.

INTRODUCTION

Genomic selection is a recent tool for genetic improvement in animal breeding. It involves the use of dense DNA markers covering the whole genome so that all QTL are in linkage disequilibrium (LD) with at least one marker. It has recently become economically feasible due to the commercial availability of dense genotyping chips, containing thousands to millions of single nucleotide polymorphisms (SNP), and the development of high throughput systems, all at cost effective prices. Genomic selection involves two main steps: first the estimation of SNP effects using phenotype and genotype data in a reference population (training data), and then the prediction of genomic breeding values (GEBV) using only marker genotypes (and the previously estimated SNP effects) in the population under selection (validation data).

Mixed model (or GS-BLUP) methods and Bayesian MCMC (Markov Chain Monte Carlo) methods are the main contenders for calculating GEBV. Bayesian MCMC methods generally have a higher accuracy of predicting GEBV than GS-BLUP methods but are slow computationally (Lund *et al.* 2009). The prior information used in a Bayesian approach can be incorporated in an Expectation Maximisation (EM) algorithm through the calculation of a posterior mode. Also EM algorithms can be significantly faster than Bayesian MCMC methods if well formulated. This paper reports on a fast EM algorithm for genomic selection by mapping QTL in genome-wide dense SNP marker data. The algorithm is called emBayesB because it is an EM implementation of the important features of the BayesB method of Meuwissen *et al.* (2001).

MATERIALS AND METHODS

Data simulation. The data analysed was the QTLMAS XII common dataset. Full details are available in Lund *et al.* (2009). Initially a population of 100 founders (50 of each sex) was simulated. For the next 50 generations, 100 progeny (50 male and 50 female) were produced by randomly sampling parents. Then for the next and last 6 generations, 15 males and 150 females were randomly selected for a hierarchical mating to produce 100 progeny per male and 10 progeny per female, giving a total of 1500 pedigreed progeny per generation. The validation data of 1200 individuals with only genotype records was produced using the last 3 generations by randomly selecting 400 progeny per generation. The training data of 4665 individuals with genotype and phenotype records consisted of the progeny from the preceding 4 generations. There were 6000

biallelic markers at 0.1 cM spacing on the six 100cM chromosomes, giving 1000 markers per chromosome. The two alleles at each marker were sampled with equal probability in the founders. QTL effects were sampled from a gamma distribution. The genomic location and allele substitution effects of the 48 simulated biallelic and additive QTL are shown in Figure 2. The number of QTL, which explain more than 0.1, 1, 5 and 10% of the total genetic variation in the training data, was 28, 15, 6 and 4 respectively. An individual's true breeding value (TBV) was the sum of the effects of all of the individual's QTL. A trait with heritability of 0.3 was produced by adding a normally distributed error term to the TBV of each individual.

EM algorithm. The methodology used in the EM algorithm emBayesB is presented in the companion paper (Shepherd and Woolliams 2009) and so is not reported here.

Validating emBayesB. The algorithm emBayesB used phenotypes and SNP genotypes of the 4665 individuals in the training data set to calculate the prediction equation $\mathbf{GEBV = B\hat{g}}$ by Gauss-Seidel iteration. The number of SNP analysed was 5726 as only SNP with a minor allele frequency greater than 0.05 were used. The initial parameter estimates assumed for the EM algorithm were $\hat{g}_j = 0$, $\hat{\gamma} = 0.01$, $\hat{\lambda} = 7.8$ and $\hat{\sigma}_e^2 = 1.4$, some of which result from an assumed heritability of 0.5 and a total phenotypic variance of 2.8. The prediction equation was used to calculate the GEBV of the 1200 individuals in the validation data set using only the genotype of their 5726 SNP. The correlation between TBV and GEBV was calculated for the validation data as well as the linear regression of TBV on GEBV, which has a slope of 1 if the GEBV are unbiased.

RESULTS AND DISCUSSION

The main advantage of emBayesB over Bayesian MCMC methods is the speed of computation. The Bayesian methods are computationally slow as they use intense MCMC techniques (eg. Gibbs sampling) to obtain, after considerable computation, random samples from each SNP's posterior distribution for the thousands of unknown SNP effects. The emBayesB algorithm is fast as it uses Gauss-Seidel iteration to compute a weighted analytical mean of the modes in each SNP's posterior distribution as the best estimate of a SNP effect (Shepherd and Woolliams 2009). Using a Bayesian MCMC approach for this data took 2 days on a Unix box (R. Pong-Wong, pers. comm.) whereas emBayesB converged in a few minutes on a laptop PC.

Using emBayesB produced a correlation of 0.85 between TBV and GEBV for the validation data. This correlation is considerably larger than the range of correlations of 0.5 to 0.77 for the GS-BLUP models used on the same data, but similar to correlations of 0.84 to 0.87 for Bayesian MCMC methods, as reported by Lund *et al.* (2009). One Bayesian method used a 5 SNP haplotype to produce a correlation of 0.92. But the correlation for this method would usually be smaller as haplotypes are estimated with error in real life, not known exactly as in this data. Using emBayesB produced a slope of 0.97 for the regression of TBV on GEBV in the validation data. This slope indicates unbiased predictions of breeding value which was also found for the Bayesian MCMC methods but in general not for the GS-BLUP methods. The Bayesian MCMC methods assume a SNP mixture distribution in which relatively few markers explain a large variance, while a large number of markers explain a very small variance. Due to the similar information in the priors of emBayesB and the Bayesian MCMC methods, it is no surprise that they produce similar predictions as measured by the accuracy and the regression coefficient of TBV on GEBV.

As expected most SNP (5702 in total) have small posterior probabilities of being in LD with QTL (Figure 1). A surprising result was that only 24 of the 5726 SNP had posterior probabilities greater than 0.1. On chromosome 6 all SNP had posterior probabilities less than 0.06 which is due to the absence of QTL on this chromosome. emBayesB detects all 15 QTL with allele substitution

effects greater than 0.2 by calculating posterior probabilities of 0.72 or more for nearby SNP (Figure 2).

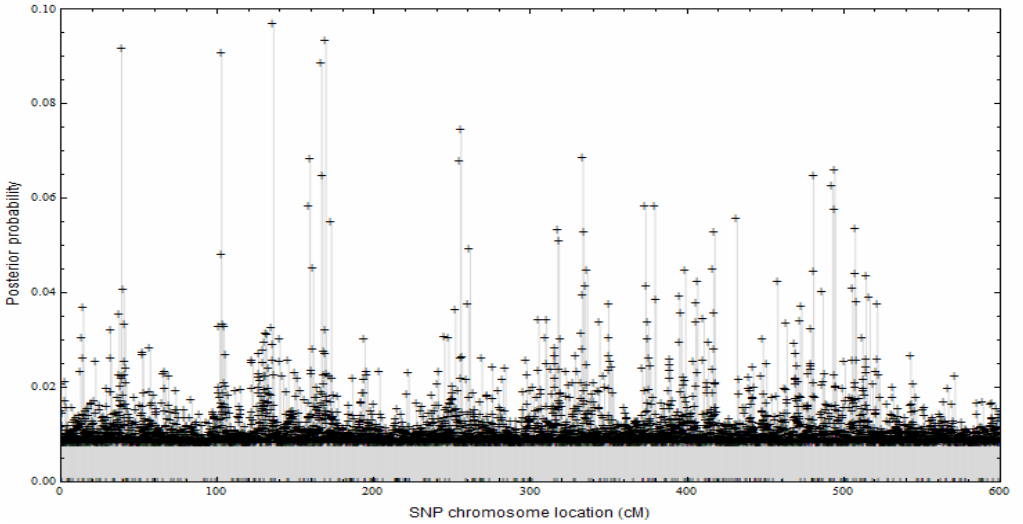


Figure 1. Posterior probability (+) for 5702 SNP of each being in LD with at least one QTL.

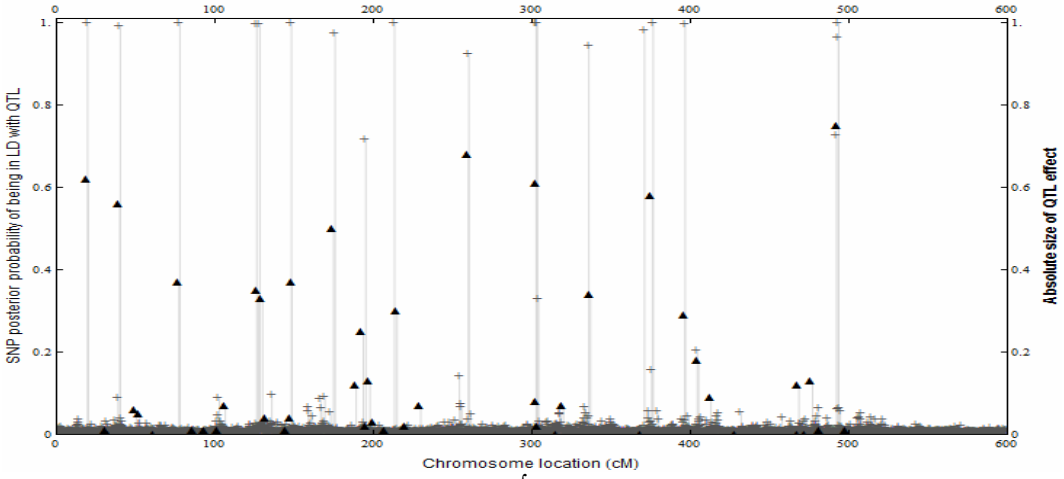


Figure 2. Absolute value of the QTL effect (∫) for the 48 simulated QTL plus the posterior probability (+) for 5726 SNP of each being in LD with at least one QTL.

There are 15 QTL which each explain more than 1% of the total additive genetic variation (V_A) and in total, explain over 95% of the total V_A. emBayesB detected each of these 15 QTL (Figure 3). The distance from each of the detected QTL to the nearest high probability SNP averaged 0.7cM, with the largest distance being 1.7cM. As the genetic variation explained by a QTL dropped below 1% so did the posterior probability of a SNP being in LD with it (Figure 3). It was found that the EM estimates of the high probability SNP effects were not very accurate for estimating the effects of individual QTL. However using only SNP with a large posterior probability as fixed effects in a multiple linear regression resulted in QTL estimates which were

quite accurate. Basically it seems that the posterior probabilities can be used to screen out most of the SNP, leaving only those in LD with QTL, for a multiple regression approach if estimates of large (> 1% V_A) QTL effects are required.

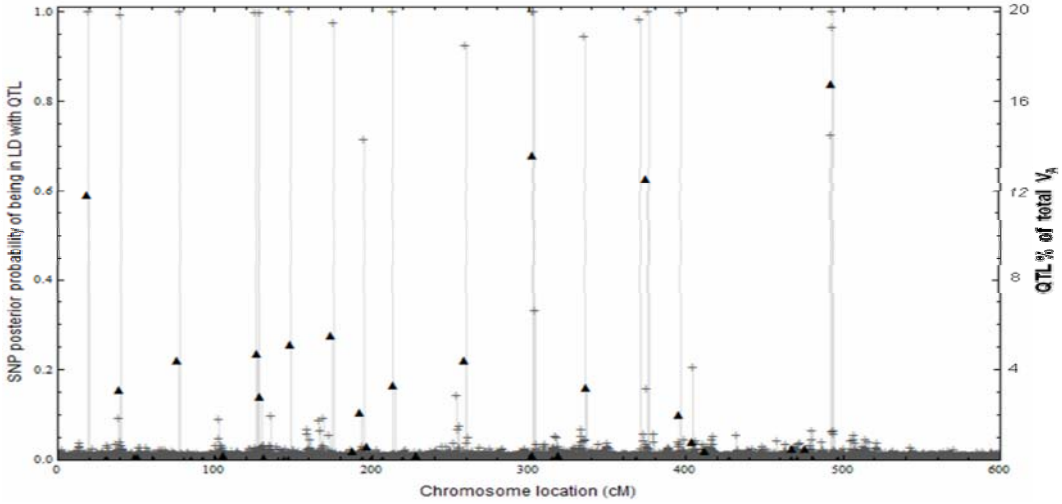


Figure 3. Percentage of total V_A (ĵ) explained by each of the 48 QTL plus the posterior probability (+) for 5726 SNP of each being in LD with at least one QTL.

In general emBayesB found only one nearby SNP with a high posterior probability for each of the 15 QTL which individually explain more than 1% of the total V_A (Figure 3). However multiple high probability SNP were reported nearby for QTL at 303cM, 376cM and 493cM. The reason for these multiple nearby high probability SNP seems to be related to the fact that these QTL explain more V_A than any other QTL and that this larger QTL variation induces higher correlation among neighbouring SNP through LD. Further investigation is needed to determine if this result is a feature of emBayesB.

As the size of SNP panels increase, the Bayesian MCMC methods will become even slower. Further research is required to see if emBayesB will be a suitable algorithm for analysing data from large SNP panels as convergence is not guaranteed with the Gauss-Seidel iteration.

CONCLUSIONS

emBayesB is a fast and accurate EM algorithm for implementing genomic selection by mapping QTL in genome-wide dense SNP marker data. Its accuracy is comparable to Bayesian MCMC methods but it takes only a fraction of the time. The current study analysed only one replicate of a common dataset so there is a need to analyse more data to understand its capabilities.

ACKNOWLEDGMENTS

This paper reports collaborative research instigated while RKS was on sabbatical at the Roslin Institute with support from CQUniversity and the Roslin Institute.

REFERENCES

Lund, M.S., Sahana, G., de Koning, D-J., Su G. and Carlborg, O. (2009) *BMC Proc.* **3**(Suppl 1):S1
Meuwissen, T.H.E., Hayes, B.J. and Goddard, M.E. (2001) *Genetics* **157:**1819.
Meuwissen, T.H.E., Solberg, T.R., Shepherd, R., and Woolliams, J.A. (2009) *Gen. Sel. Evol.* **41:**2.
Shepherd R.K. and Woolliams J.A. (2009) *Proc. Assoc. Advmt. Anim. Breed. Genet.* **18**: 80.

MAPPING THE HORNS LOCUS IN SHEEP

N.K. Pickering, P.L. Johnson, B. Auvray, K.G. Dodds and J.C. McEwan

Invermay Agricultural Centre, AgResearch, Mosgiel

SUMMARY

The horns locus has been mapped to a 200kbp region in sheep using a segregating (Merino x Romney) x Merino resource. This region contains a single candidate gene. Maximal LODs exceeded 110 for the best location. Thirty-two SNPs spanning 400kbp were subsequently genotyped over 1000 sheep from a wide variety of polled and horned breeds. A 17 SNP haplotype was associated with the polled locus which had near perfect (>97%) concordance with phenotype. The gene is currently being sequenced.

INTRODUCTION

Undomesticated sheep typically have horns. Horns are present in most *Bovidae* and resemble modified hooves. In wild populations they are used in males for mate competition in the breeding season and for feed competition in females in the peri-parturient period (Clutton-Brock and Pemberton 2004). The frequency of hornless (polled) sheep originally increased during domestication 7000-9000 yrs BP, and dramatically in English breeds after 1500 AD. Many breeds are now polled, although certain breeds, such as the Merino, still have horns. Previous work provided evidence of a Horns (Ho) locus on sheep chromosome OAR10 (Montgomery *et al.* 1996). Previous work at the Ho locus suggests it has at least three alleles: normal horns (H^+), sex limited horns (H^L), and polled (H^P). The phenotype depends on sex and alleles inherited (Figure 1). The expression of the H^+ and H^L allele can also be modified by castration as horn growth ceases immediately at this time (Marshall and Hammond 1914). Horns are generally considered an undesirable trait by sheep farmers and are selected against in most breeds. A DNA test for this locus would therefore be beneficial for two reasons. Firstly, normal Ho locus expression appears to be masked by modifiers present in dual purpose composite NZ breeds resulting in unpredictable inheritance. It is thought that these modifiers were originally present in Scandinavian breeds (A. Bray pers. comm.). Secondly, wool shedding and extended breeding season traits are being introgressed into dual purpose animals from horned breeds such as Wiltshires and Dorset Horn. These breeders want an efficient method of retaining the desirable traits while eliminating horns.

A

	Ho^LHo^L ♀ Scurred	
Ho^LHo^P ♂ Horned	Ho^PHo^L ♂ Horned	Ho^LHo^L ♀ Scurred
	Ho^PHo^L ♂ Horned	Ho^PHo^L ♀ Polled

B

	Ho^+Ho^+ ♀ Horned	
Ho^+Ho^P ♂ Horned	Ho^PHo^+ ♂ Horned	Ho^+Ho^+ ♀ Horned
	Ho^+Ho^+ ♂ Horned	Ho^PHo^+ ♀ Scurred

Figure 1. Inheritance of horns adapted from Clutton-Brock & Pemberton (2004): two models; sex-limited horns (A), and normal horns (B). Ho^PHo^P is polled for males and females in both models. P: polled, L: sex-linked horns, and +: normal horns.

MATERIALS AND METHODS

Microsatellites. The seminal report that mapped Ho locus to OAR10 used female offspring from Merino X Romney rams backcrossed to Merino dams i.e. (MxR)xM (Montgomery *et al.* 1996). Only polled and sex limited horn alleles were segregating in this resource. We screened a superset of this resource with additional years of female (MxR)xM progeny with six microsatellite markers, three of which had been used in the initial study, in order to localise a region of interest.

There were 360 F2 female offspring genotyped along with their four F1 sires and six available paternal grandparents. Horns were scored using the Dolling scale (Dolling 1970) on their presence and size.

Subsequently, 71 (MxR)xM females that had a recombination event in the region of interest (breakpoint panel) were then genotyped with a further 24 microsatellite markers spanning the region identified. In addition a further 22 animals from a separate breed resource of horned (Merino, Dorset Horn and Wiltshire Horn) and polled (Texel, Gotland, Romney) animals were also genotyped. These animals were identified by breed but often did not have individual horn phenotypes recorded.

An additional resource ("breed standards", n=698) were then genotyped with the four most informative markers. This resource consisted of animals from the following horned (NZ Wiltshire, Dorset Horn, Soay, Spanish Merino, French Merino, Awassi, Finnish Landrace, and Bighorn) and polled (Texel, Poll Dorset, Coopworth, Romney, Suffolk, Perendale, Corriedale, East Friesian, Cheviot, Gotland, and Fleischschaf) breeds. Horns status was determined from known breed characteristics. Microsatellite markers were developed from the bovine genome. The data was analysed via multi-point linkage analysis using CRI-Map in order to position the Ho locus.

SNPs. Thirty-two SNPs spanning the most likely 400kbp region containing the Ho locus were then genotyped over the breakpoint panel and breed standards. These SNPs were developed from sequence generated from initial 454 ISGC ovine genome sequencing (McEwan *et al.* 2009), and three ovine BACs selected to span the region of interest. The combined data set was then analysed via multi-point linkage analysis using CRI-Map in order to refine the location of the Ho locus.

Sequencing. Some ovine sequence for our gene of interest was available from the previously described sequencing. This was then used to design primers to sequence the mRNA and DNA isolated from a polled (Coopworth x Texel) and horned (NZ Wiltshire) ram. The last seven and the first exon of the gene was sequenced using DNA. The ovine mRNA was converted to cDNA and also sequenced.

RESULTS AND DISCUSSION

Microsatellites. From the six markers genotyped, the Ho locus was localised to a 14Mbp region (Maximal LOD 108). The second and third rounds of fine mapping with an additional 24 markers narrowed the Ho locus region down to a 200kbp region (maximal LOD 110; figure 2). Within this region a four marker haplotype was observed that best explained polled versus horned animals. This haplotype was validated on the breed standard animals and was estimated to have 92% accuracy.

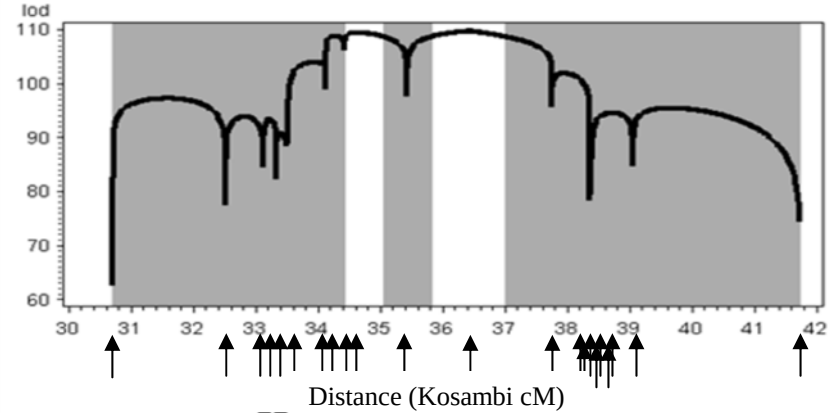


Figure 2. Multi-point linkage graph of the horns locus on OAR 10, after microsatellite genotyping. Marker BM6108 was the origin and arrows indicate microsatellite positions. White region indicates the 1 LOD score drop-off region (200kbp).

SNPs. Using the thirty-two SNP markers genotyped, the Ho locus was further localised to a 50kbp region (maximal LOD 110, figure 3). Within this region resides a single candidate gene. Analysis of breed phenotypes generated a haplotype of 17 markers for the polled haplotype; with a >97% concordance with phenotype. This haplotype extended over the 3' end of the candidate gene.

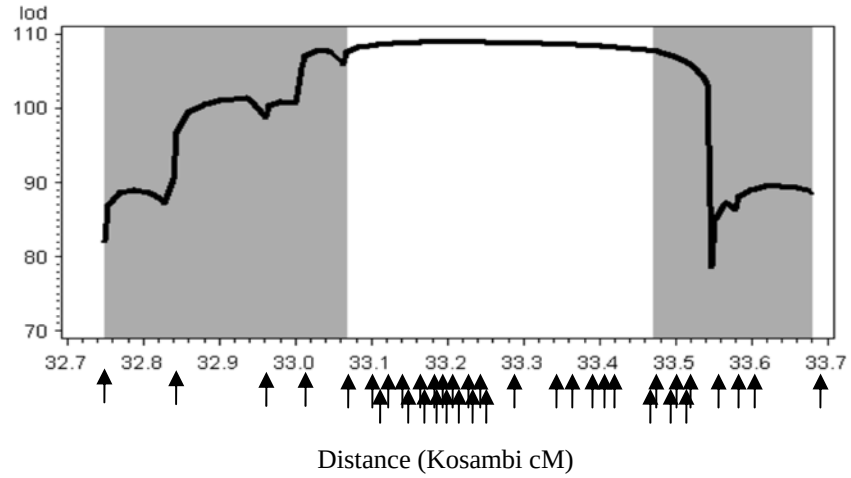


Figure 3. Multi-point linkage graph of the horns locus on OAR 10, after SNP genotyping. Marker BM6108 was the origin and arrows indicate relative SNP positions. White region indicates the 1 LOD score drop-off region (50kbp). Minor differences in scale are due to additional markers used to build the linkage map.

Sequencing. Initial sequencing identified a large ~3kb insert in the non coding 3' untranslated region of the candidate gene. This insert seemed to be present only in polled animals. The sequence of the insert indicated that a complete functional mRNA has been retrotransposed in the

reverse orientation. Comparison of the Coopworth sequence with bovine, human and Wiltshire identify a flanking thirteen base pair direct repeat. This is characteristic of a target site duplication commonly associated with the integration of a transposable element. Studies are now underway to validate this in a variety of breeds and horned phenotypes. Insertions of this nature are essentially unique within a species and it may be that the majority of polled animals in the world descend from a single common ancestor.

CONCLUSIONS

The horns locus has been fine mapped to a 50kb region in the sheep genome and predictive haplotypes for the polled H^P allele identified. A single candidate gene underlies this region and preliminary work has identified a large retrotransposed insertion in the 3' untranslated region of this gene in polled animals. This insertion appears to be functional, would be transcribed in the opposite orientation and is flanked by 13bp direct repeats. Our hypothesis is that a polled mutation has occurred in a single ancestor and this has subsequently been selected to near homozygosity in a wide variety of polled breeds. In contrast horned animals still display the full range of ancestral variability in this region.

ACKNOWLEDGMENTS

The various breeders and researchers who collected and provided the DNA samples utilised. Gemma Payne for the BAC sequencing. This research was supported by Ovita.

REFERENCES

Clutton-Brock, T., and Pemberton, J. (2004). "Soay sheep dynamics and selection in an island population", Cambridge University Press, Cambridge, UK.

Dolling, C. H. S. (1970) In "Breeding Merinos", Rigby, Adelaide, Australia.

Marshall, F., and Hammond, J. (1914). *J. physiol*. **48**:171.

McEwan, J., and International Sheep Genomics Consortium. (2009) *PAG XVII*. www.intl-pag.org/17/abstracts/W11_PAGXVII_081.html

Montgomery, G. W., Henry, H. M., Dodds, K. G., Beattie, A. E., Wuliji, T., and Crawford, A. M. (1996). *J. Heredity*. **87**:358.

GENETIC MARKERS FOR POLLED CONDITION IN CATTLE – THE CURRENT STATUS AND THE FUTURE PLANS

K.C.Prayaga[1], M. Mariasegaram[1], B. Harrison[1], B. Tier[2], J.M. Henshall[3] and W. Barendse[1]

Cooperative Research Centre for Beef Genetic Technologies
[1]CSIRO Livestock Industries, Queensland Bioscience Precinct, St. Lucia, QLD 4067
[2]Animal Genetics and Breeding Unit*, University of New England, Armidale, NSW 2351
[3]CSIRO Livestock Industries, FD McMaster Laboratory, Armidale, NSW 2350

SUMMARY

Breeding polled cattle is a non-invasive alternative to the practice of dehorning. Horn status is presumably under the control of three genes i.e. polled, African horn, and scurs, especially in *Bos indicus* cattle. A research project was initiated with an aim to develop markers for determining the genetic status of genes underpinning the inheritance of horns. The results from an initial whole genome scan and a targeted fine scale mapping experiment led to the development of a validated molecular marker predicting the genetic status at the polled locus with a high degree of accuracy (>90%). Currently, accurate phenotypes of horn status are being recorded on the progeny generated on a collaborating Brahman breeding property. Future research plans involve whole genome scan of this resource population to understand scurs inheritance and to ascertain the existence of African horn gene.

INTRODUCTION

Horns in cattle increase the risk of injury to animal handlers. They also have an economic impact through carcass bruising and hide damage. Thus dehorning has become an accepted management practice to overcome the problems associated with horns. However, owing to the extensive nature of farming practices in northern Australia, dehorning occurs in older calves of 3.5 to 10 months of age (Prayaga 2007). This has animal welfare implications. Further, the practice of dehorning has economic costs to the beef enterprise such as the cost of labour of dehorning and the cost of secondary infection and mortality, albeit in rare instances. Therefore, breeding polled cattle is a non-invasive genetic selection option to replace the practice of dehorning.

The inheritance of horns is hypothesised to be relatively complex in *Bos indicus* animals with three interacting genes i.e. polled, African horn and scurs (small loose horns) genes. The polled locus has two alleles – an allele for poll (*P*) and an allele for horn (*p*). The gene for horn is recessive so that, in the absence of other interfering loci, homozygous individuals (*pp*) are horned and other genotypes (*PP*, *Pp*) are polled. The African horn gene (*Ha/ha*) is supposedly at a higher frequency in Zebu cattle and has a sex-influenced epistatic effect on the 'polled' locus so that male heterozygotes are horned and female heterozygotes are polled. A third gene called scurs controls the development of small and loosely attached horns. The scurs gene is also thought to be sex-influenced.

The polled gene has been mapped to the centromeric region of bovine chromosome 1 (BTA1) by several groups (Georges *et al.* 1993; Brenneman *et al.* 1996). Recently, Drögemüller *et al.* (2005) used twenty microsatellite markers generated from random sequencing of BAC subclones to positionally clone the *polled* locus in *Bos taurus* cattle and mapped the locus to a 1 Mb interval between microsatellite markers *RP42-218J17_MS1* and *BM6438* on BTA1. Asai *et al.* (2004) have mapped the scurs gene to BTA19; however, this has not been confirmed later. Importantly, the scurs phenotype can only be discerned in genetically polled animals (Prayaga 2007).

---

* AGBU is a joint venture of NSW Department of Primary Industries and the University of New England

A research project with an aim to develop genetic markers for polled, African horn and scurs genes was initiated to enable faster introgression of polled condition in tropical beef cattle. This paper reports the progress made in developing these markers and outlines future research.

MATERIALS AND METHODS

Whole genome scan. A group of unrelated Brahman (n=68) and Hereford (n=20) cattle were selected across various industry and research herds with a view to exploit linkage disequilibrium based association. The Hereford sample constituted 10 polled and 10 horned with equal number of males and females in each. However, sex was deliberately confounded with horn status in the Brahman sample to increase the power of detecting allele associations given the sex-influenced pattern of inheritance of scurs and African horn. Thus, the Brahman sample consisted of 33 polled males and 35 horned females. The DNA samples from this resource were genotyped with Affymetrix GeneChip® Bovine mapping 25K panel and Brahman specific 11.5K panel. In total, 34197 assays returned genotypes from both panels. SNP with minor allele frequencies less than 0.1 were omitted from analyses. Statistical analyses involved testing the chi-square statistic for significance given the observed and expected frequencies of three genotype classes for each SNP within each breed. Because of the prior knowledge, SNP located on BTA 1 only were included for analyses of Hereford data. Appropriate threshold p-values for significance testing for a false positive rate of 1% were derived to correct for the effect of multiple testing.

Fine mapping polled locus. A concurrent experiment to fine map the polled locus was also conducted on the same animals used for whole genome scan. This was conducted in two stages. Initially, 17 microsatellites described by Drögemüller *et al.* (2005) encompassing BTA1 region between 0.9 and 3.8 Mb were genotyped. With an additional designed marker, this yielded data from 18 microsatellites for the haplotype analysis. Based on the initial results, a further 15 highly polymorphic microsatellite markers were genotyped in the region in the second stage of the experiment. These markers were identified by feeding the identified genomic sequence into Sputnik sequence annotation pipeline (http://mips.gsf.de/proj/sputnik/) and embedded within the Bovine Genome Browser of CSIRO. Genotype scoring was performed using Genemapper™ ver4.0 from Applied Biosystems. Finally, genotype data from 39 SNP (from whole genome scan) and 33 microsatellites from the above-mentioned region were available for analyses. Haplotype reconstruction was carried out using PHASE ver2.1 (Stephens *et al.* 2001; Li and Stephens 2003) that uses Bayesian method to impute haplotypes. In addition to deriving hapolotype frequencies and recombination rates, a permutation test was performed to determine the significance of differences in haplotype frequencies between case (polled) and control (horned) groups.

RESULTS AND DISCUSSION

Whole genome scan. One SNP located on BTA1 (6.3Mb) was in complete agreement with the assumed Mendelian inheritance at polled locus in Herefords (P-value=0.0001), but not in Brahmans (P-value=0.05). However, this SNP is outside the region of interest on BTA 1 based on earlier reports. Further, this SNP was later found to be either tri allelic or an artefact. In Brahmans, 8 out of 12 most significant SNP (P-value<0.0001) were located on BTA1 between 1.1 and 20.8 Mb, with some of them outside the reported region of polled locus in *Bos taurus* animals. The other significant SNP were on chromosomes 9, 10 and 29. The inability to identify a number of significant SNP within the reported polled region could be due to the insufficient number of animals, especially in Herefords, and relatively lower density of markers in the region of interest.

Fine mapping polled locus. The aim of this study was to fine-map the polled locus in Brahman by increasing the power of detection through combining SNP information from the whole genome scan study with microsatellite genotyping. The differences in haplotype frequencies between polled and horned groups were significant (P-value=0.01) in breed specific datasets and in the combined data. This supports the alternate hypothesis that the imputed polled haplotypes were more similar to other polled haplotypes than to the imputed horned haplotypes.

The PHASE analyses generated 84 and 551 haplotypes for Hereford and Brahman respectively. These haplotypes represent those with the estimated frequency greater than 1% when merging segments of consecutive loci through its partition ligation method. In the Hereford sample, out of 84 haplotypes, 76% were unique to polled and 21% were unique to horned. In the Brahman sample, out of 551 haplotypes, the observed frequencies were complementary to those observed in Hereford: 25% were unique to polled and 73% were unique to horned. This distinct difference between polled and horned haplotype frequencies in both breeds further points to the presence of polled gene in this region. The background recombination rate (population parameter, r), determined as the median estimate over 100 runs in Hereford and Brahman data, was estimated to be 4.25×10^{-6} and 4.04×10^{-5} respectively. The breed specific estimates of factors by which the recombination rate between adjacent markers (i and i-1) exceeds r is presented in Figure 1 and this depicts the propensity of certain regions as recombination hotspots. The number of markers and the number of individuals in each dataset influence these estimates.

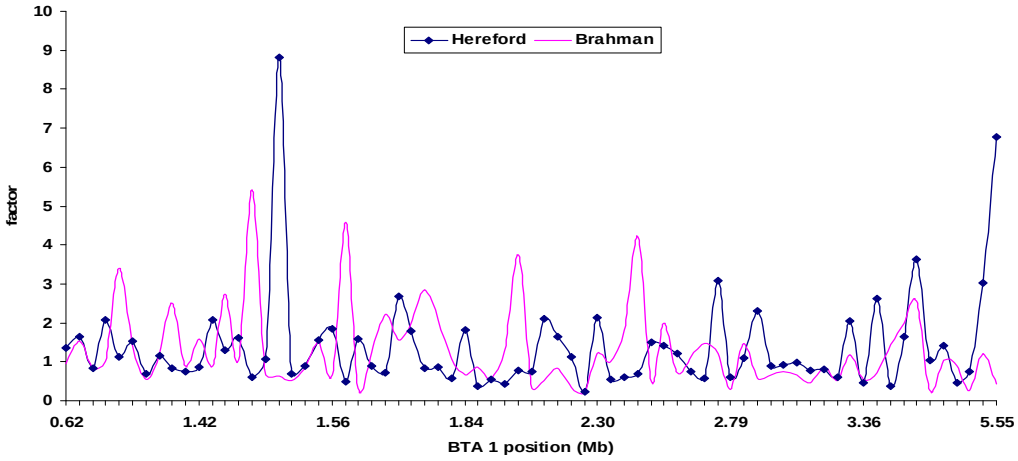


Fig 1. Estimates of variation from background recombination rate within each adjacent marker interval in Hereford and Brahman cattle

Further examination of the imputed haplotypes revealed that there is a complete association between an allele at *CSAFG29*, a newly designed microsatellite marker, and the polled condition. This particular allele is not observed in horned animals. The actual and the marker predicted horn status of the experimental population is presented in Table 1. This marker predicted all polled animals to be polled and with the exception of two animals, all horned animals to be horned. The two animals that deviated from expectation could be those with scurs and misclassified as horned at the time of dehorning. This misclassification could occur because of ambiguity associated with horn phenotype early in life.

Table 1. The actual horn status and *CSAFG29* predicted horn status of the experimental population

Breed	Actual	Predicted			Total
		Polled 'PP'	*Polled 'Pp'*	*Horned 'pp'*	
Brahman	Polled	17	16		33
	Horned		2*	33	35
Hereford	Polled	5	5		10
	Horned			10	10
Total		22	23	43	88

*deviation from expectation of zero, may be animals with scurs recorded as horned

CONCLUSIONS AND FUTURE PLANS

Breeding polled cattle could be achieved at a faster rate if the genetic status of the genes underpinning horn status is known. While the polled gene itself is elusive, markers in close association with the gene are identified. The current study proposes the use of a microsatellite marker for determining the genetic status at the polled locus even in *Bos indicus* cattle. Currently, around 1000 samples are being collected from the major Australian breeds with special focus on tropical breeds in a bid to validate this marker in the field. Efforts to develop markers for scurs and African horn are also underway. A Breeding program was conducted in 2007 at a collaborating property, Hillgrove Pastoral Company, to develop a Brahman resource population derived from mating 15 polled bulls with around 400 polled, horned and scurred heifers. Currently, around 230 calves are being regularly mustered for determining their accurate horn status phenotypes. In the coming months, this population will become the resource for the forthcoming whole genome association study. This research will enable us in identifying markers for scurs locus and in determining the presence or absence of African horn gene.

ACKNOWLEDGEMENTS

This project is supported by Meat and Livestock Australia (KCP). The Authors would like to thank Paul Williams, Geoffry Fordyce and Brian Burns for their help in conducting the breeding program and data collection. We wish to acknowledge in-kind contributions of Tom Mann, Hillgrove Pastoral Company, in generating the resource population.

REFERENCES

Asai, M., Berryere, T.G. and Schmutz, S.M. (2004) *Anim. Genet.* **35:**34.

Brenneman, R.A., Davis, S.K., Sanders, J.O., Burns, B.M., Wheeler, T.C., Turner, J.W. and Taylor, J.F. (1996) *J. Hered*. **87:**156.

Drögemüller, C., Wöhlke, A., Mömke, S. and Distl, O. (2005) *Mamm. Genome* **16:**613.

Georges, M., Drinkwater, R., King, T., Mishra, A., Moore, S.S., Nielsen, D., Sargeant, L.S., Sorensen, A., Steele, M.R., Zhao, X., Womack, J.E. and Hetzel, D.J.S. (1993) *Nat. Genet.* **4:**206.

Prayaga K.C. (2007) *Aust. J. Agric. Res.* **58:**1.

Stephens, M., Smith, N.J. and Donnelly, P. (2001) *Am. J. Hum. Genet.* **68**: 978.Li, N. and Stephens, M. (2003) *Genet.* **165**:2213.

THE GENETICS OF TEMPERAMENT TRAITS IN MERINO SHEEP

K. L. Lennon[1], M. L. Hebart[2], F. D. Brien[2] and P. I. Hynd[1]

[1] School of Animal and Veterinary Sciences, The University of Adelaide, Roseworthy SA 5371
[2] South Australian Research and Development Institute, Roseworthy SA 5371

SUMMARY

Investigations were made into the genetics of temperament in Merino ewes, with emphasis on those aspects which might have associations with maternal behaviour and postnatal survival of lambs. A data set of over 2000 animals and more than 20,000 records was analysed for estimation of genetic parameters. The heritability of ewe mothering temperament was 0.39 ± 0.02, indicating a moderate genetic component to this behavioural trait. Agitation score and flight time were less heritable (0.20 ± 0.05 and 0.12 ± 0.05 respectively). The heritability of litter survival was low (0.09 ± 0.01) and the genetic correlations between this and ewe mothering temperament, agitation score and flight time were 0.18 ± 0.08, 0.39 ± 0.18 and 0.09 ± 0.27 respectively. Estimated genetic correlations between temperament traits and wool traits were low, with the exception of staple length, which was negatively genetically correlated to agitation score (r_g = -0.26 ± 0.03). These results suggest that if temperament is used as a selection criterion, although no antagonistic results will be seen in wool production, there would be no advantage in litter survival compared with undertaking direct selection for the trait. Further, if selection is practised for low agitation score, our results suggest that litter survival may be slightly reduced in future generations.

INTRODUCTION

Lamb survival within Australia averages approximately 80% (Kilgour 1992) suggesting there is considerable opportunity for improvement. Most post partum lamb loss occurs within the first three days of life and is largely caused by starvation, mismothering and exposure (Nowak and Poindron 2006). Previous studies have highlighted that even well managed flocks may not exceed survival rates of 85% (Brand *et al.* 1985). Similar outcomes were identified within the Lifetime Wool Project (Lifetime Wool 2009), which showed when ideal management guidelines are followed twin survival still only averaged approximately 60%. These results suggest that tools in addition to management are required to maximise lamb survival rates.

Although breed and within breed differences in lamb survival have been identified in a number of studies, the heritability of lamb survival has been estimated as being only around 0.03 (Safari *et al.* 2005) suggesting genetic progress from direct selection will be slow. Use of indirect selection on a trait that has a higher heritability and is genetically correlated with lamb survival may provide a better option.

Maternal behaviour is an important determinant of lamb survival but difficult to measure commercially, however temperament could potentially be used as an indicator trait. Murphy (1999) found that in a flock of Merino ewes divergently selected for temperament using willingness to approach a human in order to re-join the flock (arena test) and agitation box scores, ewes from the 'calm' flock had a 10% higher lamb survival rate in twins compared with ewes from the 'nervous' flock.

The evidence indicates an important potential link between temperament and maternal behaviour in the immediate post natal period. Such behaviours have already been shown to significantly influence the formation of the ewe lamb bond, and subsequently lamb survival. However, before recommendations can be made to producers further investigation is required. This study aims to estimate the genetic and phenotypic parameters of temperament and correlations between this trait and important wool production traits.

MATERIALS AND METHODS

The genetic study utilised records from the Selection Demonstration Flocks (SDFs) which have been selected for wool and meat traits and had undergone no previous direct selection for temperament. For a detailed description of the flocks see (Ponzoni *et al.* 1999). Table 1 describes the traits recorded in the SDFs used to generate data sets to calculate genetic parameters for temperament and wool production traits. Ewe mothering temperament is recorded on a subjective five point scale using the distance the ewe moves from its lamb during tagging, agitation score is an objective measure of movement and vocalisations the sheep makes whilst isolated from flock mates and flight time measures the time taken for the sheep to travel 1.7 m after being released from a weight crate.

Table 1. Definition of traits for which genetic parameters were calculated, and the approximate number of animals and records used in each calculation

Trait	Abbreviation	No. of Animals	No. of Records
Litter survival from birth to weaning	LIS	3500	9000
Ewe mothering temperament	EMT	3700	5700
Agitation Score	AGIT	2000	2000
Flight Time	FT	2000	2000
Greasy Fleece Weight	GFW	6000	23500
Clean Fleece Weight	CFW	6000	23500
Yield	YLD	6000	23500
Fibre Diameter	FD	6000	23500
Standard Deviation of Fibre Diameter	SDFD	6000	23500
Coefficient of Variation of Fibre Diameter	CVFD	6000	23500
Curvature	CURV	6000	23500
Staple Strength	SS	6000	23500
Staple Length	SL	6000	23500

A pedigree file containing the sire, dam and paternal and maternal grandparents of each animal was used to form a relationship matrix. Variance and covariance components were estimated using a bivariate animal model in ASREML. The model included the fixed effects of age (2…7), type of birth and rearing of ewe (11, 21, 22), flock (1…5), drop (1990…2005) and the lamb type of birth and rearing (11, 21, 22) was fitted for ewe mothering temperament only. Any significant interactions were also fitted. Litter survival was analysed as a trait of the dam.

RESULTS

Ewe mothering temperament was moderately heritable ($h^2 = 0.35 \pm 0.02$; Table 2) and exhibited a positive, genetic correlation with both agitation score ($r_g = 0.26 \pm 0.12$) and litter survival ($r_g = 0.18 \pm 0.08$) implying that ewes with higher mothering temperament would have progeny with higher agitation scores and increased litter survival. Agitation score was also moderately heritable ($h^2 = 0.20 \pm 0.05$) and was negatively correlated with flight time ($r_g = -0.26 \pm 0.23$), but positively with litter survival ($r_g = 0.39 \pm 0.18$; Table 2), thus the more agitated the animal, the quicker the flight time and the higher the litter survival in its progeny. The heritability of flight time and litter survival were low ($h^2 = 0.12 \pm 0.05$ and 0.09 ± 0.01 respectively; Table 2) suggesting a low genetic component to these traits. All phenotypic correlations were small with the exception being a positive relationship ($r_p = 0.17 \pm 0.02$) between ewe mothering temperament and litter survival (Table 2).

Table 2. Heritability ± standard error, phenotypic and genetic correlations of ewe mothering temperament (EMT), agitation score (AGT), flight time (FT) and litter survival (LIS)*

	EMT	AGT	FT	LIS
EMT	**0.35 ± 0.02**	0.01 ± 0.02	0.01 ± 0.02	0.17 ± 0.02
AGT	0.26 ± 0.12	**0.20 ± 0.05**	0.03 ± 0.02	0.05 ± 0.02
FT	-0.08 ± 0.15	-0.26 ± 0.23	**0.12 ± 0.05**	0.01 ± 0.06
LIS	0.18 ± 0.08	0.39 ± 0.18	0.09 ± 0.27	**0.09 ± 0.01**

*Phenotypic correlations are presented above the diagonal, genetic correlations below the diagonal and heritabilities in bold on the diagonal

Most genetic and phenotypic correlations between temperament traits and wool production traits were not significantly different to zero (P>0.05, Table 3) with the exception of a negative genetic correlation between agitation score and staple length (r_g = -0.26 ± 0.03, P < 0.05). Thus, the less agitated the ewe the longer the staple length in its progeny.

Table 3. Genetic and phenotypic correlations between ewe mothering temperament (EMT), agitation score (AGT) and flight time (FT) and important wool production traits

	EMT		AGT		FT	
	Phenotypic	Genetic	Phenotypic	Genetic	Phenotypic	Genetic
GFW	-0.01± 0.04	0.01± 0.02	0.01± 0.11	0.01 ± 0.03	0.02± 0.12	0.16± 0.03
CFW	-0.01± 0.04	0.01± 0.02	0.01± 0.11	0.03 ± 0.03	0.02± 0.12	0.16± 0.03
YLD	-0.01± 0.04	0.01± 0.02	-0.03 ± 0.01	0.01 ± 0.02	-0.01± 0.12	-0.03± 0.03
FD	-0.01± 0.04	0.03± 0.02	0.06 ± 0.01	0.16 ± 0.03	-0.01± 0.12	-0.04± 0.02
CVFD	-0.01± 0.04	-0.01± 0.02	-0.06 ± 0.01	-0.05 ± 0.03	0.02± 0.12	0.18± 0.02
SS	0.03 ± 0.04	-0.03± 0.02	0.13 ± 0.11	0.1± 0.03	-0.01± 0.13	-0.03± 0.02
SL	-0.01± 0.04	0.01± 0.02	-0.12 ± 0.01	-0.26± 0.03	-0.01± 0.12	0.08± 0.02

DISCUSSION

Overall, temperament heritability estimates in this study were low to moderate, consistent with other studies. In our study, the heritability estimated for EMT of 0.35 was higher than the estimate of 0.09 found by Everett-Hincks *et al*. (2005); several possible causes for this disparity are suggested. Firstly, there was a difference in the timing of EMT scores with those in our study allocated to ewes within 12 hours of birth, whilst the flock analysed by Everett-Hincks *et al.* (2005) recorded EMT 12-36 hours after birth. Secondly, there were differences in prior selection for maternal ability, with no history of such selection in the flock we investigated, whereas the flock studied by Everett-Hincks *et al.* (2005) had undergone culling for poor rearing ability which may have reduced genetic variation. Thirdly, the varying estimates may reflect actual differences in available genetic variation between the Merino flock we investigated and the Coopworth flock investigated by Everett-Hincks *et al.* (2005). The current study estimated a lower heritability of agitation score than found in others (h^2 = 0.41; Blache and Ferguson 2005) however standard errors were high in previous estimates due to limited number of records. In contrast to studies in cattle but in agreement with those conducted in sheep, flight time heritability estimates were low, which coupled with its low genetic correlation with litter survival and low repeatability (Blache and Ferguson 2005) suggests that selection for this temperament trait will result in little genetic gain, both for the trait itself and for lamb survival.

The correlation between agitation score and litter survival was positive, thus the more agitated the ewe (or nervous in temperament) the higher the litter survival. This was unexpected as Murphy

(1999) reported that that ewes selected for calm temperaments had increased lamb survival. The disparity between studies may partly be explained by the fact that the divergent flocks used in Murphy (1999) were selected not only on agitation score but also on arena test results in addition to the smaller number of sires and lambs used in the study. The result of more agitated ewes displaying higher litter survival is concerning as selection for temperament has been suggested as a means of improving meat quality (Voisinet *et al*, 1997), an event which would have detrimental consequences for lamb survival.

The lack of correlations suggest that temperament can be selected for without impacting wool production. In fact, the correlation between agitation score and staple length implies that calmer ewes will tend to have progeny with increased staple length. A similar result in beef cattle has been witnessed whereby less docile animals are less productive (Gauly *et al*, 2001). Additionally, increased corticosterone levels have been shown to retard growth in broiler chickens (Post *et al.* 2003). The results in this study coupled with those obtained in other species suggest that animals that differ in their temperament also differ in physiological characteristics.

Litter survival heritability was low, which is consistent with other studies (Everett-Hincks *et al*, 2005, Fogarty *et al.* 1994) with heritability estimates of lamb survival being similar (Safari *et al.* 2005). These estimates indicate only slow genetic progress when direct selection is employed, explaining why indirect methods of increasing lamb survival were investigated in this study. Although the overall low to moderate heritability of temperament suggests genetic gain in the trait itself can be made through selection, for indirect selection to be more efficient than direct selection, the product of heritability of temperament and the genetic correlation needs to be greater than the heritability for the direct trait (litter survival). This is not the case for any measures of temperament analysed in this study, and coupled with the antagonistic correlation between agitation score and litter survival, suggests that other methods of improving lamb survival should be explored.

ACKNOWLEDGEMENTS

Data collection was conducted by the South Australian Research and Development Institute at Turretfield Research Centre. This research was supported by the University of Adelaide, the South Australian Research and Development Institute and RSPCA Australia.

REFERENCES

Blache, D. and Ferguson, D. (2005) *Sheep Updates*, Department of Agriculture Western Australia

Brand, A. A., Cloete, S. W. P. and De Villiers, T. T. (1985) *S. Afr. J. Anim. Sci.* **15**:555.

Everett-Hincks, J. M., Lopez-Villalobos, N., Blair, H. T. and Stafford, K. J. (2005) *Livest. Prod. Sc*. **93**:51.

Fogarty, N. M., Brash, L. D. and Gilmour, A. R. (1994) *Aust. J. Agric. Res*. **45**:443.

Gauly, M., Mathiak, H., Hoffmann, K., Kraus, M. and Erhardt, G. (2001) *Appl. Anim. Behav. Sci.* **74**:109.

Kilgour, R. J. (1992) *Aust. J. Exp. Agric.* **32**:311.

Lifetime Wool (2009) Accessed 2nd March, 2009 < http://www.lifetimewool.com.au/index.aspx>

Murphy, P. M. (1999) PhD Thesis, The University of Western Australia

Nowak, R. and Poindron, P. (2006) *Reprod. Nutr. Dev.* **46**:431.

Ponzoni, R. W., Jaensch K. S., Grimson, R. J., Smith, D. H., Ewers, A. L. and Ingham, V. H. (1999) *Wool Tech. Sheep Bree*. **47**:83.

Post, J., Rebel, J. M. and ter Huurne, A. A. (2003) Poultry Sc. **82**:1313.

Safari, A., Fogarty, N. M. and Gilmour, A. R. (2005) *Livest. Prod. Sc*. **92**:271.

Voisinet, B. D., Grandin, T., O'Connor, S. F., Tatum, J. D. and Deesing, M. J. (1997) *Meat. Sc.* **46**:367.

THE ROLE OF CYTOCHROME P450 17α-HYDROXYLASE/17,20-LYASE (CYP17) IN THE STRESS COPING ABILITY OF A DIVERGENTLY SELECTED MERINO SHEEP POPULATION

D. van der Walt[1], S.W.P. Cloete[2,3], K. Storbeck[1] and P. Swart[1]

[1]Department of Biochemistry, University of Stellenbosch, Stellenbosch 7602, South Africa
[2]Department of Animal Sciences, University of Stellenbosch, Stellenbosch 7602, South Africa
[3]Institute for Animal Production: Elsenburg, Private Bag X1, Elsenburg 7607, South Africa

SUMMARY

South African Merino sheep were selected divergently from the same base population for their ability to rear multiples. Two distinct populations were formed over a period of more than 20 years of selection. Reproduction (and therefore presumably fitness) in the line selected in the upward direction (H-line) was substantially improved compared to the line selected in the downward direction (L-line). In the present study, it was demonstrated that the H-line was more stress-tolerant than the L-line in terms of their glucose and cortisol response when challenged with insulin. Sheep from the breeding program were genotyped according to one of two cytochrome P450 17α-hydroxylase/17-20 lyase (CYP17) alleles, as these genotypes were previously linked to the ability of Angora goats to cope with external stressors. However, no association was found between CYP17 genotype and selection line. The difference in insulin induced stress response between the H- and the L-line can therefore not be attributed to CYP17 genotype.

INTRODUCTION

Fitness of farm animals (defined as reproduction and survival) has long been identified as being of economic importance. Yet fitness traits have seldom been incorporated in selection programs for livestock (Goddard 2009). These fitness traits can be linked to genotypic markers that can be identified within livestock breeding programmes. These genetic targets can then be recorded and included in selection criteria to ultimately improve livestock fitness. In this study we look at cytochrome P450 17α-hydroxylase/17-20 lyase (CYP17) genotype as a possible genetic target to link to stress coping ability, a fitness trait.

CYP17 plays a critical role in the production of mineralocorticoids, glucocorticoids and androgens by the adrenal cortex in mammals (Vander *et al.* 2001). These steroid hormones are involved in fitness, since they play a vital role in the control of water and mineral balance, stress management and reproduction, respectively. CYP17 catalyses two distinct reactions, namely: a 17α-hydroxylation and a C17-C20 lyase reaction (Nakajin *et al.* 1981). This dual enzymatic activity places CYP17 at key branch points in the biosynthesis of adrenal steroid hormones.

Cortisol and corticosterone are the glucocorticoid hormones produced in the adrenal gland, which play an important role in stress management (Vander *et al.* 2001). As in humans, cortisol is by far the main glucocorticoid in sheep that counters a stress stimulus. Cortisol production is stimulated when the adrenal gland receives a "stress-signal" through the hypothalamus-pituitary-adrenal axis, via adrenocorticotropic hormone. The decreased ability of an animal to produce cortisol will lead to a reduced ability to counteract stress associated with the environment. Such an example was observed by Engelbrecht and Swart (2000) in Angora goats. These animals had a decreased ability to produce cortisol compared to Merino sheep and Boer goats, and accordingly exhibited a reduced ability to cope with insulin-induced stress.

Stress has been shown to reduce fitness, as reflected by growth, reproduction and survival of farm animals. Divergent selection for number of lambs weaned in Merino sheep, an example of a composite fitness trait, has resulted in marked differences in responses between the lines in this

trait (Cloete *et al*. 2004). Differences between the lines in survival of lambs and behavioural adaptations conducive to lamb survival were also observed (Cloete and Scholtz 1998).

Two CYP17 alleles have previously been identified in Merino sheep (Genbank accession no. L40335/WT1 and AF251388/WT2) and confirmed by Storbeck *et al.* (2008). However, it has not been established in this species whether a specific CYP17 genotype would enhance cortisol production, and thus stress coping ability, relative to the other. In this study, we investigated whether the observed divergence in fitness (as reflected by number of lambs weaned) observed in a Merino selection experiment can be related to the genotypic composition of ovine CYP17.

MATERIALS AND METHODS

Breeding program. A Merino sheep breeding program has been undertaken since 1986 in which sheep have been divergently selected for their ability to rear multiples (alternatively defined as number of lambs weaned per mating). The selection lines were derived from the same base population and selection within each line based on maternal ranking values for number of lambs weaned per lambing opportunity (Cloete *et al*. 2004). Number of lambs weaned per mating in the line selected in the upward direction (H-line) has been proved to be near to double that of the line selected in the downward direction (L-line) (Cloete *et al.* 2004).

Stress test. Stress coping ability was tested on 24 rams from this breeding program (13 H-line and 11 L-line sheep), housed at the Elsenburg Research farm near Stellenbosch, South Africa. These rams were injected intravenously with human insulin (Actrapid® HM, Novo Nordisk, Johannesburg, South Africa) after which 6 blood samples of each animal were collected over a 2 hour period and placed on ice. Blood samples were centrifuged at 2 500x*g* for 10 minutes (4^{o}C) to acquire representative plasma samples from each animal. Plasma glucose and cortisol levels were determined by PathCare Reference Laboratory (PathCare Park, N1 City, Goodwood, Cape Town, South Africa). Ethics approval for this stress test was obtained from the Departmental Ethics Committee for Research on Animals (DECRA reference R08/21).

Genomic DNA isolation. Blood samples of both H- (n=105) and L-line (n=31) sheep were collected in EDTA treated collection tubes (BD Vacutainer® Blood Collection Tubes; Pronto™ Quick Release Holder and Eclipse™ Blood Collection Needles). Blood samples were also acquired from the heart chamber of 36 lambs that had died during the 2008 lambing season. Genomic DNA was isolated using the Wizard® Genomic DNA isolation kit (Promega, Madison, Wisconsin) according to the instructions provided by the manufacturer.

CYP17 genotyping with real time polymerase chain reaction (RT-PCR). All 172 sheep were genotyped using the RT-PCR method developed by Storbeck *et al*. (2008). The primers and hybridisation probes (TibMolBio, Berlin, Germany) were as follows: LCLP, 5'-CCTGAAGGCCATACAAA-3'; LCRP, 5'-GGATACTGTCAGGGTGTG-3'; fluorescein-labelled CYP17 sensor probe, 5'-TTCTGAGCAAGGAAATTCTGTTAGA-FL; LC640-labelled CYP17 anchor probe, 640-TATTCCCTGCGCTGAAGGTGAGGA-3'. RT-PCR was carried out using a LightCycler® 1.5 instrument. Amplification reactions (20µl) contained 2 µl LightCycler® FastStart DNA Master HybProbe Master Mix (Roche Applied Science, Mannheim, Germany), 3 mM $MgCl_2$, 0.5 µM of each CYP17 primer, 0.2 µM fluorescein-labelled CYP17 sensor probe, 0.2 µM LC640-labelled CYP17 anchor probe and 10 to 100 ng genomic DNA. Following an initial denaturation at 95°C for 10 min to activate the FastStart *Taq* DNA polymerase, the 35-cycle amplification profile consisted of heating to 95°C with a 8 s hold, cooling to 52°C with a 8 s hold and heating to 72°C with a 10 s hold. The transition rate between all steps was 20°C/second. Data

were acquired in single mode during the 52°C phase using the LightCycler® software (version 3.5). Following amplification, melting-curve analysis was performed as follows: denaturation at 95°C with a 20 s hold, cooling to 40°C with a 20 s hold and heating at 0.2°C/s to 85°C with continuous data acquisition.

The sensor probe was designed to be a perfect match for the WT1 sequence and dissociated at 58°C when bound to the perfectly matched WT1 sequence. However, when bound to the mismatched sequence (WT2) dissociation occurred at 54°C. A no-template control (negative control) was also included in each assay.

Statistical analysis. For the stress test, plasma glucose and cortisol response over time for the H- and L-lines were analyzed with a regular two-way ANOVA with selection line and the period that passed since the insulin injection as factors. Differences between breed lines at specific time points were examined with Bonferonni's post-test (glucose: 95% confidence; cortisol: 94% confidence). The Chi-square test was used to analyze CYP17 genotype frequencies. GraphPad Prism (version 4) software (GraphPad Software, San Diego, California) was used for all statistical analysis.

RESULTS AND DISCUSSION

Stress test. Plasma glucose (mmol/L; which served to monitor the progress of the stress response) and plasma cortisol (log nmol/L) responses to insulin-induced stress are depicted in Figure 1. The H-line reached a hypoglycaemic state earlier than the L-line, with glucose levels of 1.9 mmol/L 30 min post insulin challenge, and recovered to baseline glucose (3.7 mmol/L) 2 hours post insulin challenge (3.3 mmol/L). Cortisol levels increased rapidly from 30 min post insulin challenge (60.3 log nmol/L), reached maximum at 60 min (120.2 log nmol/L) and returned to baseline concentrations after 2 hours. The stress response of H-line animals was completed after 2 hours with both glucose and cortisol concentrations having recovered to baseline levels.

The L-line reached maximum hypoglycaemic state at 60 min post insulin challenge with glucose levels of 2.1 mmol/L, but did not recover to baseline concentrations (3.6 mmol/L) by 2 hours post insulin challenge (2.7 mmol/L). L-line cortisol levels increased from 30 min post insulin challenge (53.3 log nmol/L), but at an apparently slower rate than the H-line. Maximum cortisol in the L-line was observed 90 min post insulin challenge (100.5 log nmol/L).

The interaction between breed line and time of measurement was highly significant ($P<0.0001$) for glucose ($F=9.22$, dfn=5, dfd=132), but not for cortisol ($P=0.3028$, $F=1.22$, dfn=5, dfd=132) responses to insulin-induced stress.

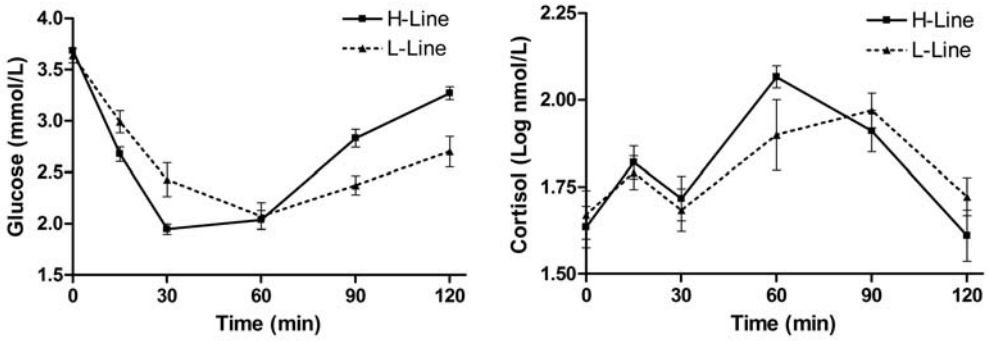


Figure 1. Merino sheep plasma glucose and cortisol response to insulin challenge.

Glucose levels were lower in the H-line at 30 minutes post insulin challenge (P<0.01), the opposite trend being observed after 90 (P<0.01) and 120 min (P<0.001). Cortisol levels were higher (P=0.0590) in the H-line 60 min post insulin challenge. The H-line animals thus had an improved ability to cope with insulin induced hypoglycaemia than the L-line, as reflected by their quicker glucose recovery to baseline and earlier peaking of cortisol at 60 min. An improved cortisol collection and detection method might limit variation in future tests.

CYP17 genotyping. Interestingly, no homozygous WT2 sheep were detected in either the H- or L-lines or among the lambs that died in 2008. Relative DNA copy number determination of sheep CYP17 has previously been done (Storbeck *et al.* 2008), indicating that the two CYP17 genetic sequences are two alleles of one gene. This finding thus warrants further investigation.

Table 1 summarizes the genotyping results obtained for the H-, L-line and lamb mortalities. There was no significant association (P=0.7617, Chi-square=0.5444, df=2) between CYP17 genotypes and designation of sample population (H-, L-line or lamb mortalities). On average 83.4 % sheep in the breeding program was heterozygous, while 16.6 % were homozygous WT1.

Table 1. Frequency distribution of CYP17 genotype in the Merino sheep breeding program

Merino flock	Homozygous WT1		Heterozygous WT1/WT2		Homozygous WT2	
	Number of sheep	Percentage	Number of sheep	Percentage	Number of sheep	Percentage
H-line	15	14.3	90	85.7	0	0
L-line	5	16.1	26	83.9	0	0
Lamb mortalities	7	19.4	29	80.6	0	0

CONCLUSIONS

The divergent breeding program was shown to result in differences in insulin-induced stress coping ability, the H-line having a higher stress tolerance than the L-line. This difference in stress tolerance could not, however, be ascribed to CYP17 genotype, since there was no association between CYP17 genotype and selection lines. One CYP17 isoform is not more advantageous for cortisol production in the adrenal gland than the other. This study rules out CYP17 as possible genotypic marker to use during selection, and suggest investigating other factors along the HPA axis or adrenal steroidogenesis that could be implicated in the stress response difference observed.

ACKNOWLEDGMENTS

This research was supported by South African Wool Industry, South African Meat Industry and the National Research Foundation.

REFERENCES

Cloete, S.W.P., Gilmour, A.R., Olivier, J.J. and Van Wyk, J.B. (2004) *Aust. J. Exp. Agric.* **44**:745.
Cloete, S.W.P. and Scholtz, A.J. (1998) *Aust. J. Exp. Agric.* **38**:801.
Engelbrecht, Y. and Swart, P. (2000) *J. Anim. Sci.* **78**:1036.
Goddard, M. (2009) "Adaptation and Fitness in Animal Populations – Evolutionary and Breeding Perspectives on Genetic Resource Management", Springer Science and Business Media, www.springer.com.
Nakajin, S., Shively, J.E., Yuan, P. and Hall, P.F. (1981) *Biochemistry.* **20:**4037.
Storbeck, K., Swart, A.C., Snyman, M.A. and Swart, P. (2008) *FEBS J.* **275**:3934.
Vander, A., Sherman, J. and Luciano, D. (2001) "Human physiology: the mechanisms of body function", 8th ed., McGraw-Hill Companies, Inc., New York, NY.

THE IMPROVEMENT OF LAMB SURVIVAL OF MERINO SHEEP AS A CORRELATED RESPONSE TO DIRECT SELECTION FOR REARING ABILITY

S.W.P. Cloete[1,2], I. Misztal[3] and J.J. Olivier[4]

[1]Department of Animal Sciences, University of Stellenbosch, Stellenbosch 7602, South Africa
[2]Institute for Animal Production: Elsenburg, Private Bag X1, Elsenburg 7607, South Africa
[3]Department of Animal and Dairy Science, University of Georgia, Athens, GA 30605, USA
[4]ARC: Livestock Business Division, Private Bag X5013, Stellenbosch 7599, South Africa

SUMMARY

Data were available for peri-natal lamb survival (PNS), lamb survival from 4 days to weaning (S4DTW), birth weight (BW) and birth coat score (BCS) in South African Merino sheep that were selected divergently from the same base population for their ability to rear multiples. These data spanned the period 1986-2008, and were subjected to genetic analyses involving two three-trait linear-threshold analyses, entailing PNS and S4DTW analysed along with the other traits. The lamb survival trait considered was treated as a categorical trait in both instances. The line selected in the upward direction was denoted the high line (H line) and the line selected downward as the low line (L line). Heritability was estimated at 0.15 for PNS, 0.20 for S4DTW, 0.16 for BW and 0.60 for BCS. Corresponding estimates for maternal heritability amounted to 0.16, 0.10, 0.25 and 0.06 and to 0.11, 0.05, 0.08 and 0.03 for the dam permanent environment. Expressed relative to the overall mean, direct and maternal breeding values for PNS improved at respectively 0.61% p.a. and 0.46% p.a. in the H line. Corresponding genetic trends in the L line amounted to respectively 0.26% p.a. and -0.96% p.a. Direct and maternal genetic trends for S4DTW amounted to respectively 0.70% p.a. and 0.16% p.a. in the H line and to respectively -0.26% p.a. and -0.26% p.a. in the L line. It was concluded that genetic change in lamb survival would accrue when selection is based on a trait like maternal multiple rearing ability.

INTRODUCTION

Ovine lamb survival has long been considered a trait of economic importance (Haughey 1991). Apart from the obvious loss of monetary income incurred by lamb deaths, there is also an important animal welfare component associated with it. Various analysts considered the prospect of improving lamb survival genetically as being relatively poor (Morris *et al*. 2000; Everett-Hincks *et al*. 2005). The binomial distribution of survival traits and generally low levels of genetic variation in these traits have contributed to this assessment. However, recent analyses involving threshold models have reported noteworthy genetic variation in survival (Welsh *et al*. 2006; Riggio *et al*. 2008), while divergence were reported in age-specific lamb survival (survival of birth, survival from birth to docking and survival from docking to weaning) of Merino sheep selected divergently for ewe multiple rearing ability (Cloete *et al*. 2009).

Peri-natal lamb survival is defined as survival of the peri-parturient period, up to an age of three days (Haughey 1991). Survival of this phase is often seen as critical to ensure good survival to weaning. This study therefore reports genetic variation in peri-natal lamb survival, as well as lamb survival from four days to weaning. These traits were analysed along with lamb birth weight and birth coat score, enabling the estimation of (co)variance components for all traits.

MATERIALS AND METHODS

The resource population involved Merino sheep that were divergently selected for their ability to rear multiples since 1986. Selection was based on maternal ranking values for number of lambs weaned per parity (Cloete *et al*. 2004; 2009). The origin, environment, management and selection

practices in the resource population are described in the literature cited, and will not be elaborated further.

Age-specific lamb survival of 5,686 lambs born from 1986 to 2008 was available in the resource population, as described by Cloete *et al*. (2009). These data were used to construct individual records for peri-natal survival (PNS – defined as the survival prior to, during and up to 3 days after birth), and survival from 4 days of age up to weaning (S4DTW). Birth weight (BW) was recorded within 24 hours of birth in all these lambs, with the exception of a small number that were mutilated by vermin. Birth coat score (BCS) was recorded for lambs that were born from 1995. Scores ranged from 1 (hairy) to 5 (woolly), as detailed by Cloete et al. (2003).

Two three-trait animal models were fitted with either PNS or S4DTW together with BW and BCS. Lamb survival was defined as a binary trait with two categories (1 for lambs dying in the particular category and 2 for those surviving), whereas BW and BCS data were treated as continuous. The fixed effects included year of birth (1986 – 2008), sex (male and female), dam age (2 to 7+ years) and birth type (single and multiple). Animal, maternal genetic and dam permanent environment terms were included as random effects. The genetic correlation between direct and maternal genetic effects was included initially. It was not significant for either trait, and was excluded from the final runs. The software used was THRGIBBSF90 (Misztal *et al*. 2002). This software is suitable for the estimation of variance components and genetic parameters in threshold animal mixed models for any combination of categorical and continuous traits. The programme POSTGIBBSF90 was used for Post Gibbs analysis (Misztal *et al*. 2002). In all cases, a single chain of 200,000 cycles were run, with the first 40,000 cycles used as the burn-in period.

RESULTS AND DISCUSSION

Data description: The proportion of lamb surviving amounted to 0.88 for PNS and 0.86 for S4DTW (1 is subtracted from the mean value in Table 1 to account for the coding). Sawalha *et al*. (2007) reported a proportion of 0.95 lambs surviving the period during or shortly after birth. Riggio *et al*. (2008) accordingly found a proportion of 0.92 lambs surviving birth. Postnatal survival was high in the study of Sawalha *et al*. (2007), namely a proportion of 0.98 surviving from 1 to 14 days and 0.96 surviving from 15 days to 120 days of age. Cumulative survival in the study of Riggio *et al*. (2008) amounted to 0.87 after four weeks and to 0.85 after 12 weeks. Survival was thus generally somewhat poorer in the present study.

BW and BCS were normally distributed and averaged respectively 3.8 kg and a score of 3.2 (Table 1). The overall mean for BW was within the range of 3.6 to 4.9 kg reported for seven Australian Merino resource flocks (Safari *et al*. 2007a). At 3.85 kg for BW and 3.22 for BCS, the previous means reported for a smaller data set of the same resource population (Cloete *et al*. 2003) also accorded with the present study. The variability of the BW data used in the present study compared well with coefficients of variation ranging from 17 to 22% in the study of Safari *et al*. (2007a).

Table 1. Descriptive statistics for the raw data analysed

Trait	Number of observations	Mean ± s.d.	Skewness	Kurtosis	Range
PNS	5,686	1.88 ± 0.32	n.a	n.a.	1 – 2
S4DTW	5,024	1.86 ± 0.35	n.a.	n.a.	1 – 2
BW	5,665	3.79 ± 0.86	0.02	0.10	1 – 7
BCS	3,521	3.19 ± 0.92	-0.95	1.68	1 – 5

n.a. – as means for survival are incidence dependent, statistical information is not supplied

Genetic parameters: Lamb survival on the underlying scale was moderately heritable, at 0.15 (PNS) and 0.20 (S4DTW) (Table 2). Corresponding estimates for the maternal genetic variance ratio (m^2) were respectively 0.16 and 0.10. Sawalha *et al*. (2007) analysed lamb viability (defined as survival at birth and up to 24 hours and thus corresponding with PNS), and obtained estimates of 0.05 for h^2, 0.10 for m^2 with a litter effect of 0.19. Riggio *et al*. (2008) estimated h^2 of survival at birth and up to 24 hours at 0.33, using a sire model and the probit link function. Literature estimates comparable to S4DTW were 0.13 for h^2 and 0.14 for m^2 with 0.25 for the litter effect for lamb survival from 1 to 14 days of age in the study of Sawalha *et al*. (2006). These estimates generally corresponded with the present study.

(Co)variance ratios for BW and BCS were within 0.01 when jointly estimated with either PNS or S4DTW (Table 2). Estimates for BW were 0.16 for h^2, 0.25 for m^2 and 0.08 for c^2. Genetic parameters for BW are relatively scarce for Merino lambs. Previous estimates of h^2 ranged from 0.12 to 0.19 (Duguma *et al*. 2002; Cloete *et al*. 2003; Safari *et al*. 2007b). Corresponding estimates for m^2 ranged from 0.19 to 0.25, and those for the dam permanent environment (c^2) from 0.07 to 0.10. Variance ration estimates for BCS were 0.60 for h^2, 0.06 for m^2 and 0.03 for c^2. An earlier study on a smaller data set of the same resource population yielded estimates of 0.70 for h^2 and 0.04 for c^2 (Cloete *et al*. 2003). Part of the maternal variation was partitioned towards m^2 when using the larger database of the present study. Kemper *et al*. (2003) accordingly reported the h^2 of BCS to be 0.65, while Ponzoni *et al*. (1996) reported an estimate of 0.66.

Table 2. Phenotypic variances (σ^2_P) and genetic, maternal and dam permanent environmental (co)variance ratios (± s.e.) for PNS and S4DTW with BW and BCS

Trait / Trait	PNS			S4DTW		
	Survival	BW	BCS	Survival	BW	BCS
σ^2_P	1.708	0.508	0.683	1.525	0.506	0.684
Genetic correlations (h^2 in bold on diagonal)						
Survival	**0.15±0.07**	-0.28±0.19	0.07±0.20	**0.20±0.08**	-0.09±0.19	-0.22±0.18
BW		**0.16±0.03**	0.06±0.11		**0.16±0.03**	0.05±0.11
BCS			**0.61±0.06**			**0.60±0.06**
Maternal genetic correlations (m^2 in bold on diagonal)						
Survival	**0.16±0.05**	0.27±0.17	0.24±0.23	**0.10±0.03**	0.22±0.17	0.22±0.21
BW		**0.25±0.04**	0.02±0.17		**0.25±0.04**	0.02±0.16
BCS			**0.06±0.02**			**0.06±0.02**
Dam PE correlations (c^2 in bold on diagonal)						
Survival	**0.11±0.04**	0.14±0.27	0.17±0.35	**0.05±0.02**	0.21±0.25	-0.82±0.28
BW		**0.08±0.03**	-0.72±0.25		**0.08±0.02**	-0.71±0.22
BCS			**0.03±0.01**			**0.03±0.01**

Ranging between -0.28 and 0.22, direct and maternal genetic correlations were consistently smaller than twice their standard errors (Table 2). Lamb viability to 1 day of age (0 for survivors and 1 for mortalities) was related to BW by Sawalha *et al*. (2007). The derived genetic correlation was unfavourable at 0.21. The sign and magnitude of this correlation were consistent with those reported in the present study (-0.26, remembering that the sign of the correlation has to be reversed because of different coding). No conclusive genetic relationship was found between S4DTW and BCS. Ponzoni *et al*. (1996) accordingly reported a limited impact of BCS on lamb survival. The corresponding dam permanent environmental correlation was high and significant at -0.82.

Genetic trends. Expressed relative to the overall phenotypic mean, the direct genetic trend for PNS amounted to 0.61% per annum in the H line and to 0.26% per annum in the L line.

Corresponding maternal genetic trends amounted to 0.46% and -0.96% per annum for the respective lines. Direct genetic trends for S4DTW amounted to 0.70% of the overall phenotypic mean per annum in the H line and to -0.26% per annum in the L line. Maternal genetic trends in the respective lines accordingly amounted to respectively 0.16% and -0.26 per annum. Previous research indicated divergence in genetic trends for survival of birth (maternal) and in survival from birth to docking (direct) (Cloete *et al*. 2009). A substantial direct improvement was observed in the survival of H line lambs in the period from docking to weaning while maternal breeding values in the L line declined during this period.

Table 3. Regressions (b ± s.e.) of individual predicted breeding values on year of birth for PNS and S4DTW of 3,548 H line lambs and 1,664 L line lambs. Regressions were forced through the origin in all instances.

Trait	PNS		S4DTW	
Line	H line	L line	H line	L line
Direct	0.0114±0.0005	0.0049±0.0011	0.0131±0.0006	-0.0049±0.0011
Maternal	0.0087±0.0006	-0.0281±0.0012	0.0031±0.0005	-0.0048±0.0009

CONCLUSIONS

Results from the present study suggest that genetic improvement in lamb survival is feasible when selection is based on a related trait like maternal rearing ability. Sheep production is seen to benefit from such selection, since ewes capable of rearing their lambs with minimal external efforts would be desirable from an economic as well as from an animal welfare perspective.

REFERENCES

Cloete, S.W.P., Gilmour, A.R., Olivier, J.J. and Van Wyk, J.B. (2004) *Aust. J. Exp. Agric.* **44**:745.

Cloete, S.W.P., Misztal, I. and Olivier, J.J. (2009) *J. Anim. Sci.* **87**:2196.

Cloete, S.W.P., Olivier, J.J., van Wyk, J. B., Erasmus, G.J. and Schoeman, S.J. (2003) *S. Afr. J. Anim. Sci.* **33**:248.

Duguma, G.J., Schoeman, S.J., Cloete, S.W.P. and Jordaan, G.F. (2002) *S. Afr. J. Anim. Sci.* **32**:66.

Everett-Hincks, J.M., Lopez-Villalobos, N., Blair, H.T. and Stafford. K.J. (2005) *Livest. Prod. Sci.* **93**:51.

Haughey, K.G. (1991) *J. S. Afr. Vet. Assoc.* **62**:78.

Kemper, K.E., Smith, J.L. and Purvis, I.W. (2003) *Proc. Assoc. Advmnt Anim. Breed. Genet.* **15**: 139.

Misztal, I., Tsuruta, S., Strabel, T., Auvray, B., Druet, T. and Lee, D. H. (2002) *Proc. 7th World Congr. Gen. Appl. Livest. Prod.* **33**:743. Montpellier, France.

Morris, C.A., Hickey, S.M. and Clarke, J.N. (2000) *N. Z. J. Agric. Res.* **43**:515.

Ponzoni, R.W., Grimson, R.J., Jaensch, K.S., Smith, D.H. and Hynd, P.I. (1996) SARDI Research Report Series, No 11. pp. 44.

Riggio, V., Finocchiaro, R. and Bishop, S.C. (2008) *J. Anim. Sci.* **86**:1758.

Safari, E., Fogarty, N.M., Gilmour, A.R., Atkins, K.D., Mortimer, S.I., Swan, A.A., Brien, F.D., Greeff, J.C. and Van der Werf, J.H.J. (2007a) *Aust. J. Agric. Res.* **58**:169.

Safari, E., Fogarty, N.M., Gilmour, A.R., Atkins, K.D., Mortimer, S.I., Swan, A.A., Brien, F.D., Greeff, J.C. and Van der Werf, J.H.J. (2007b) *Aust. J. Agric. Res.* **58**:177.

Sawalha, R. M., Conington, J., Brotherstone, S. and Villanueva, B. (2006) *Proc. 8th World Cong. Gen. Appl. Livest. Prod.*, Belo Horizonte, Brazil. CD communication 04-18.

Sawalha, R.M., Conington, J., Brotherstone, S. and Villanueva, B. (2007) *Animal* **1**:151.

Welsh, C.S., Garrick, D.J., Ens, R.M. and Nicoll, G.B. (2006) *N. Z. J. Agric. Res.* **49**: 411.

GENETICS OF LAMB SURVIVAL: PRELIMINARY STUDIES OF THE INFORMATION NUCLEUS FLOCK

F.D. Brien[1,2], M.L. Hebart[1,2], J.E. Hocking-Edwards[1,3], J.C. Greeff[1,4], K.W. Hart[1,5], G. Refshauge[1,6], G. Gaunt[1,7], R.Behrendt[1,8], K. Thomson[1,8,4], G.N. Hinch[1,9], K.G. Geenty[1,9] and J.H.J. van der Werf[1,9]

[1]Sheep CRC and [2]South Australian Research and Development Institute, Roseworthy, SA, 5371
[3]South Australian Research Institute, Struan Research Centre, Naracoorte, SA, 5271
[4]Department of Agriculture and Food Western Australia, South Perth, WA 6151
[5]Department of Agriculture and Food Western Australia, Narrogin, WA, 6312
[6]New South Wales Department of Primary Industries, Cowra, NSW, 2794
[7]Department of Primary Industries, Victoria, Rutherglen, VIC, 3685
[8]Department of Primary Industries, Victoria, Hamilton, VIC, 3300
[9]University of New England, Armidale, NSW 2351

SUMMARY

Poor lamb survival is recognised in Australia as a major contributor to reproductive inefficiency in the national flock. This paper provides preliminary estimates of phenotypic and genetic parameters on lamb survival and related traits for the 2007 and 2008 data from the CRC for Sheep Industry Innovation's Information Nucleus. The paper focuses on the potential utility of measurements on related traits for achieving genetic improvement in lamb survival. Although timed lamb behaviour traits show consistently high genetic correlations with lamb survival to 3 days of age, no obvious candidates have yet been identified for commercial use as indicator traits to genetically improve lamb survival.

INTRODUCTION

Poor lamb survival is a major contributor to reproductive inefficiency in sheep flocks in Australia (Alexander 1984), with survival rarely exceeding 90% and 80% in singles and twins, respectively (Hinch *pers.com.*). This paper provides preliminary estimates of phenotypic and genetic parameters on lamb survival and related traits from the CRC for Sheep Industry Innovation's Information Nucleus (IN) (Fogarty *et al.* 2007). The paper focuses on the potential utility of measurements on related traits for achieving genetic improvement in lamb survival.

MATERIALS AND METHODS

Information Nucleus. The data is from records of the 2007 and 2008 lambings of the IN, which operates as a series of linked flocks at research sites in widely differing environments around Australia, genetically testing key young industry sires for an extensive range of traits. The IN is directly linked to breeders and industry through the Sheep Genetics database (Fogarty *et al.* 2007). The 8 sites of the IN and the approximate number of ewes mated at sites each year are Armidale (1000 ewes), Trangie and Cowra (500 ewes each) in NSW, Rutherglen and Hamilton (500 ewes each) in Victoria, Struan and Turretfield (500 ewes each) in SA and Katanning (1000 ewes) in WA. Of the total ewe numbers, 80% are Merinos and 20% are Border Leicester-Merinos. All sites except Trangie had their first mating in 2007. Except for 2007 (when 65 sires were mated), approximately 100 sires are mated each year by AI, with most sires represented across all sites. Matings in the IN included Merino sires by Merino ewes, terminal sires by Merino ewes, maternal sires by Merino ewes and terminal sires by Border Leicester-Merino ewes. Further details of the sire and dam genotypes mated in the IN are provided by Fogarty *et al.* (2007).

Data collection. Prior to lambing, pregnant ewes were randomly allocated to lambing paddocks of 1-20 ha each. Commencing in 2008, to assist in pedigree accuracy, no lambing paddock contained ewes representing more than one individual sire for each sire breed used in the IN. Twice-daily lambing rounds were conducted, with lambs identified with their dams within 18 hours of birth. The lamb measurements and scores recorded are outlined in Tables 1 and 2. Dead lambs, where possible, were measured in the field as appropriate and collected for later autopsy. Lambs were marked and mulesed at an average age of 40 days and then weaned from their dams at an average age of 91 days, with the identity of all surviving lambs recorded at both times. Before marking, daily checks were conducted, with less frequent checking after marking. All deaths were recorded. For more details of the data collected within the IN, see Geenty *et al.* (2009).

Table 1. Lamb measurements/scores recorded from lambing till weaning at all sites in 2008

Time of collection	Measurement/Score
Lambing (lambing rounds lasted 10 days)	Lamb identification, date of birth, birth weight (BWT), type of birth, birth coat score (BCS), estimated lamb age, rectal temperature, lamb vigour scores (overall score and 5 component scores), sex, dam identification, 3 skeletal measurements, cause of death from autopsy on dead lambs and lamb survival to 3 days of age
Marking	Survival to marking
Weaning	Date, lamb weight, type of rearing, survival to weaning

At the 2007 lambing of the IN, no site collected skeletal measurements and only four sites collected data on rectal temperatures, detailed lamb vigour scores and performed autopsies on dead lambs. All other remaining measurements and scores were collected in 2007 as well as 2008.

Table 2. Measurement and scoring systems on lambs for assessed traits

Trait	**Measurements/Scores**
Birth Coat (BCS)	Score range 1 to 7, 1 is no halo hair, 7 is extreme halo hair
Estimated Lamb Age	0= Wet – limited membrane breakage on feet (new born) 1= 1-4 hours old – has walked – still wet and at birth-site 2= > 4 hours old – dry – difficult to catch – follows mother
Rectal Temperature (RT)	At lamb tagging
Lamb Vigour:-	
Overall Score (OBV)	0= lamb still wet – new born - invalid record, 1= Constant struggle, 5= Little movement when held – lies on release
Time-based values (sec)	Based on time taken from release to bleat, stand, contact ewe, contact udder and follow ewe. Recordings stopped after 3 minutes
Skeletal measures (cm):-	
• Crown Rump (CRL)	Distance from back of the skull/nape of the neck to base of the tail
• Metacarpal (ML)	Length of the lower leg bone (front leg – knee to fetlock)
• Thorax	Maximum abdominal circumference around the rib cage
Cause of death	Autopsy procedure outlined by Holst (2004)

*Data Analysis***.** An animal model was fitted to the data using ASREML (Gilmour *et al.* 2006). The specific analytical model fitted IN site, lamb genotype, sex, year of birth, age of dam and any significant interactions as fixed effects, with individual sire fitted as a random effect. Correlations

were calculated using bivariate analyses with fitted effects as described above. Estimated lamb age was used as a covariate for rectal temperature and the lamb behaviour traits (OBV and time-based values). Analyses were performed with and without birth weight in the model as a covariate.

RESULTS AND DISCUSSION

Estimates of heritabilities for traits measured at birth and their phenotypic and genetic correlations with lamb survival (based on 2007 and 2008 IN data) are given in Table 3, where birth weight has not been fitted as a covariate. Although the standard errors for estimates of heritabilities and phenotype correlations are low, indicating good precision, those for genetic correlations are much higher, so the estimates should be regarded as preliminary.

Table 3. Trait heritabilities (h^2) and phenotypic (r_P) and genetic correlations (r_G) between lamb survival to 3 days and to weaning and traits recorded at birth in 2007 and 2008. Results shown are without birth weight fitted as a covariate. Standard errors are shown in brackets

Trait	h^2	Lamb Survival to 3 days		Lamb Survival to Weaning	
		r_P	r_G	r_P	r_G
BWT	0.20 (0.03)	0.21 (0.01)	-0.51 (0.27)	0.24 (0.01)	-0.45 (0.16)
BCS	0.52 (0.06)	0.11 (0.02)	0.37 (0.21)	0.12 (0.01)	0.25 (0.14)
OBV	0.09 (0.02)	0.04 (0.01)	0.24 (0.27)	0.06 (0.01)	0.14 (0.18)
Bleat	0.12 (0.03)	-0.01 (0.02)	-0.57 (0.27)	0.03 (0.02)	-0.38 (0.20)
Stand	0.09 (0.02)	-0.04 (0.02)	-0.39 (0.30)	0.03 (0.02)	-0.45 (0.21)
Contact	0.06 (0.02)	-0.01 (0.02)	-0.48 (0.35)	0.02 (0.02)	0.04 (0.24)
Udder	0.12 (0.04)	0.04 (0.02)	-0.39 (0.32)	0.03 (0.02)	-0.20 (0.23)
Follow	0.06 (0.02)	-0.04 (0.02)	-0.59 (0.40)	0.02 (0.02)	-0.39 (0.27)
RT	0.14 (0.03)	0.36 (0.01)	0.28 (0.23)	0.32 (0.01)	0.31 (0.17)
CRL	0.30 (0.05)	-0.02 (0.02)	-0.55 (0.29)	0.05 (0.02)	-0.29 (0.20)
ML	0.34 (0.06)	0.11 (0.02)	-0.37 (0.25)	0.13 (0.02)	-0.32 (0.20)
Thorax	0.15 (0.03)	0.22 (0.02)	-0.39 (0.31)	0.24 (0.02)	-0.31 (0.23)

Heritabilities were moderate for birth weight and high for birth coat score, similar to earlier reports (Ponzoni *et al.* 1997; Safari *et al.* 2005). However, although the phenotypic correlations between birth weight and both lamb survival traits are low, similar to earlier reports, the genetic correlations are moderate to high and negative, in contrast to the zero or near zero estimates obtained from 2 recent large-scale studies in Merinos (Hatcher *et al.* 2009; Brien *et al.* 2009). Similarly, whilst the phenotypic correlations of birth coat score with both lamb survival traits are low and positive, the genetic correlations are moderate, especially for lamb survival up to 3 days of age and are higher than earlier reports (Hatcher *et al.* 2009; Brien *et al.* 2009). These differences in results may be partly due to the inclusion of several lamb genotypes and multiple sites in our data. Our results require further investigation and analysis, especially once a more complete set of dam pedigrees can be assembled on IN foundation ewes.

Heritabilities for lamb behaviour traits, including OBV and the time-based values, are all in the low range. This likely reflects a high degree of environmental variation, some of which is unavoidable as part of the intervention by humans when collecting the data. The phenotypic correlations of these behaviour traits with both lamb survival traits are zero or near zero, but the genetic correlations between the timed lamb behaviours and lamb survival are in general moderate to high and negative (the shorter the time elapsed from release of the lamb to the expression of the behaviour, the higher the lamb survival), with the estimates being higher for lamb survival to 3 days of age than for the longer period up to weaning. The fitting of birth weight as a covariate in

the analysis of the data made no difference to the size of the genetic correlations between timed lamb behaviour traits and lamb survival. The genetic correlations between OBV and both lamb survival traits are lower and positive (favourable) and could reflect the subjective scoring system and its discrete categories in contrast to the quantitative basis of the timed traits.

Rectal temperature had low heritability but was both phenotypically and genetically moderately correlated with both lamb survival traits in a positive (favourable) direction. Little difference was made to these correlation estimates by the fitting of birth weight as a covariate. Finally, the skeletal measures of lambs had moderate heritability, although thorax circumference was the lowest at 0.15, and despite zero to lowly positive phenotypic correlations with lamb survival traits, all had moderately negative genetic correlations. Fitting birth weight as a covariate more than halved the genetic correlations with lamb survival to 3 days of age (to -0.18, -0.15 and -0.15 for CRL, ML and Thorax, respectively) indicating some remaining correlation with lamb survival that is independent of birth weight. These skeletal measures are a proxy for lamb shape and suggest that lamb shape itself is genetically linked to lamb survival.

The consistent high negative genotypic correlations between timed lamb behaviour traits and survival to 3 days are interesting and suggest further evaluation in future years is justified to improve precision of the correlation estimates. Although these traits are useful in further research, they do not immediately appeal as easy, cheap and quick to measure indicator traits in commercial breeding programs. Although rectal temperature and skeletal measures of lambs at birth are less time consuming to measure and well genetically correlated with lamb survival, especially to 3 days of age, they also require further investigation to determine their commercial usefulness.

In conclusion, whilst these preliminary results from the IN show promise, to date there are no obvious candidates for commercial use as indicator traits for genetically improving lamb survival.

ACKNOWLEDGEMENTS

The IN and associated research programs are supported by the Australian Government's Cooperative Research Centres Program, Meat and Livestock Australia and Australian Wool Innovation Ltd. Resources are provided at IN sites by NSW DPI, the University of New England, Vic DPI, SARDI and DAFWA. Participation by Sheep Genetics and the Sheep Genomics Program is integral. Development and maintenance of the IN database is supported by AGBU.

REFERENCES

Alexander, G. (1984). In "Reproduction in Sheep", p.199, editors D.R. Lindsay and D.T. Pearce, Australian Academy of Sciences: Canberra.

Brien, F.D., Hebart, M.L., Jaensch, K.S., Smith, D.H. and Grimson, R.J. (2009). *Proc. Assoc. Advmt. Anim. Breed. Genet.* **18:** *these proceedings.*

Fogarty, N.M., Banks, R.G., Van der Werf, J.H.J., Ball, A.J. and Gibson, J.P. (2007). *Proc. Assoc. Advmt. Anim. Breed. Genet.* **17:**29.

Geenty, K.G., van der Werf, J.H.J., Gore, K.P., Ball, A.J. and Gill, S. (2009). *Proc. Assoc. Advmt. Anim. Breed. Genet.* **18:** *these proceedings.*

Gilmour, A.R., Gogel, B.J., Cullis, B.R. and Thompson, R. (2006). 'ASReml User Guide Release 2.0.' (VSN International Ltd: Hemel Hempstead).

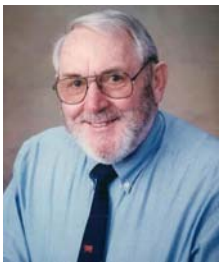
Hatcher, S., Safari, A. and Atkins, K.D. (2009). *J.Anim.Sci.* (Submitted).

Holst, P.J. (2004). Lamb Autopsy. Notes on a procedure for determining cause of death. *www.dpi.nsw.gov.au/agriculture/livestock/sheep/health/other/lamb-autopsy*

Ponzoni, R.W., Grimson, R.J., Jaensch, K.S., Smith, D.H. and Hynd, P.I. (1997). *Int.J.Sheep.Wool.Sci.* **45**:12.

Safari, E., Fogarty, N.M. and Gilmour, A.R. (2005). *Livest.Prod.Sci.* **92**:271.

JOHN EDWARD VERCOE AM FSTE FAIAST
1936 – 2005



Dr John Vercoe was a former president of AAABG who, over five decades, made an outstanding contribution to the northern cattle industry in Australia and livestock production internationally.

John graduated in Agricultural Science from the University of Melbourne, worked with the Victorian Department of Agriculture at Werribee and was awarded a MAgrSc (Hons) for his research on feed intake of grazing sheep, before returning to University of Melbourne to complete his PhD studies with Professor Derek Tribe. He completed his post-doctoral research at the Hannah Research Institute in the UK under the guidance of Sir Kenneth Blaxter, FRS.

On joining CSIRO Rockhampton in 1964, John studied the genetic and non-genetic factors that influence beef production in stressful tropical environments, examining the effect of a range of environmental stresses on nitrogen and energy metabolism of different cattle breeds. With research colleague, Dr John Frisch, he developed a concept that greatly influenced the understanding of the origins of genotype x environment interactions and consequently the breeding and management strategies necessary to increase cattle production in tropical and other stressful environments.

As Assistant Chief of CSIRO Tropical Animal Production (1982-1993) and Officer-in-Charge of the Tropical Cattle Research Centre for more than a decade, John had a profound influence on the direction of beef research in northern Australia. His own research demonstrated the need and the value of interdisciplinary research, and highlighted the limitations of a single-discipline approach to solving industry problems. He was a long-time advocate of cooperative research efforts and initiated moves to implement them. As a member of the Cooperative Research Centre's Life Sciences Assessment Panel and the Official Visitor to several agriculturally and environmentally related CRCs, John provided great insights into the value and benefits of collaborative R&D activities and structures.

Of unfailing good humour, John's cup was always half full. His ability to relate to people and work with them was one of John's most outstanding qualities. He was also a great mentor of junior scientists in CSIRO and elsewhere, nationally and internationally. Many young, and more not-so-young, scientists have John to thank for their initial scientific profiles.

John retired from CSIRO in 1996 but continued to pursue his passion for improving livestock production in Australia and abroad. He was a consultant with the Australian Centre for International Agricultural Research (ACIAR) and the World Bank in Sudan. He served two 3-year terms as Head, Animal Production & Health, Joint FAO/IAEA in Vienna (1972-1974 and 1978-1980), establishing a wide network of contacts that were extremely valuable later in his career. For varying periods, he was Chairman of the Board of Trustees for the International Livestock Research Institute (Nairobi), Chairman of the Committee of Board Chairs of the Consultative Group on International Agricultural Research (CGIAR, Washington) and Queensland Coordinator of the Crawford Fund, which supports international agricultural research.

In 1997, John was conferred a Member of the Order of Australia (AM). The citation read: *For service to the cattle industry through research into genetic factors affecting the efficiency of beef production in tropical environments.* He was also a recipient of a Centenary Medal in 2001 - *for service to primary industries and science and technology;* a Fellow of the Australian Academy of Technological Sciences and Engineering; a Fellow of the Australian Institute of Agricultural Science; a Fellow of the Australian Society of Animal Production; an Honorary Professor at the University of Central Queensland; and an Honorary CSIRO Research Fellow. Above all, John was a real gentleman and a true industry politician.

Proc. Assoc. Advmt. Anim. Breed. Genet. 18:113-116

INTERBREED EVALUATION OF BEEF CATTLE PRODUCTIVITY UNDER LOW AND MODERATE DRY MATTER AVAILABILITIES

T.G. Jenkins

CaTo Consulting LLC, Clay Center NE, USA

SUMMARY

Costs limit live animal experimentation to quantify genotype by environment interactions in herd productivity due to the number of production environments and beef cattle germ plasm combinations. Life cycle simulation models enable producers to evaluate breeds differing in genetic potential for productivity under a variety of environmental conditions; e.g., inter- or intra-breed by environment interactions. A simulation model developed for within ranch evaluations, DECI, was parameterized to evaluate productivity for a constant calving population inventory herd of two biological types of beef cattle across two production environments differing only in annual yields of dry matter (DM), low and moderate. The two biological types evaluated differed in genetic potentials for mature size and peak milk yield with each type having associated attributes; e.g., propensity to fatten, measures of fertility, dystocia and maintenance requirements. Ranking for weight of calf marketed per cow exposed of the two biological types differed depended on environment. Under restrictive DM the biological type with greater genetic potential for growth and milk production marketed fewer kilograms of product per cow exposed. At moderate DM availability, the biological type with greater genetic potential yielded more of this product. Application of dynamic simulation models provides an opportunity to investigate genotype by environment interactions.

INTRODUCTION

Hammond (1947) stated "… thus environmental condition existing at any given time will lead to the natural selection of genes giving rise to characters in harmony with the environment concerned." Failure to realize greatly increased levels of productivity when using animals with greater genetic potential is well documented (Lin and Togash 2002); this is due to emphasis on output rather on the "harmony" of the animal with the challenges to be encountered in the production environment. For beef cattle producers, environment is made up of all non-genetic factors; those normally considered include green grass days, temperature, humidity, parasites, etc., but factors not normally considered are the primary product and the marketing end points for the primary product. The challenge is to evaluate productivity of biological types with diverse genetic potentials for production across a wide range of environmental conditions. The present study applies a dynamic simulation model to characterize the productivity of two biological types differing in genetic production potentials interacting with differing feed environments.

MATERIALS AND METHODS

Model information. A dynamic, mechanistic life cycle herd level simulation model capable of tracking daily events for individual animals was parameterized to evaluate productivity of two biological types differing in genetic potentials in two nutritional environments. The Decision Evaluator for the Cattle Industry (DECI) is a herd inventory model incorporating mechanistic and empirical equations at the tissue level to predict cattle performance (Jenkins and Williams 1998; Williams *et al.* 2006). DECI was designed to allow producers to evaluate the impact of strategic management decisions on herd productivity over time, including the interaction of genetic potentials of cattle with various environmental conditions (management, nutritional, and market

endpoint) and to identify constraints that limit full expression of genetic potential. Feed environments are defined by defining forage availability and quality, either grazed or harvested, and times of availability during the production year. Breed means derived from breed evaluations in the Germ Plasm Evaluation and Germ Plasm Utilization projects at MARC (Cundiff *et al*. 1998; Gregory *et al*. 1999) are used as the genetic potentials for 20 characters contributing to an animal's phenotypic performance. These potentials are stored in a data set accessible internal to the program. Mating systems provide the user with the ability to use heterosis and breed differences. Maximum cowherd size per simulation is 500 pregnant cows at the beginning of the inventory year.

Approach. Two production environments were parameterized to differ only in availability (pasture and harvested supplements) of dry matter (DM), all other environmental factors being held constant. Forage resources available to the animals were programmed to reflect a normal forage production year for expected forage protein and energy content based on dry matter yields. This approach allows seasonal variation in the two dry matter availability (indicator of energy intake) environments that were simulated: 4000 kg and 6700 kg per cow per year. Other management factors in common were simulated herd size (set at 150 pregnant females at the beginning of the inventory year), breeding seasons were 90 days using bull genetic potentials within biological type, simulated male calves were castrated soon after birth, calves did not receive supplemental feeds, and all calves were weaned on a single day. Two biological types differing in genetic potential, Moderate (M) and Greater (G) for weight and maturity and peak milk were identified (Table 1). For each type the simulation model was parameterized for all characters using the genetic potentials of breeds derived from the GPE and GPU (studies imbedded in a data base of DECI) associated with each type. Three 15 year simulations created a final herd of the correct genetic potential that was at a herd age equilibrium within each nutritional environment.

Table 1. Genetic potentials for two biological types of cattle

	Moderate	Greater
Potential		
Birth weight, kg	40	42
Mature weight, kg*	630	744
Peak milk yield, kg**	10.4	11.4

*at 26% body fat
**at time of peak lactation

RESULTS AND DISCUSSION

Predicted performance. Predicted values for traits of interest within each nutritional environment are reported in Table 2. Within the 4000 kg DM environment, predicted responses for cows 5 years or older of the two biological types are reported for weights and condition scores at time of calving, start of the breeding season, and weaning time. Under restricted DM availability, predicted body weights of M mature cows varied across the production year as demand for production energy changed. Response in condition scores (9-point system as an indicator of energy status at each of the three reporting times) suggests DM availability on average would have limited impact on expression of M genetic potential for productivity for either cows or calves in the herd. The G mature cows with greater genetic potential of mature weight and peak milk yield than the M, had predicted weights either less than or equal to those of the M. Low body condition scores suggest the probability of reproductive success for G females would be lower in

the 4000 kg environment. For each herd, predicted mean postpartum interval of the M cows was 16 days shorter than that of G cows (68 d and 84 d; respectively) which could be attributed to negative energy balance of the G cows. With a restricted breeding season longer postpartum intervals contribute~~s~~ to lower pregnancy rates, so that predicted herd mean pregnancy rates were 86 and 74% for the M and G; respectively. To sustain a cow herd inventory of 150 pregnant females at the start of the inventory year, a mean of 182 females were exposed each year for the M herd in the 4000 kg environment compared to 200 females for the G herd. The difference in females exposed reflects a difference in the number of heifers that need to be retained to meet the requirement of 150 pregnant females at the start of the inventory year.

Table 2. Predicted performance of two biological types in two DM availability environments

	4000 kg annual DM		6700 kg annual DM	
	Moderate	Greater	Moderate	Greater
Mean 5-year-old cow				
Calving time wt, kg	555	509	647	684
Breeding time wt, kg	505	505	661	683
Weaning time wt, kg	522	483	627	635
*Mean 5-year-old cow CS**				
Calving time	5.5	2.8	7.0	5.8
Breeding time	4.5	2.6	7.1	5.8
Weaning time	4.8	2.5	6.6	5.0
Mean herd reproduction				
Postpartum interval, d	68	84	62	67
Number exposed	182	200	181	180
Pregnancy rate at weaning, %	86	74	93	94
Mean herd productivity				
Milk yield, kg/d	6.6	6.7	6.6	7.7
Number calves weaned	141	141	144	142
Weight/calf weaned, kg	195	224	212	253
Weight of calf marketed/ cow exposed	151	137	168	199

*CS condition score 9-point system 1 very thin – 9 very obese

With increased dry matter availability, both groups had higher predicted cow weights at each production segment and G cows with greater genetic potential were heavier than M cows although the M population continued to have higher CS. Both M and G exhibited improvement in reproduction and productivity traits in the better nutritional environment with the greatest improvement predicted for the G cows. Sufficient DM availability in the 6700 DM environment enabled the G genetic potentials for growth and milk production to be expressed and reproductive rate improved (fewer females exposed to maintain herd size) allowing 45% more weight of calf marketed per cow exposed. The M cattle had improved reproductive rates in the 6700 DM environment with slight improvement in predicted calf weight at weaning (17 kg) resulting in 11% improvement in weight of calf marketed per cow exposed.

Ranking of the M and G populations was dependent on the nutritional environment creating a biological type by environment interaction. Dry matter availability that varied across the

production year was the constraining factor on herd productivity for the G population in the 4000kg DM environment with the availability during critical times of the year limiting the fertility and milk production of the cows and full expression of the growth potential of the calves. Maintaining a constant herd size of pregnant females required more replacement females to be retained in the breeding herd each year leading to reduced number of weaned calves for sale-reducing kg of calves marketed per female exposed. A CS of 2.6 for G population of mature females would have a smaller proportion of these females fertile during the restricted breeding season. Conversely, the M population's CS of 4.5 at the start of the breeding season indicate a higher probability for a greater proportion of the mature cows to cycle and conceive. Nutritional requirements more closely met the needs for expression of milk production in cows and for the growth potential of the calves, thus fewer replacement heifers were required to maintain herd inventory resulting in a higher yield of calf weight marketed per female exposed the previous breeding season. At increased DM availability (6700 kg), the rank of the two populations for this measure of efficiency changed. The reproduction rate of the G population increased by 20% resulting in an increase of 32 kg per female exposed (more heifers available for market). Sufficient body energy reserves and forage availability to not limit reproduction or the expression of genetic potential for milk production of the cow or the growth protential of the calves. The response in productivity at the 6700 kg of the M population was positive for reproduction and growth traits but to a lesser degree. The greater CS scores of mature cows at all times during the production year indictes cows were consuming energy in excess of their genetic potential for production. Even with unlimited energy availability in the 6700 kg environment, the M population average calf weight at weaning only increased 9% indicating limited genetic potential for growth of the calves.

CONCLUSIONS

Genotype by environment interactions affect herd productivity of cattle. Rank changes of importance are those that affect the relevant measure of productivity for a producer. The greatest impact of nutritional constraint in a cow herd is on the reproduction potential. For those interactions resulting from differing feed need attributable to greater genetic potential of one biological type relative to others, dynamic herd simulation models represents a tool to quantify the feed resources needed to benefit from the greater genetic potentials. With the diversity of genetic potential within and among cattle breeds, live animal experimentation is not feasible to investigate all potential genotype by environment interactions. Application of this tool expands the opportunity to evaluate the interactions for many breeds or breed crosses across a wide range of environmental conditions.

REFERENCES

Cundiff, L.V., Gregory, K.E. and Koch, R.M. (1998) *J. Anim. Sci*. **76**:2528.
Gregory, K.E., Cundiff, L.V. and Koch, R.M. (1999) USDA Tech. Bull. Nu. 1875.
Hammond, J. (1947) *Biol. Rev*. **22**: 195.
Jenkins, T.G. and Williams, C.B. (1998) *Proc. 6th World Congress on Genetics Applied to Livestock Production* **27**:461.
Lin, C.Y. and Togash, K. (2002) *Anim. Sci. J*. **73**:3.
Williams, C.B. and Jenkins, T.G. (2003) *J. Anim. Sci*. **81**:1371.

Proc. Assoc. Advmt. Anim. Breed. Genet. 18:117-120

MATERNAL PRODUCTIVITY IN INDUSTRY HERDS: PRELIMINARY RESULTS

K. A. Donoghue[1] and P. F. Parnell[2]

Cooperative Research Centre for Beef Genetic Technologies
[1] NSW Department of Primary Industries, Agricultural Research Centre, Trangie, NSW 2823
[2] NSW Department of Primary Industries, Beef Industry Centre, Armidale, NSW 2351

SUMMARY

This paper presents results of analyses of 60% of the data from the initial 3 years of a large maternal productivity study involving 8 Angus co-operator herds. The genetic parameter estimates for weight and body composition measured repeatedly in Angus females at different time points during the first 2 parities tended to be consistent with expectations based on previous single point estimates reported in the literature. Further analyses of these data, together with information on early life performance as yearlings and subsequent measures of maternal productivity, will provide unique information on the likely consequences of selection for traits associated with body composition.

INTRODUCTION

Growth and carcase traits have received much attention in beef genetic improvement programs, and Australian beef producers have been successful in achieving significant genetic gains in these traits (Johnston 2007). The Cooperative Research Centre for Beef Genetic Technologies (Beef CRC) is investing in the development of new DNA based tools to enable beef producers to achieve genetic improvement from their breeding programs in "difficult to measure" traits eg. feed efficiency, meat tenderness. However, breeders may not fully utilise new opportunities for genetic improvement without better knowledge about the potential impacts on performance traits of the breeding cow herd, here termed maternal productivity.

This paper describes the preliminary results of a Beef CRC project designed to address concerns about the potential negative impacts on maternal productivity that may accompany genetic selection for traits associated with body composition (e.g. feed conversion efficiency and carcase traits). The project will determine genetic relationships between body composition and components of breeding herd productivity and explore potential new early-life selection criteria for improving maternal traits. This paper describes results from analyses of a subset of the data from the initial 3 years of the project, where body composition measurements have been recorded in 8 Angus co-operator herds.

MATERIALS AND METHODS

Data. A total of 7,226 liveweight and ultrasound body composition measurements were made on 3,775 Angus females in 8 co-operator herds. These measurements were taken at 2 points during the annual production cycle (Pre-Calving and Weaning) during the first and second parities and represent approximately 60% of the final data expected. In this analysis, four traits were considered: weight (Wt); P8 (P8) and rib (Rib) fat measures; and eye muscle area (EMA). Each female could have up to 4 measurements of these traits: Pre-Calving (PC1) and Weaning (W1) in Parity 1, and Pre-Calving (PC2) and Weaning (W2) in Parity 2. There were various reasons for females not having a complete set of measurements; age of animal (Parity 2 not yet recorded), disposal from herd, missing on measurement day. Animals with incomplete pedigrees, missing birth date and trait measurements greater than 3 standard deviations from the contemporary group mean were removed from the analysis.

Model of analysis. Variance components were estimated using ASReml (Gilmour *et al*. 2006). The fixed effect of contemporary group (CG) was included in the model and single record contemporary groups were excluded from the analysis. Contemporary group definition included herd, calving season (spring or autumn), date of measurement and breeder-defined management group. Age of the animal on the date of first measurement was included as a covariate. Random effects fitted included a term for direct genetic effects. Pedigree records for all animals with records and 2 further generations of ancestors were used. Multivariate analyses of the same trait across the 4 different time points were conducted separately for each trait. In addition, multivariate analyses among the four different traits were conducted separately for each of the time points.

RESULTS AND DISCUSSION

Data. Table 1 contains summary statistics of the data for each parity. Phenotypic variance and trait averages for P8 fat increased steadily over the 4 time points. There was a trend in the other traits, however, for decreased phenotypic variance and/or trait average at Pre-Calving in Parity 2 compared to Weaning in Parity 1.

Table 1. Descriptive statistics, phenotypic variances (σ^2_p) and heritabilities (h^2) for traits of females recorded at Pre-Calving and Weaning in 1st and 2nd parities

	Average (SD)	Minimum	Maximum	σ^2_p *(SE)*	h^2 *(SE)*
Parity 1					
Pre-Calving (n=3,016)					
Age (days)	685 (35)	585	905	-	-
P8 fat (mm)	5.2 (2.6)	1	21	3.1 *(0.10)*	0.47 *(0.06)*
Rib fat (mm)	4.1 (1.9)	1	13	1.7 *(0.06)*	0.54 *(0.07)*
Weight (kg)	476 (68)	291	748	1,115 *(37)*	0.48 *(0.06)*
Eye Muscle Area (cm^2)	56 (11)	20	89	39 *(1.1)*	0.28 *(0.05)*
Weaning (n=1,527)					
Age (days)	895 (52)	785	1,151	-	-
P8 fat (mm)	6.2 (2.7)	1	28	5.8 *(0.25)*	0.40 *(0.09)*
Rib fat (mm)	5.5 (2.1)	1	24	3.7 *(0.16)*	0.34 *(0.08)*
Weight (kg)	530 (64)	303	760	2,117 *(83)*	0.21 *(0.07)*
Eye Muscle Area (cm^2)	63 (8)	30	89	38 *(1.5)*	0.26 *(0.07)*
Parity 2					
Pre-Calving (n=1,755)					
Age (days)	1,049 (38)	884	1,261	-	-
P8 fat (mm)	6.4 (3.3)	2	26	6.6 *(0.25)*	0.32 *(0.07)*
Rib fat (mm)	5.0 (2.3)	2	20	3.2 *(0.12)*	0.19 *(0.06)*
Weight (kg)	562 (65)	351	774	1,913 *(71)*	0.23 *(0.06)*
Eye Muscle Area (cm^2)	62 (10)	33	94	47 *(1.7)*	0.24 *(0.06)*
Weaning (n=924)					
Age (days)	1,258 (53)	1,129	1,454	-	-
P8 fat (mm)	9.1 (4.1)	2	26	9.3 *(0.51)*	0.43 *(0.11)*
Rib fat (mm)	7.6 (3.1)	2	20	5.6 *(0.31)*	0.39 *(0.11)*
Weight (kg)	597 (74)	402	846	2,673 *(139)*	0.31 *(0.10)*
Eye Muscle Area (cm^2)	67 (10)	37	97	41 *(2.1)*	0.23 *(0.09)*

Heritabilities (h^2) for P8 fat and EMA across the 4 time points were relatively stable and are similar to published estimates for yearling ultrasound scan measurements (eg. Meyer *et al.* 2004). The heritabilities for rib fat and weight fluctuated across the time points, with very high estimates

for both traits at the first time point (Pre-Calving in Parity 1). All heritabilities had large associated standard errors, particularly traits measured at Weaning in Parity 2.

Genetic correlations (r_g) and their associated standard errors between measurements of the same trait (weight, P8 fat, rib fat or EMA) across the 4 different time points are reported in Table 2. Very high positive r_g were observed among weight measurements at the 4 different time points, while r_g for P8 fat, rib fat and EMA were high to very high across all time points. For all traits, the highest r_g was observed between the same time point in different parities, while the lowest r_g occurred between Pre-Calving in Parity 1 and Weaning in Parity 2. For some traits, this latter r_g would indicate that significant re-ranking of animals would occur between those different time points, and that performance at Pre-Calving in Parity 1 for these traits is not a good indicator of performance at Weaning in Parity 2.

Table 2. Genetic correlations (SE) across time points for the same trait

Trait*	PC1-W1**	PC1-PC2**	PC1-W2**	W1-PC2	W1-W2	PC2-W2
Wt	0.91	0.98	0.90	0.92	0.90	0.92
	(0.06)	*(0.03)*	*(0.07)*	*(0.06)*	*(0.10)*	*(0.05)*
P8	0.78	0.91	0.73	0.85	0.94	0.87
	(0.08)	*(0.06)*	*(0.09)*	*(0.06)*	*(0.05)*	*(0.06)*
Rib	0.67	0.78	0.67	0.92	0.92	0.90
	(0.10)	*(0.09)*	*(0.10)*	*(0.05)*	*(0.07)*	*(0.07)*
EMA	0.76	0.93	0.64	0.84	0.89	0.82
	(0.11)	*(0.07)*	*(0.16)*	*(0.09)*	*(0.14)*	*(0.12)*

* Weight (Wt), P8 fat (P8), rib fat (Rib) and eye muscle area (EMA)

** Pre-Calving (PC1) and Weaning (W1) in Parity 1; Pre-Calving (PC2) and Weaning (W2) in Parity 2

Estimates of genetic correlations (and associated standard errors) between the 4 different traits within the same time point (Pre-Calving or Weaning in Parity 1, or Pre-Calving or Weaning in Parity 2) are reported in Table 3. The r_g between P8 and rib fat was consistently high across all time points and similar to earlier literature estimates (eg. Meyer *et al.* 2004). The r_g estimates between the fat traits and weight tended to be low and positive (highest at Weaning in both parities), except for weak negative correlations at Pre-Calving in Parity 2. Similar trends were observed for genetic correlation estimates between the fat measures and EMA. The r_g estimates between weight and EMA were low to moderate and positive, but fluctuated considerably across time.

Only approximately 60% of the target number of records for this project were available for this analysis. As more data become available, the parameter estimates reported above will be re-estimated with greater precision. In addition, further analyses will examine the phenotypic and genetic relationships between these measures and prior ultrasound and weight measurements taken at yearling age, and with subsequent maternal performance, including calf weaning weights, fertility (e.g. re-conception rates) and longevity (herd life). Genetic relationships between the present measures and structural soundness scores will also be examined on females that have been assessed for structural traits. All early-life measures will be examined especially for their usefulness as indicators of lifetime productivity.

Table 3. Genetic correlations (SE) between traits within time points

Time*	P8-Rib**	P8-Wt**	P8-EMA**	Rib-Wt	Rib-EMA	Wt-EMA
PC1	0.87	0.18	0.30	0.13	0.16	0.53
	(0.03)	*(0.10)*	*(0.11)*	*(0.10)*	*(0.12)*	*(0.09)*
W1	0.83	0.38	0.49	0.34	0.43	0.30
	(0.05)	*(0.17)*	*(0.15)*	*(0.19)*	*(0.17)*	*(0.20)*
PC2	0.82	-0.04	0.04	-0.12	-0.10	0.38
	(0.06)	*(0.19)*	*(0.18)*	*(0.23)*	*(0.22)*	*(0.16)*
W2	0.74	0.34	0.44	0.36	0.22	0.23
	(0.09)	*(0.20)*	*(0.20)*	*(0.20)*	*(0.24)*	*(0.25)*

* Pre-Calving (PC1) and Weaning (W1) in Parity 1; Pre-Calving (PC2) and Weaning (W2) in Parity 2
** Weight (Wt), P8 fat (P8), Rib fat (Rib) and eye muscle area (EMA)

CONCLUSIONS

Results of preliminary analyses of a large Beef CRC maternal productivity study showed genetic parameter estimates for traits of Angus females, measured at different time points during their first 2 parities, tended to be consistent with previous single point estimates from the literature. Further analyses of these data, together with measures of yearling performance and subsequent maternal productivity, will provide information on the likely consequences of selection for traits associated with body composition changes. In addition to identifying early indicators of lifetime maternal performance, this project will develop a comprehensive phenotypic database of breeding female performance and body composition on temperate genotypes which can be later used for evaluation and validation of the effect of future DNA markers.

ACKNOWLEDGMENTS

This work was funded by the Cooperative Research Centre for Beef Genetic Technologies. The authors gratefully acknowledge the co-operator herds and the ultrasound technicians who collected the data.

REFERENCES

Johnston, D. (2007) *Proc. Assoc. Advmt. Anim. Breed. Genet.* **17**:8.

Gilmour, A.R., Gogel, B.J., Cullis, B.R., Welham, S.J. and Thompson, R. (2006) "ASReml User Guide Release 2.0" VSN International Ltd, Hemel Hempstead.

Meyer, K., Johnston, D. and Graser, H. (2004) *Aust. J. Agric. Res.* **55**:195.

Proc. Assoc. Advmt. Anim. Breed. Genet. 18:121-124

GENETIC VARIATION IN GROWTH, HORMONAL AND SEMINAL TRAITS OF YOUNG TROPICALLY ADAPTED BULLS

N.J. Corbet[1], B.M. Burns[2], D.H. Corbet[2], D.J. Johnston[3], J. M. Crisp[4], M. R. McGowan[4], K. C. Prayaga[1], B. K. Venus[5] and R.G. Holroyd[2]

Cooperative Research Centre for Beef Genetic Technologies
[1] CSIRO Livestock Industries, Rendel Laboratory, Rockhampton, QLD 4701
[2] Queensland Primary Industries and Fisheries, Parkhurst, Rockhampton, QLD 4701
[3] Animal Genetics and Breeding Unit, University of New England, Armidale, NSW 2351
[4] School of Veterinary Science, The University of Queensland, St. Lucia, QLD 4072
[5] Queensland Primary Industries and Fisheries, Animal Research Institute, Yeerongpilly, QLD 4105

SUMMARY

Except for scrotal circumference there is little published information on the genetic variation of traits measured on young tropical breed bulls. The current study assesses growth, hormonal and seminal traits measured in Brahman and tropical Composite bulls (n=2212) from weaning through to 2 years. Heritability of scrotal size (39 to 75%), prepubertal serum concentration of inhibin (40 to 71%) and semen quality traits (19 to 27%) indicated potential to genetically improve these traits in both Brahman and tropical Composite populations. Genetic correlations between male traits varied (-46 to 78%), some suggesting antagonism between traits while others indicated that some traits measured in bulls early in life could potentially be indicators of post-pubertal semen quality.

INTRODUCTION

Reproductive performance of a cattle population has, arguably, the single greatest influence on beef economic returns. Reproductive performance has many components and BREEDPLAN currently provides EBVs for scrotal circumference (SC) and days to calving (DTC) to aid selection for improved herd fertility. Studies have reported favourable genetic relationships between SC and female fertility traits (Morris *et al*. 1993; Martinez-Velazquez *et al*. 2003). However, there is only sparse information published on the genetic relationships among desirable traits measured on bulls in Australian tropical beef breeds. The current study provides a preliminary report on genetic variation in early-life male fertility traits and estimates of genetic correlations.

MATERIALS AND METHODS

Animals. Data were obtained from bulls of two genotypes (908 Brahmans and 1304 tropical Composites) which were progeny of cows bred for the Beef CRC northern Australia breeding project (Johnston *et al*. 2009). Composites bulls comprised admixes of Belmont Red, Charbray, Santa Gertrudis and Senepol breeds. The bulls were bred on 5 properties across central, northern and western Queensland over 4 years using sires selected to ensure representation of industry populations and genetic linkage across years and properties within genotype. At weaning, 1494 bull calves (average of 374 per year) were relocated by road transport to Brigalow Research Station (170km SW of Rockhampton). The remaining 718 bulls were born at Belmont Research Station (25km NW of Rockhampton) and remained there post-weaning. At Brigalow and Belmont all bulls weaned in the same year were managed as a single group (defined as a cohort) until completion of data collection as 2 year olds. Breeding of a further 3 cohorts is planned.

Measurements. At ~4 months of age, blood samples were taken to measure pre-pubertal serum inhibin (INH4), a trait linked to reproductive function (Phillips 2005). When weaned at ~6 months, flight time (FLT6) was recorded. Scrotal circumference was measured at 9 (SC9), 12 (SC12) and at 24 months of age (SC24). Weights were also recorded at these intervals (WT9, WT12, WT24), and body condition score (BCS12; scored 1 to 5) was recorded at 12 months. At 15 months the bulls were scanned to measure fat depth at the P8 site (Fat15) and eye-muscle area (EMA15). An ejaculate from bulls with SC ≥ 20cm was collected by electro-ejaculation at 12 and 24 months. Traits recorded on the ejaculate included presence or absence of sperm at 12 months (Sperm12; 1 or 0), and percent morphologically normal sperm at 24 months (Norm24; 0 to 100%). Table 1 lists the descriptive statistics of the traits measured.

Table 1. Descriptive statistics of traits measured on young tropical breed bulls

Trait*	Units	Brahman			Composite		
		N	Mean	SD	N	Mean	SD
INH4	ng/ml	580	7.8	1.97	790	8.2	1.91
FLT6	seconds	908	1.17	0.59	1304	1.16	0.45
SC9	cm	796	18.0	1.76	1136	21.4	2.90
SC12	cm	902	21.6	2.95	1297	26.5	3.28
SC24	cm	892	30.4	3.19	1281	31.2	2.78
WT12	kg	904	247	33.1	1295	267	39.5
WT24	kg	898	380	42.5	1284	378	46.0
BCS12	score 1 to 5	903	2.5	0.31	1299	2.5	0.31
Fat15	mm	903	1.4	0.57	1295	1.1	0.32
EMA15	cm^2	903	46	8.4	1293	49	8.1
Sperm12	binomial	798	0.10	0.30	1130	0.52	0.50
Norm24	%	765	69	23.2	1112	74	18.9

* See text for trait definitions; INH4 not measured on first 2 cohorts; SC9 not measured on first cohort.

Statistical analyses. Significant fixed effects were identified separately for each genotype using linear mixed model procedures of GenStat (10^{th} Edition). Models included the fixed effects of year (2004 to 2007), birth location (5 properties), post-weaning location (Brigalow or Belmont), dam age (3, 4 or 5+ years), birth month (Sep. to Jan.), their interactions and sire as a random effect. WT was included as a covariate for SC traits to adjust to a common body mass. For the Composite data, sire and dam breed groups were included to account for heterosis effects. Non-significant terms were sequentially removed from the model to yield the final model for each trait. Phenotypic variances and trait heritabilities were estimated in univariate analyses using ASReml (Release 2.0). The animal models used included the final fixed effects identified above with additional maternal random terms to model genetic and permanent environment effects of the dam. Genetic correlations among bull traits were estimated in a series of bivariate analyses between paired traits. The relationship matrix was derived from a pedigree of 6081 animals going back 5 generations.

RESULTS AND DISCUSSION

Estimates of phenotypic and genetic variance parameters for the traits measured are presented in Table 2. The direct heritability of most traits was generally moderate indicating that improvement of the traits could readily be made by selection. The low heritability estimates of

BCS12 in both breeds and Fat15 in Composites could reflect that these pasture-fed bulls did not vary greatly in body fatness. Heritability estimates of INH4, FLT6, WT12 and Fat15, although associated with high error, may suggest breed differences reflecting underlying genetic and physiological differences between Brahmans and Composites for these traits. The heritabilities for INH4 and SC24 were high in Brahmans and although no published estimates for either trait were cited, the heritability of SC24 was greater than published estimates for SC at 18months (Burrow 2001). Further data are needed to validate these estimates. Maternal genetic and non-genetic effects were important for some traits, including those measured post-weaning, indicating the need to consider maternal effects in analytical models to ensure accurate estimation of direct heritability. Martinez-Velazquez (2003) reported maternal effects of similar magnitude for SC12 in *Bos taurus* breed bulls.

Table 2. Phenotypic variance (V_p), direct heritability (h^2), maternal heritability (m^2) and permanent environment effect (c^2) estimated for traits measured on young bulls

Trait*	Brahman				Composite			
	V_p	h^2	m^2	c^2	V_p	h^2	m^2	c^2
INH4	3.17	0.71 (0.21)	0.04	0.00	2.65	0.40 (0.13)	0.23	0.06
FLT6	28.4	0.18 (0.08)	0.02	0.00	17.1	0.33 (0.11)	0.02	0.00
SC9	2.11	0.41 (0.13)	0.22	0.00	5.64	0.49 (0.12)	0.00	0.11
SC12	4.97	0.59 (0.15)	0.15	0.00	6.32	0.39 (0.10)	0.00	0.11
SC24	6.6	0.75 (0.18)	0.04	0.01	5.4	0.46 (0.11)	0.00	0.16
WT12	554	0.31 (0.12)	0.13	0.00	671	0.18 (0.07)	0.23	0.00
WT24	950	0.35 (0.12)	0.14	0.00	1104	0.39 (0.11)	0.17	0.00
BCS12	0.05	0.14 (0.07)	0.00	0.09	0.04	0.15 (0.06)	0.00	0.00
Fat15	0.310	0.21 (0.09)	0.02	0.11	0.100	0.03 (0.04)	0.04	0.00
EMA15	28.9	0.25 (0.10)	0.09	0.00	30.3	0.26 (0.09)	0.09	0.00
Sperm12	0.084	0.27 (0.11)	0.11	0.00	0.223	0.19 (0.07)	0.00	0.00
Norm24	512	0.20 (0.09)	0.00	0.00	355	0.24 (0.08)	0.00	0.01

***** See text for trait definitions; standard error shown in parentheses.

Estimated genetic correlations between traits measured on the bulls are presented in Table 3. For brevity only traits of higher economic importance or heritability have been included here. High standard errors generally indicate fewer observations available for analyses. INH4 had a positive genetic correlation with SC9 and SC12 but, except with Sperm12 in Brahmans, tended to show a negative relationship with semen quality traits (Sperm12 and Norm24). These relationships suggest higher levels of inhibin at 4 months are associated with the development of larger testes in pre-pubertal bulls but, conversely, associated with poorer semen quality post-puberty. However, high standard error warrants further data before dismissing INH4 as a predictor of male fertility.

FLT6 was positively correlated with SC in both breeds and Norm24 in Brahmans indicating that selection for less flighty bulls (longer flight time) would improve SC in both breeds and improve percent normal sperm in Brahmans. In Composites, however, there was no relationship between flight time and percent normal sperm. The genetic relationships between WT12 and semen traits were negative suggesting genetic antagonism between growth and semen quality. The relationships of Fat15 and EMA15 with SC12 and Sperm12 were also negative and, although behavioural studies were not incorporated for confirmation, may be influenced by depletion of

body tissue reserves due to mounting activity of bulls reaching puberty. However, negative relationships between EMA15 and Norm24 in Composites and between WT12 and Norm24 in both breeds are of concern if selection emphasis is on improved EMA and weight.

Table 3. Estimates of genetic correlation (and standard error) among bull traits

Trait	INH4	FLT6	WT12	Fat15	EMA15	Sperm12	Norm24
Brahman							
SC9	0.50 (0.20)	0.42 (0.25)	0.23 (0.25)	-0.28 (0.28)	-0.01 (0.25)	0.36 (0.25)	-0.11 (0.29)
SC12	0.39 (0.20)	0.62 (0.19)	0.29 (0.23)	-0.22 (0.26)	-0.16 (0.24)	0.76 (0.14)	0.36 (0.24)
Sperm12	0.26 (0.26)	0.10 (0.40)	-0.01 (0.43)	-0.46 (0.26)	-0.21 (0.29)	-	-
Norm24	-0.33 (0.28)	0.32 (0.31)	-0.41 (0.28)	0.02 (0.34)	0.03 (0.32)	0.13 (0.33)	-
Composite							
SC9	0.21 (0.22)	0.34 (0.20)	-0.01 (0.26)	0.05 (0.46)	-0.20 (0.23)	0.49 (0.20)	0.40 (0.22)
SC12	0.29 (0.23)	0.31 (0.22)	0.21 (0.25)	-0.24 (0.48)	-0.35 (0.24)	0.64 (0.18)	0.32 (0.23)
Sperm12	-0.28 (0.26)	0.17 (0.31)	-0.29 (0.29)	-0.80 (0.60)	-0.39 (0.26)	-	-
Norm24	-0.15 (0.29)	-0.06 (0.27)	-0.27 (0.31)	0.22 (0.55)	-0.40 (0.26)	0.45 (0.25)	-

* See text for trait definitions.

The moderate genetic correlations between SC (particularly SC12) and semen quality traits indicated that improved SC in Brahman and tropical Composite bulls is genetically associated with more animals producing sperm at 12 months and higher percent normal sperm at 24 months. These genetic correlations are important given that percent normal sperm is one of the better predictors of calf output by bulls in multiple sire mating groups (Holroyd *et al*. 2002).

CONCLUSION

As a result of moderate heritability and favourable genetic association with semen quality traits, yearling SC could be flagged as an indicator trait for desirable male reproductive traits in tropical breeds. High standard errors associated with most genetic parameters estimated warrant the collection of more data to test the validity of these genetic associations.

ACKNOWLEDGMENTS

We acknowledge the contribution of the Beef CRC industry collaborators, particularly those that loaned cattle for the project. We also thank Warren Sim, Tim Grant, Rob Young and staff at the research stations for their supervision of cattle and help with data collection.

REFERENCES

Burrow, H.M. (2001) *Livest. Prod. Sci.* **70**:213.

Holroyd, R.G., Doogan, V.J., De Faveri, J. Fordyce, G., McGowan, M.R., Bertram, J.D., Vankan, D.M., Fitzpatrick, L.A. Jayawardhana, G.A. and Miller, R.G. (2002) *Anim. Reprod. Sci*. **71**:67.

Johnston, D.J., Barwick, S.A., Corbet, N.J., Fordyce, G.M., Holroyd, R.G., Williams, P.J and Burrow, H.M. (2009) *Anim. Prod. Sci.* **49**:399.

Martinez-Velazquez, G., Gregory, K.E., Bennett, G.L. and Van Vleck, L.D. (2003) *J. Anim. Sci.* **81**:395.

Morris, C.A., Bennett, G.L. and Johnson, D.L. (1993) *Proc. New Zeal. Soc. An.* **53**:427.

Phillips, D.J. (2005) *Domest. Anim. Endocrin.* **28**:1.

Proc. Assoc. Advmt. Anim. Breed. Genet. 18:125-128

GENOME WIDE ASSOCIATION STUDIES FOR NET FEED INTAKE, BODY WEIGHT AND HIP HEIGHT IN BEEF CATTLE

S. Bolormaa[1] and M.E. Goddard[1,2]

CRC for Beef Genetic Technologies
[1] Department of Primary Industries, Victoria
[2] University of Melbourne

SUMMARY

We report on a genome-wide association study (GWAS) from the CRC for Beef Genetic Technologies using the 50k Illumina SNP chip. Here, we present the results for net feed intake (NFI), body weight and hip height. The aims of this analysis are to discover SNPs associated with all traits but especially NFI and to test the consistency of SNP effects across datasets and breed types using the weight and height data. The data were analysed within datasets and within breed type using a mixed model and fitting one SNP at a time. In each case the number of significant SNPs was more than expected by chance alone. However, the SNP effects for weight and height were consistent between datasets only when estimated in the same breed type (*B.inducus* x *B.taurus* composite breeds). While NFI was only measured in one dataset, we found 9 SNPs associated with NFI on BTA 3, 5, 7 and 8 with $P \leq 6.0 \times 10^{-5}$.

INTRODUCTION

In genomic selection, the estimation of breeding values is based on genetic markers. This would be particularly useful for traits that are very expensive to measure such as net feed intake (NFI). In beef cattle, some studies (Barendse *et al.* 2007; Nkrumah *et al.* 2007; Sherman *et al.* 2009) have reported associations between markers and NFI. For instance, Barendse *et al.* (2007) analysed 8,786 polymorphic SNPs in 189 Australian beef cattle, selected for high and low NFI, and detected 161 SNPs associated with NFI at P<0.01. However, the availability of 50K SNP chip from Illumina provides us with an opportunity to conduct a more powerful genome-wide association study (GWAS) for NFI.

Before genomic selection can be implemented with confidence, it is necessary to confirm in independent populations the associations that have been discovered in one population. Often such attempts at confirmation have been unsuccessful. Failure to confirm associations can be due to breed-specificity. That is, an association is only found in one breed or group of breeds. Here we use data on weight and hip height in cattle to test the consistency of associations across breed types. To do this, we have conducted GWAS in cattle from two datasets including the Beef CRC I dataset comprising *Bos taurus*, *B. indicus* and crosses between the two subspecies, and the Beef CRC II dataset with *B. indicus* and crosses between B. *indicus* and *B. taurus* cattle.

MATERIALS AND METHODS

SNP data. In total, 53,798 SNPs were genotyped. Preliminary edits were carried out at the Animal Genetics and Breeding Unit of the University of New England in Armidale, NSW. In brief, genotypes were discarded if they did not have high quality scores (>95%) and the proportion of missing genotypes did not pass defined criteria (≥20%). Sixteen thousand and eight SNPs had minor allele frequency (MAF) < 0.05 and 8,469 SNPs deviated from Hardy-Weinberg equilibrium (HWE; P<0.0001). However, these were not removed from the further analyses. Out of the initial 53,798 SNPs, 50,650 were polymorphic and included in the GWASs.

Animals and population structure. Eight hundred and fifty two steers from the Beef CRC I population had both genotype and phenotype data for NFI, weight and height (Table 1). These steers were from 7 different pure breeds. Four breeds (Angus, Murray Grey, Shorthorn and Hereford) were *Bos taurus* (BT), 1 breed (Brahman) was *Bos indicus* (BI) and two breeds (Santa Gertrudis and Belmont Red) were BT×BI synthetic breeds. Additionally, 1,456 cows from the Beef CRC II population had weight and most had height data. These cows were sourced from two different breed types, BI (Brahman) and BI×BT crosses.

Traits. CRC I steers were recorded in the feedlot for four traits: net feed intake (NFI), average daily gain (ADG), daily feed intake (DFI), and mean metabolic weight (mMWT); and prior to the feedlot period, for post weaning hip height (pwHH). The CRC II heifers were recorded for first wet season weight (w1WGT) and hip height (w1HH) (Table 1).

Table 1. Number of records (N), mean, standard deviation (SD) and estimates of heritability (h^2) and associated standard errors (SE) for all traits studied

Trait	CRC	N	BT	BI	BT×BI	ALL-mean	ALL-SD	ALL-h^2	ALL-h^2_SE
NFI	I	852	486	78	288	-2.6	1.2	0.18	0.13
DFI	I	852	486	78	288	12.3	2.1	0.16	0.13
ADG	I	852	486	78	288	1.4	0.4	0.24	0.14
mMWT	I	852	486	78	288	93.8	11.4	0.31	0.15
W1WGT	II	1456	-	590	866	301.3	44.3	0.61	0.11
pwHH	I	812	466	65	281	116.4	6.5	0.25	0.18
W1HH	II	1224	-	360	864	126.0	5.8	0.60	0.12

NFI = net feed intake (kg/day); DFI = daily feed intake (kg/day); ADG = average daily gain (kg/day); mWGT = mean metabolic weight ($kg^{0.75}$); w1WGT ="end of wet season 1" weight (kg); pwWGT = post-weaning hip height (cm); w1HH = ="end of wet season 1" hip height (kg); - = not available

Statistical analyses. The association between each SNP and each of the traits was assessed by a regression analysis using the ASReml software (Gilmour *et al.* 2002). The mixed model applied was as follows: trait ~ mean + fixed effects + SNP_i + animal + error; with animal and error fitted as random effect. Fixed effects were different for the CRC I and CRC II datasets. For CRC I data, breed, origin of herds, sex, year of measurement, season, market-weight destination and nutritional treatment were fitted as fixed effects, age deviation from group mean and SNP_i were fitted as covariate effects. Whereas for CRC II data the effects of breed, origin of herds, sire group, cohort, calving month and their first degree interactions were fitted as fixed effects. Using the same model without fitting SNP_i, estimates of heritability were calculated based on the genotyped animals (Table 1) and their ancestors.

The two datasets were analysed separately. Additionally, the three breed types within the CRC I data (BT, BI and BTxBI) were also analysed separately as well as in a joint analysis. Similarly, the two breed types (BI and BTxBI) represented in the CRC II dataset were analysed separately as well as jointly. Weight and hip height were measured in both datasets and therefore we could compare the SNP solutions from the two datasets. We calculated the number of SNPs that were significant in both datasets and for these SNPs we calculated two parameters to assess the agreement between the results: the correlation between SNP solutions in the two datasets, and the proportion of SNPs in which the solutions were in the same direction; that is, the proportion in which the same SNP allele increased the trait.

SNPs effects with high standard errors are sometimes large but the effects are poorly estimated. Therefore, the SNP effects were divided by their standard error before correlations of the SNP effects were estimated.

In a GWAS there are many thousands of significance tests performed. Therefore, we compared the number of SNPs that were significant to the number expected by chance using a False Discovery Rate (FDR) as described by Storey (2002).

RESULTS AND DISCUSSION

Heritability (h^2) estimates for NFI based on genotyped animals was a bit lower (0.18) than, but within a SE, of the one found for the complete CRC I dataset (David Johnston, pers. comm.). Overall, h^2 estimates were low to moderate for feedlot traits and moderate to high for weight and height (Table 1).

The number of significant SNPs at threshold of P<0.001 (Table 2) was 78 for weight and 75 for height in CRC I, but 156 and 134 in the larger CRC II data. Consequently, the FDRs were lower in CRC II (32% and 37%) compared with CRC I (65% and 67%).

Table 2. Number of significant SNPs and FDR at different thresholds based on all animals studied

	No. of SNPs at P				FDR (%) at P			
Trait	<0.0001	<0.001	<0.01	<0.05	<0.0001	<0.001	<0.01	<0.05
NFI	11	75	615	2826	46	67	82	89
DFI	8	76	624	2733	63	67	81	92
ADG	11	83	698	2995	46	61	72	84
mMWT	6	78	694	3141	84	65	73	80
w1WGT	29	156	935	3578	17	32	54	69
pwHH	13	75	632	2907	39	67	80	86
w1HH	26	134	833	3368	19	38	60	74

Table 3 shows the correlations of SNP effects for weight and height between CRC I and CRC II data sets across breed types, as well as the proportion of SNPs whose effects are in the same direction. Significance analysis was evaluated separately for each breed type. For instance, the number of significant SNPs for weight in CRC I BTxBI and CRC II BTxBI were 176. This is little more than expected by chance given the number tested. However, the correlation between the effects of these 176 SNPs in the two datasets and breeds was 0.43, and in 72% of SNPs the direction of the effect was the same. If these SNPs were significant in both datasets by chance alone we would expect the correlation to be zero and 50% of the SNP effects to be in the same direction. Thus, we can conclude that the two analyses are finding some SNPs with a genuine effect on weight. Similar results were found for height (Table 3).

However, when SNPs that were significant in different breed types were compared, the correlation of SNP effects dropped to near zero and the proportion in the same direction was ~50% (Table 3). The negative correlations involving CRC I Brahman cattle could be due to the very small number of animals (78) in this breed. We conclude that the direction of SNP effects is not consistent across breed types.

A total of 75 SNPs were significant (P<0.001) for NFI giving a FDR of 68%, which means that 32% of total findings (or ~24 SNPs) are expected to be true positives. The 9 most significant SNPs were detected on BTA 3, 5, 7 and 8 with $P \leq 6.0 \times 10^{-5}$. In previous studies (Barendse *et al.* 2007; Nkrumah *et al.* 2007; Sherman *et al.* 2009), NFI QTL were identified on BTA 1, 7, 18 and 19

across the 3 studies. Barendse *et al.* (2007) also detected NFI QTL (P=0.0006) on BTA 8. Sherman and *et al.* (2009) in Canadian cattle (Angus, Charolais & composites) mapped QTLs for NFI ($P=7.6\times10^{-5}$) and DMI (P<0.001) on BTA 3 and 7, respectively. However, they found no QTL on BTA 8. This may be due to several reasons including a small number of markers, differences in breed composition across studies, G×E interaction, and even a lack of power in all studies.

Table 3. Validation of SNPs for body weight and hip height (P<0.05) between CRC I and CRC II datasets across breed types

P<0.05	ALL_{CRC1} : ALL_{CRC2}	BT_{CRC1} : BI_{CRC2}	BT_{CRC1} : $BTxBI_{CRC2}$	BI_{CRC1} : BI_{CRC2}	BI_{CRC1} : $BTxBI_{CRC2}$	$BTxBI_{CRC1}$: BI_{CRC2}	$BTxBI_{CRC1}$: $BTxBI_{CRC2}$
Body weight							
No. of SNPs	242	206	209	245	203	159	176
correlation	0.26	-0.10	-0.04	-0.21	0.03	0.07	0.43
% of same dir	61	46	48	39	50	53	72
Hip height							
No. of SNPs	208	161	149	213	226	197	198
correlation	0.12	0.00	0.02	0.09	-0.17	0.10	0.36
% of same dir	53	47	51	55	43	54	66

No. = number of SNPs; correlation = correlation of corrected SNPs; and % of same dir = proportion of the direction of the effect, which was the same; ALL = all CRC I or CRCI I animals; BT = *Bos taurus*; BI = *Bos indicus*; BTxBI = crosses of *Bos taurus* and *Bos indicus*

CONCLUSIONS

For all traits, we found more significant SNPs than expected by chance. FDR was lower in the larger population and when using a more stringent significance level. Comparing CRC I and CRC II results for body weights and hip height, we found agreement between significant SNPs effects only when the same breed type was used in both studies. This implies that the power to detect SNPs when all breed types are analysed together is reduced because the effects in different breed types are different. The SNPs that were significant for NFI could not be tested in CRC II cattle because these heifers were not measured for NFI. However, we will test them in other cattle populations. To achieve greater power in GWAS we have agreed to collaborate with scientists in North America so that a larger sample can be included in one analysis.

ACKNOWLEDGEMENTS

We are grateful for comments from Toni Reverter, Cynthia Bottema and Bill Barendse.

REFERENCES

Barendse, W., Reverter, A., Bunch, R.J., Harrison, B.E., Barris, W. and Thomas, M.B., (2007) *Genetics* **176**:1893.

Gilmour, A.R., Gogel, B.J., Gullis, B.R., Welham, S.J. and Thompson, R. (2002) *VSN International Ltd*. Hemel Hempstead, U.K.

Nkrumah, J.D., Sherman, E.L., LI, C., Marques, E., Crews, D.H., Bartusiak, J.R., Murdoch, B., Wang, Z., Basarab, J.A. and Moore, S.S. (2007) *Journal of Animal Science* **85**:3170.

Sherman, E.L., Nkrumah, J.D., Bartusiak, C.L.R., Murdoch, B., Moore, S.S. (2009) *Journal of Animal Science* **87**:37.

Storey, J.D. (2002) *Journal of the Royal Statistical Society, Series B* **64**:479.

Proc. Assoc. Advmt. Anim. Breed. Genet. 18:129-132

VALIDATION AND ESTIMATION OF ADDITIVE GENETIC VARIATION ASSOCIATED WITH DNA TESTS FOR QUANTITATIVE BEEF CATTLE TRAITS

A. L. Van Eenennaam[1], R. M. Thallman[2], R. L. Quaas[3], K. Hanford[4], and E. J. Pollak[3]

[1] Department of Animal Science, University of California, Davis, CA, USA
[2] US Meat Animal Research Center, Clay Center, NE, USA
[3] Department of Animal Science, Cornell University, Ithaca, NY, USA
[4] Departments of Statistics and Animal Science, University of Nebraska, Lincoln, NE, USA

SUMMARY

The U.S. National Beef Cattle Evaluation Consortium (NBCEC) has been involved in the validation of commercial DNA tests for quantitative beef quality traits since their first appearance on the U.S. market in the early 2000s. This paper discusses pre-validation, analysis, and reporting issues based on our validation experiences. Estimates of DNA test performance (e.g. proportion of genetic variation accounted for by a DNA test panel) in representative populations will be required for incorporation of DNA data into the existing genetic evaluation infrastructure. Such incorporation is appealing as it presents results in an EBV format that is familiar to producers, and eliminates the choice that is implicit when EBVs and marker information are published in tandem.

INTRODUCTION

Prior to moving genetic markers from discovery populations to commercialization, it is important to validate their purported effects on the trait of interest in different breeds and environments, and assess them for correlated responses in associated traits (Barendse 2005). The biggest challenge to achieving this objective is the paucity of cattle populations with sufficient phenotypic data to assess the association between various traits and newly discovered genetic markers, and this makes it difficult and expensive to do large-scale field validations. Results from such validation studies to date have not been widely published (Burrow and Bindon 2005). The validation of panels of DNA markers that are proposed to be used commercially is not simply a repeat of the discovery process, but rather a critical activity to test the strength of support for the testing companies published claims based on independent data.

The NBCEC originally used the term "having validated" to mean finding a significant association "between genetic tests and traits as claimed by the commercial genotyping company based on phenotypes and genotypes derived from reference cattle populations" (Van Eenennaam *et al.* 2007). This process sometimes revealed that tests did not perform as expected, and in certain cases companies chose to withdraw their plans to market those tests.

During the past decade, the DNA testing industry matured from single gene tests to panels involving an ever-increasing number of markers with purported effects on multiple traits, and/or in specific cattle subpopulations. The NBCEC and DNA testing companies have struggled to find appropriately-phenotyped populations that were not involved in the discovery process for validation studies. Additionally, results from different validation populations genotyped with the same SNP panel were often inconsistent with respect to the significance of the association between the test and the trait(s), and sometimes even with respect to the direction of the association. This complicated the interpretation of validation results, and created confusion as to whether "validation" meant a test "worked" (i.e. was significantly associated with the trait) in one or more of the test populations, or had simply been tested by an independent third party. This exposed the process to marketing zeal and left producers somewhat stymied because the data generated did not help to inform decisions about the value proposition associated with investing in specific DNA tests. With the imminent commercialization of a multiplicity of products derived from high density

SNP assays, it seems an opportune time to address some of these concerns. We believe that the validation process needs to evolve from simply reporting the finding of a significant association between DNA test results and the trait of interest, towards an independent calibration approach that estimates the parameters that will be required to facilitate the incorporation of DNA test-based predictions of genetic merit into national genetic evaluation schema. Additionally, results have to be disseminated and interpreted in such a way as to provide industry with the information they require to make the best use of DNA information. Issues associated with the process of validation can be divided into three categories; pre-validation issues, analytical issues and issues associated with presenting results to scientific and industry audiences.

PRE-VALIDATION

DNA-test developers will typically use a "discovery" dataset(s) where a large number of SNP have been assayed on phenotyped animals to develop their test (Allan and Smith, 2008). Genotypes and phenotypes from these discovery populations will be used to develop "molecular breeding value" (MBV) prediction equations by summing the individual SNP additive effects of those loci that show the strongest association with the trait of interest. This DNA test is then the focus of a validation study where a representative sample of animals is genotyped for the markers included in the panel, and the resultant MBVs are compared to phenotypes to assess the accuracy of the test (Goddard and Hayes, 2007). It is essential that discovery populations not be used in validation studies (Barendse, 2005). Prediction equations will perform best in discovery populations in which the SNP associations were discovered and/or populations in which the SNP effects were estimated. Upon requesting a validation, it is therefore crucial that developers fully document which version of a SNP panel and prediction equation they plan to commercialize, and the populations that were used for discovery and training.

As a result of the global investment in SNP genotyping, it is likely that a large number of marker panels associated with various traits will be commercialized in the next couple of years. In genomic studies where many loci are tested against many traits, false positives are inevitable. It is important to delineate test claims and proposed target populations prior to the commencement of any validation studies. Otherwise there is an obvious temptation to use the validation process for discovery. Requiring the disclosure of results from all populations included in validation studies is also important to guard against the understandable temptation to go public with only favourable results. It will be a challenge to keep track of which SNPs are included in the various commercial offerings, especially as they mature and additional SNPs are added to previously tested panels. Therefore some system of nomenclature such as version numbers of tests is essential. While it might be assumed that adding additional SNPs to a panel will improve the accuracy of tests, it is not clear whether new panels should be tested in the same validation populations as the original panel to demonstrate the magnitude of this improvement, or in new populations. Equally unclear is what course of action should be taken if the new panel proves inferior to the old panel.

VALIDATION ANALYSIS

Initial validations performed by the NBCEC looked at individual marker effects to conclude whether the genotype at each individual locus, and the combined effect of all loci in the test, were significantly associated with trait phenotype(s) as claimed by the commercial genotyping companies (Van Eenennaam *et al.* 2007). A similar approach was taken by the "SmartGene for Beef" project in Australia, except in that case the amount of variation accounted for by each DNA marker in each cattle population was estimated using these gene frequencies and gene effects (http://agbu.une.edu.au/SmartGene%20Report11.pdf, accessed 10/6/09). As marker panels grew in size and intellectual property concerns regarding disclosure of the specific marker loci involved in a genetic test emerged, validation moved from testing the effect of individual loci towards testing a

single marker score, or molecular breeding value (MBV). The validation data analysis moved to a determination of whether the regression of phenotype on marker score for a single trait model (in which the marker score was a covariate) differed from zero. While reassuring to see this regression coefficient be non-zero, the significance of this result did not provide useful information in terms of decisions related to the value proposition of the test. A test that has a significant association with the trait of interest may nonetheless explain only a minor proportion of the genetic variance.

The proportion of variation accounted for by a DNA test seems to currently be the best metric available with which to quantitatively evaluate the merit of commercial products. The (co)variance estimates from which the proportion of additive genetic variation accounted for by a DNA test are computed will be requisite for the incorporation of DNA information into national beef cattle genetic evaluations as described by Kachman (2008b). This approach uses a two trait model with the marker score and phenotype included as correlated traits. A theoretically robust estimator of this statistic is the REML estimate of the genetic correlation squared (R_g^2) in a bivariate animal model for the target trait and the MBV, as the second trait (Thallman *et al.* 2009). This estimator has the advantage of producing estimates within the parameter space, and also should be computationally feasible given the size of typical validation data sets.

Simulation studies using this approach (Thallman *et al.* 2009) showed that this R_g^2 estimator tended to be closer to the simulated value of the proportion of variation accounted for by a DNA test than other statistical estimators. Australian researchers recently reported results from a marker panel (Pfizer Animal Genetics 56 SNP panel) evaluation run on four different populations using this approach. These data showed that the proportion of genetic variation accounted for by the molecular value predictions (i.e. MBVs) ranged from ~ 0 - 0.3 depending upon the population and the target trait. However, Kachman (2008b) warned that such estimates may be inaccurate in small data sets (< 1,000 records), and that this error will be exacerbated in traits with low heritability.

Unlike the dairy industry which has the advantage of large, single-breed phenotyped populations for marker discovery and validation (Van Raden *et al.* 2009), the data sets that have typically been used for validation in beef cattle are far from ideal for estimating additive variances and covariances. For example, in the U.S. the NCBA Carcass Merit Project used in the original validations, most breeds had between 400 and 600 progeny represented with records but these were progeny of relatively few sires (≤10). There is a clear need for large, well-organized, thoroughly-phenotyped populations for estimating genetic (co)variances. The development of such populations may require collaborative efforts, and the expenses involved are likely beyond the resources of any single company, or even a single country.

REPORTING

Despite analyses and validation work in both Australia and the US, it is not clear if the data that is currently being reported is providing users with the type of information they need to make informed decisions about the use of a particular test. Validation teams have objectively presented their findings on websites with the optimistic view that "decisions very much depend on the individual business' attitude to risk and can only be made effectively by the individual business." While this is undoubtedly true, most producers are unlikely to have had sufficient training in quantitative genetics to correctly interpret the results and get to the point of making such decisions based on risk considerations. Many producers are already wary of this historically-oversold technology based on past experience (Barendse *et al.* 2005; Casas *et al.* 2005). It is not evident how reporting findings that a test explains a proportion ranging from 0 to 0.15 of the additive genetic variation associated with the target trait, has a regression coefficient of 0.26 (±0.3), and a *p* value of 0.001 provides information that helps in the decision-making process. The accuracy of a DNA test at predicting the true genetic merit of an animal is primarily driven by the amount of additive genetic variation accounted for by the DNA test. Current estimates for this correlation

and the proportion of the variation accounted for by existing tests (the square of that correlation) are generally low with the exception of DNA tests for tenderness where available estimates include 0.25 (Allan and Smith, 2008) and 0.016-0.299 from an Australian evaluation of the Pfizer Animal Genetics 56 SNP panel, (http://www.beefcrc.com.au/Aus-Beef-DNA-results, accessed 10/6/09). Over time it is envisioned that genetic tests will have markers associated with the majority of important genes influencing a trait and the marker effects will have been assessed in large enough training populations to provide for accurate SNP effect estimates meaning genotyping results will be highly predictive of the true breeding value (BV) of an animal.

In the interim however, from the user's perspective, perhaps the most useful information that could be provided is how much a DNA test improves the accuracy of expected breeding values (EBV). That is, how much improvement in the accuracy associated with an EBV can be expected from adding DNA test information. Publishing traditional EBVs and marker information separately, as is currently the case, is confusing and can lead to incorrect selection decisions when marker scores predict only a small proportion of the genetic variance (Crews *et al.* 2008). Developing an approach to calculate marker-assisted EBVs would seem to be the next logical step.

Selection index methodology has been applied to combine marker scores and polygenic EBVs using a linear index weighted on the accuracy of traditional BV and the proportion of genetic variance attributable to the marker score (Amer 2007; Crews *et al.* 2008). Kachman (2008a) simulated the effect of DNA marker scores on the accuracy of tenderness expected progeny difference (EPD). Assuming a heritability of 0.4 and a 0.45 genetic correlation between the marker score and shear force, he found that the beef improvement association accuracy of the sire tenderness EPD increased from 0.27 to 0.31 when DNA marker information was combined with information from 10-phenotyped progeny. Expressing the effect of marker information on accuracy is appealing because it makes use of the existing genetic evaluation infrastructure, and presents a metric that is familiar to producers.

CONCLUSIONS

The validation process and analyses have evolved as the DNA testing industry has matured from single gene tests to panels involving an ever-increasing number of markers. With products derived from high density SNP assays on the horizon, it seems an opportune time to reassess the process and consider how it can best be used to provide the estimates of the genetic (co)variance parameters that will be required to facilitate the incorporation of information into national genetic evaluation systems. Additionally, careful consideration must be given to industry dissemination of independent validation studies.

REFERENCES

Allan, M.F., and Smith, T.P.L. (2008) *Meat Sci*. **80**:79.
Amer, P.R. (2007) *Proc. Assoc. Advmt. Anim. Breed. Genet*. **17**:479.
Barendse, W. (2005) *Aust. J. Exp. Agric.* **45**:831.
Barendse, W., Bunch, R.J., and Harrison, B.E. (2005) *Anim. Genet*. **36**:86.
Burrow, H.M., and Bindon, B.M. (2005) *Aust. J. Exp. Agric*. **45**:941.
Casas, E., White, S. N., Riley, D. G. , et al.. (2005*) J. Anim. Sci.* **83**:13.
Crews, D. H., Jr., Moore, S. S., and Enns, R. M. (2008) *Proc. 40th Beef Improv. Fed.* **40**:44.
Goddard, M. E. and Hayes, B. J. (2007) *J. Anim. Breed. Genet*. **124**:323
Thallman, R.M., Hanford, K.J., Quaas, R.L., et al. (2009) *Proc. 41st Beef Improv. Fed.* **41**:184.
Kachman, S. D. (2008a). *40th Beef Improv. Fed.* http://www.bifconference.com; accessed 10/6/09
Kachman, S. D. (2008b). *CD-ROM 9th Genetic Prediction Workshop, Beef Improv. Fed.*
Van Eenennaam, A.L, J. Li, J., Thallman, R.M., Quaas, R.L., et al. (2007) *J. Anim. Sci.* **85**:891.
VanRaden, P. M., Van Tassell, C. P., Wiggans, G. R., et al. (2009) *J. Dairy Sci*. **92**:16-24.

Proc. Assoc. Advmt. Anim. Breed. Genet. 18:133-142

GENOTYPE BY ENVIRONMENT INTERACTIONS IN INTERNATIONAL GENETIC EVALUATIONS OF DAIRY BULLS

J. H. Jakobsen, J. W. Dürr, H. Jorjani, F. Forabosco, A. Loberg and J. Philipsson

Interbull Centre, Swedish University of Agricultural Science, Box 7023, 75007 Uppsala, Sweden

SUMMARY

For international genetic dairy evaluations different environments have been defined as country borders and the genetic component has been progeny tested bulls. The correlations estimated between the countries are affected by differences in trait definition, differences in genetic evaluation model and genotype-environment interactions. Also, genetic correlations estimated between countries can be affected by the amount of exchange of semen expressed as genetic links between the countries. Seven different trait groups and of these 38 different sub-traits are currently evaluated internationally for six dairy breed types. In order to describe the complex of genotype by environment interactions in international genetic evaluations this paper will deal with three of the sub-traits: protein yield, longevity, female fertility (lactating cows' ability to conceive) for the Holstein breed with respect to average correlations, genetic links and model differences between countries that currently subscribe to the international genetic evaluation services for these traits. These three sub-traits were picked to represent 1) *Yield*: a long history of data collection, trait harmonization and genetic evaluation; 2) *Longevity*: a change in trait definition over time due to change in main culling reasons; 3) *Female fertility*: a complex and novel trait in international genetic evaluations. Based on the traits considered in this study it can be indicated that correlations among countries are highest for a trait with a long history of harmonization, and lower for a trait with change in trait definition over time and for a novel trait.

Most countries currently participating in the evaluations have certain environmental similarities. Countries without a national genetic evaluation are buying semen from these countries through semen vendors but breeding values of these bulls are based on progeny tests in totally different environments. Different studies have been looking at inclusion of descriptive variables in order for these countries to get a sire ranking probably more suitable for their environment. The results look promising but are not yet implemented for routine use.

INTRODUCTION

Increased worldwide trade in dairy bull semen in the last 2-3 decades has resulted in a demand for international comparison of bulls from other countries on each country's own scale. For this reason, accurate genetic evaluation of bulls on an international basis requires correct sire rankings for all environments. Differences between sires in the genetic ability of their daughters to perform in different environments may result from true interactions between genotype and environment (GxE). However, performance in neighboring countries with similar production environments cannot be treated as genetically identical traits, due to possible genotype x environment interactions between countries caused by different definitions of the traits or different evaluation models (Fikse *et al.* 2003).

National evaluation bulls' proofs from all member countries are processed together by **Interbull** (International Bull Evaluation Service) and international breeding values are returned to each country on its own scale. The Interbull evaluations started in 1994 with four countries, one breed and three sub-traits. The Interbull evaluation currently (2009) includes 28 countries, 6 breeds and 38 different sub-traits. The numbers of bulls getting an international proof for production traits (the trait group with the largest participation) were in January 2009: Brown Swiss (BSW) 7,922

bulls; Guernsey (GUE) 944; Holstein (HOL) 105,904; Jersey (JER) 8,692; Red Dairy Cattle (RDC) 11,643; and Simmental (SIM) 22,275.

Globally, dairy selection was for several decades almost solely on production traits. However, due to the negative correlations between production and many functional traits, one-sided selection for production has proven to have a detrimental effect on functional traits (Philipsson and Lindhé, 2003). Functional traits were therefore one after the other included in the national breeding objectives of many countries. This resulted in a demand to have the same traits evaluated internationally as were included in many national genetic evaluations.

The first objective of this paper is to present evidence for GxE for traits of three different types: protein, longevity and female fertility, as examples of traits currently evaluated internationally and the practical implications for the member countries and the industry when ranking the bulls. The second objective is to indicate a method by which countries currently not participating in the international genetic evaluation could benefit from utilizing the information on the different country scales.

COUNTRIES CURRENTLY PARTICIPATING IN THE INTERNATIONAL GENETIC EVALUATION

The aim of international genetic evaluations is to compute international breeding values for all bulls with data sent by countries participating in the evaluations and to deliver back predicted breeding values of bulls from all the countries to each individual country scale.

The process of international genetic evaluations can be divided into two steps. *Step 1* deregression (Jairath *et al*, 1998) and estimation of genetic correlations among countries (Klei and Weigel 1998) and *step 2* deregression, sire variance estimation (Sullivan 1999) and prediction of international breeding values. Steps 1 + 2 are done twice per year during the "Test Evaluations" where new countries or new traits or new breeds can join the evaluations, while step 2 only is performed for the three yearly "Routine Evaluations".

Multi-trait Across Country Evaluation (MACE) software (Klei, 1998; Klei and Weigel 1998) is used to obtain the genetic correlations. Genetic correlations among countries are post-processed using prior information about previously estimated correlations, national production systems and trait definitions. These post-processed correlations are used for prediction of breeding values across countries. The average correlations per country for protein yield, longevity and female fertility are shown in Tables 1, 2 and 3, respectively, for the Holstein breed. The tables also show how far back each country includes data for these traits in their national genetic evaluation, the trait definition, their national genetic evaluation model, heritability of the trait as well as average number of common bulls. Apart from the average number of common bulls and average genetic correlations all the information in these three tables is provided by the Interbull member countries in the so called GE-forms where they describe the national genetic evaluation system for each specific trait group (Interbull 2009).

Tables 1 to 3 show that average correlations are the highest for production traits, with a mean of 0.84, and medium for both female fertility (0.63) and longevity (0.66). The history of milk recording is much longer than the collection of information about culling and fertility, and production traits have therefore had a greater opportunity to reach a higher degree of trait and model harmonization than the other traits.

Longevity is a trait that is very difficult to harmonize partly because the time of culling may be determined by agricultural policy or product price changes in a country. This may easily cause the main culling reasons to change over time. A survey sent to the Interbull community (Forabosco *et al.* 2009) on the main culling reasons showed that poor female fertility is the main reason for culling today in most of the member countries. However, some countries do not collect the information about culling reasons at the national level. It can only be speculated what the

responses would have been if the survey had been conducted 15-20 years ago, when more cows may have been culled due to low production. Tarres *et al.* (2007) estimated genetic correlations between longevity in Germany and longevity in France using different left censoring for data inclusion. They found that correlations between these two countries for longevity changed from 0.694 to 0.731 to 0.845 when 1985, 1990 and 1995, respectively,were used for left censoring of data. This indicates a change of the longevity trait definition over time for the countries included in their analysis. A change in trait definition over time for productive life in the US was confirmed by Tsuruta *et al.* (2004). Some countries have productive longevity as their official trait while others have functional longevity. The Netherlands changed their trait definition from functional longevity to productive life in 2008 and noticed an average decrease in genetic correlations to other countries of -0.18 (Van der Linde and de Jong 2007), where the decrease was largest to countries having functional longevity. This indicates that a change in trait definition can change correlations to other countries and that the similarity in trait definition among countries will affect the correlation.

Currently, International genetic evaluations for female fertility are offered for five different sub-trait groups: *T1:* Maiden heifers' ability to conceive; *T2*: Lactating cows' ability to recycle after calving; *T3*: Lactating cows' ability to conceive expressed as a rate trait; *T4*: Lactating cows' ability to conceive expressed as an interval trait (all countries are expected to submit a trait for this trait group and can even submit a rate trait for this sub-trait); *T5*: Lactating cows' measurements of the interval traits calving-conception. In this study we looked at sub-trait group T4 because it includes all countries, but the differences in trait definition impacted the average genetic correlation reported for female fertility. Computing the average correlation for a pure rate trait as T3 gave a correlation of 0.71 and slightly higher than the average for the T4 sub-trait group.

National evaluation models used to analyze protein, longevity and female fertility are also shown in Tables 1 to 3. Emanuelson *et al.* (1999) showed that changes in national genetic evaluation models can impact correlations to other countries in both directions depending on the nature of the model change.

The size of the correlations among countries affects re-ranking of bulls on other country scales. If the genetic correlations are unity the ranking of the bulls will be the same on all country scales, but as soon as the correlation is smaller than one some re-ranking will occur. In general, the smaller the correlation, the more re-ranking of bulls. Figure 1 illustrates the number of bulls appearing on a top 100 list for protein and longevity in any participating country for each of the evaluations from February 2006 to January 2009. If the correlation had been one, the same 100 bulls would have been selected for any country, but as the correlations are less than unity the figure shows that 362 and 655 bulls appeared on a top 100 list for one of the countries participating in the January 2009 evaluation for protein and longevity, respectively.

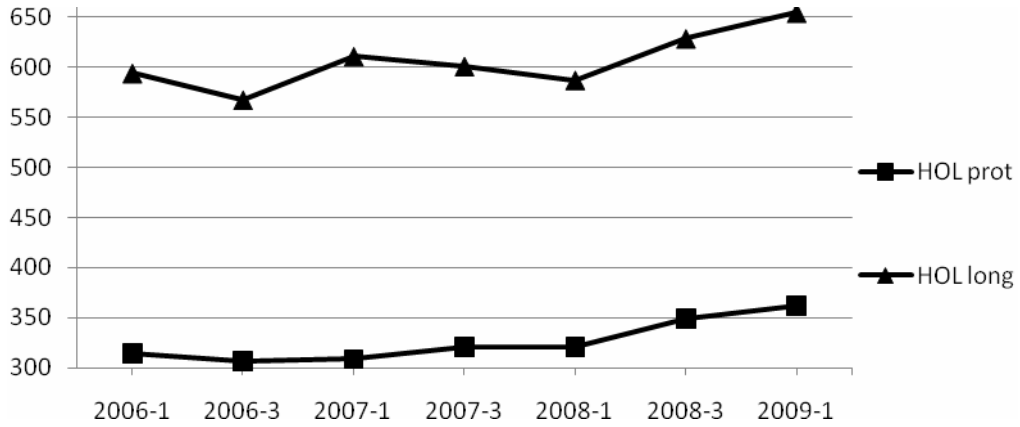


Figure 1. Trend of the total number of top 100 Holstein (HOL) bulls for protein (prot) and longevity (long) traits when r_g <1.

Average number of common bulls for each of the countries participating in the evaluations for production, longevity and female fertility are also shown in Tables 1 to 3. Good links to other countries are required for estimation of reliable correlations. Semen of the Holstein breed has been distributed over large parts of the world and for most populations links are rather strong. The average number of common bulls ranged from 31 (Israel) to 472 (United States) for protein, from 8 (Denmark – Red HOL) to 476 (United States) for longevity and from 44 (Israel) to 762 (United States) for female fertility. Average number of common bulls for individual countries cannot be compared across traits as the value is influenced by which countries are participating in the evaluation, how much semen exchange there has been among these countries, how well the international animal identification is harmonized and also the novelty of the trait in respect of national evaluations. Large populations have by definition a larger chance of having large number of common bulls than small populations. However, the issue is also what proportion of tested bulls are used in common with other countries.

COUNTRIES CURRENTLY NOT PARTICIPATING IN THE INTERNATIONAL GENETIC EVALUATION

Many countries around the world may for political, economical or infrastructural reasons not have a national genetic evaluation in place and neither an organized animal identification nor a milk recording system. With that in mind there is no opportunity of joining the international genetic evaluation for dairy bulls in the near future. However, semen from other countries is sold in these countries but bulls´ breeding values cannot be presented on a country scale of the purchasing country. Bulls may have been progeny tested in a temperate environment with one sort of management system (e.g. lots of concentrate feeding and/or low fibre grass diet) and sold to a tropical environment with a different management system (e.g. very limited concentrate feeding and/or high fibre grass diet) and daughter performance of these bulls may be very different in these environments. This is confirmed by, among others, Ojango and Pollot (2002) who found a genetic correlation between breeding values of bulls tested in both the UK and Kenya of 0.49.

Also, environmental factors, such as temperature and humidity, may differ a lot between the different environments. G x E exists within large countries such as the United States and therefore Weigel *et al.* (2001) and Bohmanova *et al.* (2008) suggested including effects of temperature and humidity in the US national genetic evaluation model for milk production in order to account for the heat stress / heat tolerance factor in the model.

It would be very desirable for the importing non-Interbull member countries to have a tool that can be used when selecting bulls that would perform the best under their environmental conditions.

Mark *et al.* (2006) and Torsell *et al.* (2007) worked on procedures to produce a good estimate of a correlation between the countries currently participating in the Interbull evaluations and non-participating countries. Torsell *et al.* (2007) applied the currently used REML procedure (Sigurdsson *et al.* 1996) to several different scenarios without inclusion of data from a pilot country and compared the results of the different scenarios to the results obtained using data from a pilot country. Among the scenarios tested was the use of different fixed correlations between the country with no data and all other countries but also inclusion of different environmental descriptors as climatic variables, production system indicators and information about national genetic evaluation systems in their model. They found that the scenario including environmental descriptors gave a better result than the scenarios using fixed correlations.

These procedures are not yet implemented for routine use, but show the opportunities for countries not participating in Interbull evaluations to benefit from the international genetic evaluation system. In order to get that system to be operational it is necessary to include relevant environmental descriptors for all countries participating in Interbull evaluations as well as for those countries wanting to participate without own data. Selection of breeding animals using genomic selection will have an impact on selection of the dairy breeds also in these countries, yet the G x E among countries need to be considered as apparently different environments require different animals. The issue is to get as good estimates as possible of the genetic correlations, even when national evaluation systems are lacking.

CONCLUSION

Genetic correlations less than one exist among the countries currently enrolled in the Interbull evaluation and are affected by trait definition, national genetic evaluation model, and G x E interactions. Three of the thirty-eight sub-traits currently evaluated were described in this paper to represent different dimensions of the genetic by environment interaction complex for international genetic evaluations. Currently, environments are separated by national country borders, due to the fact that most genetic evaluation units are national, but could theoretically be defined by other environmental descriptors. Many non-Interbull member countries import semen from bulls tested in a different environment and daughters of these bulls seem to perform differently in the local environment indicating G x E interactions. These countries may not be able to participate in the International evaluation with their own data but may be able to convert breeding values from other countries to their own scale including predefined environmental descriptors as prior information.

REFERENCES

Bohmanova, J., Misztal, I., Tsuruta, S., Norman, H.D., and Lawlor, T.J. (2008*) J. Dairy Sci.* **91**:840.

Emanuelson, U., Fikse, W.F., and Banos, G. (1999) *Interbull Bull.* **20**:114.

Fikse, W.F., R. Rekaya, and Weigel, K. A. (2003) *J. Dairy Sci.* **86**:1821.

Forabosco, F., Jakobsen, J.H., and Fikse, W.F. (2009). *J. Dairy Sci.* **92**: 2338

Interbull (2009). Description of national genetic evaluation systems for dairy cattle traits as applied in different Interbull member countries. www.interbull.org/national_ges_info2

Jairath, L., Dekkers, J.M.C., Schaeffer, L.R., Liu, Z., Burnside, E.B., and Kolstad, B. (1998). *J. Dairy Sci.* **81**:550.
Klei, B. (1998) *Interbull Bull.* **17**:3.
Klei, B. and Weigel, K. A. (1998) *Interbull Bull.* **17**:8.
Mark, T., Sullivan, P.G., Fikse, W.F., and VanRaden, P.M. (2006). *Interbull Bull* **35**: 72.
Ojango, J.M.K. and Pollot, G.E. (2002). *Livest Prod Sci.* **74**: 1.
Philipsson, J. and Lindhé, B. (2003). *Livest Prod Sci.* **83**: 99.
Sigurdsson, A., Banos, G., and Philipsson, J. (1996). Acta Agric. Scand., Sect. A, Anim. Sci. **46**: 129.
Sullivan, P. (1999) *Interbull Bull.* **22**:146.
Tarres, J., Liu, Z., Reinhardt, F., Reents, R., and Ducrocq, V. (2007). *Interbull Bull.* **37**: 98.
Torsell, A., Jorjani, H., and Fikse, W.F. (2007) *Interbull Bull.* **37**: 111.
Tsuruta, S., Misztal, I., and Lawlor, T.J. (2004). *J Dairy Sci*. **87**: 1457.
Van der Linde, C. and de Jong, G. (2007). *Interbull Bull* **37**: 203.
Weigel, K. A., Rekaya, R., Zwald, N. R., and Fikse, W. F. (2001). *J. Dairy Sci*. **84**:2789.
Zwald, N. R., Weigel, K. A., Fikse., W. F., and Rekaya, R. (2003). *J. Dairy Sci*. **86**:1009.

Table 1. Country (COU), first year of data inclusion for current national genetic evaluation, trait definition, genetic evaluation model, heritability (h^2), average number of common bulls (μ_{CB}) and average post-processed correlations to other countries (μ_{RG}) for international genetic evaluation for protein yield for the Holstein breed in April 2009.

COU[1]	Data since	Trait definition[2]	Model[3]	h^2	μ_{CB}	μ_{RG}
AUS	1975	305 d yield; complete and extended lactations	ST RP AM	0.250	246	0.77
BEL	1973	Yield within a 24 hour test period; EBV: average 305 day yield across Lact. 1-3	MT ML RR TD AM	0.410	164	0.84
CAN	1988	Yield within a 24 hour test period; proofs are average yields across Lact. 1-3	MT ML RR TD AM	0.370	308	0.86
CHE	1985	Test day yield. Lact 1-3; records < 330 DIM	MT ML RR TD AM	0.320	134	0.86
CHR	1987	Test day yield. Lact 1-3; records 5-330 DIM	MT ML RR TD AM	0.360	123	0.86
CZE	1995	Yield within a 24 hour test period; Lact 1-3.	ST ML RR TD AM	0.370	295	0.84
DEU	1990	Yield within a 24 hour daily basis; Lact. 1-3. Records 5-330 DIM	ST ML RR TD AM	0.480	448	0.84
DFS	1990(DNK),1988(FIN), 1995(SWE)	Test day yield (DNK+FIN), lactation records (SWE), Lact 1-3 (DNK+SWE), all lactations (FIN)	MT ML RR TD AM	0.390	293	0.86
ESP	1986	305 d yield; complete and extended lactations; Lact 1-5	ST RP AM	0.280	230	0.85
EST	1994	Test day yields 5-365 DIM. Lact 1-3.	ST ML RR TD AM	0.480	41	0.86
FRA	1980	305 d yield; complete and extended lactations; Lact 1-3	ST RP AM	0.300	253	0.86
FRR	1980	305 d yield; complete and extended lactations; Lact 1-3	ST RP AM	0.300	10	0.86
GBR	1975	Test day yields. Lact 1-5	ST ML RR TD AM	0.510	359	0.85
HUN	1985	305 day yield. Lact 1-3	ST RP AM	0.200	232	0.85
IRL	1970	305 day yield. Lact 1-5	ST RP AM	0.350	174	0.78
ISR	1985	305 day yield; complete and extended lactations	ST ML AM	0.410	31	0.78
ITA	1985	Test day yield within 24 hour period. Lact 1-3; records 5-305 DIM	MT ML RR TD AM	0.300	298	0.85
JPN	1985	305 day yield; Lact 1-5.	ST RP AM	0.270	113	0.87
LVA	1996	Yield within a 24 hour test period. Lact 1-3.	ST ML RR TD AM	0.480	36	0.85
NLD	1990(NLD), 1981(FLA), 1995(LUX)	Yield within a 24 hour test period. Lact 1-3. Records 5-335 DIM.	ST ML RR TD AM	0.500	446	0.85

Tabel 1 *continued*. Country (COU), first year of data inclusion for current national genetic evaluation, trait definition, genetic evaluation model, heritability (h^2), average number of common bulls (μ_{CB}) and average post-processed correlations to other countries (μ_{RG}) for international genetic evaluation for protein yield for the Holstein breed in April 2009.

COU[1]	Data since	Trait definition[2]	Model[3]	h^2	μ_{CB}	μ_{RG}
NZL	1986	Test day records 3-270 DIM. Age groups 2, 3, 4, and 5-7 years	ST ML RR TD AM	0.310	234	0.76
POL	1995	Test day records 5-305 DIM. Lact 1-3.	ST ML RR TD AM	0.290	216	0.84
SVK	1992	Yield within a 24 hour test period. Records 5-365 DIM. Lact 1-3.	ST RR TD AM	0.300	102	0.85
SVN	1997	Test day records 6-305 DIM. Lact. 1-5.	ST RP FR TD AM	0.210	38	0.85
USA	1960	305 day yield; complete and extended lactations. Lact 1-5	ST RP AM	0.300	472	0.86
ZAF	1988	Yield within 24 hour test period. Records: 5-305 DIM. Lact 1-3.	MT RP FR TD AM	0.140	174	0.84

1. AUS=Australia; BEL=Belgium, CAN=Canada, CHE=Switzerland (black & white), CHR=Switzerland (red&white), CZE=Czech Republic, DEU=Germany, DFS=Denmark+Finland+Sweden, ESP=Spain, EST=Estonia, FRA=France, FRR=France (red&white),GBR=United Kingdom, HUN=Hungary, IRE=Ireland, ISR=Israel, ITA=Italy, JPN=Japan, LVA=Latvia, NLD=The Netherlands, NZL=New Zealand, POL=Poland, SVK=Slovak Republic, SVN=Slovenia, USA=United States, ZAF=South Africa

3. DIM=Days in milk 3) ST=Single trait, MT=multiple trait, ML=multiple lactations, RR=random regressions, RP=repeatability, FR=fixed regression, AM=animal model, TD=test-day model

Table 2. Country(COU), first year of data inclusion for current national genetic evaluation, trait definition, genetic evaluation model, heritability (h^2), average number of common bulls (μ_{CB}), and average post-processed correlations to other countries (μ_{RG}) for longevity evaluation for the Holstein breed in April 2009.

COU[1]	Data since	Trait definition	Model[2]	h^2	μ_{CB}	μ_{RG}
AUS	1975	Probability of surviving from one year to the next	AM RP	0.025	309	0.61
BEL	1973	Survival over successive lactations	AM RR	0.106	157	0.66
CAN	1980	Survival in the first three lactations	AM MT	0.098	322	0.76
CHE	1980	Productive life span of the cow in months	SM SA	0.077	154	0.75
CHR	1984	Productive life span of the cow in days	SM SA	0.110	88	0.67
DEU	1985	Functional herd life in days	SM SA	0.166	335	0.75
HUN	1988	Productive life span of the cow in days	SM SA	0.050	232	0.51
DNK	1984	Productive life span of the cow	SM SA	0.117	169	0.74
DNR	1984	Productive life span of the cow	SM SA	0.117	8	0.69
ESP	1986	Productive life span of the cow in days	SM SA	0.115	240	0.66
FIN	1988	Stayability from first calving	SM SA	0.120	49	0.65
FRA	1988	Productive life of the cow in days	SM SA	0.108	260	0.66
NLD	1988	Productive life span of the cow in days	SM SA	0.120	377	0.61
NZL	1987	Survival from first to fifth lactation	AM MT	0.055	256	0.48
GBR	1986	Lifespan	AM MT	0.064	442	0.75
IRL	1980	Survival to the next lactation (lactation 1 to 4)	AM MT	0.016	221	0.71
ISR	1985	Days from first calving to 2922d.	AM ST	0.110	34	0.57
ITA	1980	Productive life span of the cow in days	SM SA	0.097	307	0.61
SWE	1984	Survival rate at second calving	SM MT	0.080	230	0.65
USA	1960	Productive life	AM ST	0.080	476	0.76

1) AUS=Australia; BEL=Belgium, CAN=Canada, CHE=Switzerland (black & white), CHR=Switzerland (red&white), DEU=Germany, HUN=Hungary, DNK=Denmark (black&white), DNR=Denmark (red&white), ESP=Spain, FIN=Finland, FRA=France, NLD=The Netherlands, NZL=New Zealand, GBR=United Kingdom, IRL=Ireland, ISR=Israel, ITA=Italy, SWE=Sweden, USA=United States

2) AM=animal model, SM=sire model, ST=single trait, MT=multiple trait, SA=survival analysis

Table 3. Country (COU), first year of data inclusion for current national genetic evaluation, trait definition, genetic evaluation model, heritability (h^2), average number of common bulls (μ_{CB}), and average post-processed correlation to other countries (μ_{RG}) for female fertility (lactating cows ability to conceive) evaluation for the Holstein breed in April 2009.

COU[1]	Data since	Trait definition	Model[2]	h^2	μ_{CB}	μ_{RG}
BEL	1980	Pregnancy Rate	ST AM	0.040	463	0.68
CAN	1996	Interval first insemination-conception in cows	MT AM	0.077	292	0.70
CHE	1994	Non return rate after 56 days	MT AM	0.010	277	0.43
CHR	1994	Cows' non return rate after 56 days	MT AM	0.010	186	0.46
CZE	1993	Cows' conception rate (pregnant or not after 3 months)	SM GSM	0.030	226	0.63
DEU	1995	Interval from first to last insemination cows	MT ML RP AM	0.010	714	0.74
DFS	1990(DNK),1988(FIN),1995(SWE)	Interval from first to last insemination cows (days)	SM RP	0.020	461	0.73
ESP	1986	Days open	ST RP AM	0.045	686	0.72
FRA	1995	Cows' conception rate (binary trait) for cows	MT AM	0.020	582	0.65
GBR	1992	Days between 1st and 2nd calving	MT AM	0.033	632	0.65
IRL	1980	Calving interval	MT AM	0.037	405	0.65
ISR	1985	Inverse of the number of insemination to conception	MT AM	0.067	44	0.52
ITA	1990	Calving Interval (days)	MT AM	0.038	512	0.69
NLD	1978(NLD), 1975(FLA)	Calving Interval	MT AM	0.145	704	0.65
NZL	1990	Lactating cow's ability to conceive	MT ML AM	0.030	354	0.49
USA	1960	Daughter pregnancy rate	ST RP AM	0.040	762	0.75

1) BEL=Belgium, CAN=Canada, CHE=Switzerland (black & white), CHR=Switzerland (red&white), CZE=Czech Republic, DEU=Germany, DFS=Denmark+Finland+Sweden, ESP=Spain, FRA=France, GBR=United Kingdom, IRL=Ireland, ISR=Israel, ITA=Italy, NLD=The Netherlands, NZL=New Zealand, USA=United States

2) ST=Single trait, MT=multiple trait, ML=multiple lactations, RP=repeatability, AM=animal model, SM=sire model

Proc. Assoc. Advmt. Anim. Breed. Genet. 18:143-146

UPDATED INDEX WEIGHTS FOR THE AUSTRALIAN PROFIT RANKING IN DAIRY CATTLE

J.E. Pryce[1], J.H.J. van der Werf[2], M. Haile-Mariam[1], B. Malcolm[3] and M.E. Goddard[1, 3]

[1]Department of Primary Industries, 1 Park Drive, Bundoora, Vic., 3083
[2]School of Rural Science and Agriculture, University of New England, NSW, 2351
[3]Institute of Land and Food Resources, University of Melbourne, Parkville, Vic., 3010

SUMMARY

Several alternative selection indexes for dairy cattle are reviewed as possible replacements for the current version of Australia Profit Ranking (APR) including: an economically optimal solution (the economic index) and two alternative indexes where the relative emphasis on fertility and fitness (compared to protein) is increased. As survival is positively correlated to yields of milk, fat and protein (production), the economic index is close to the maximum response which could be achieved for survival. However, fertility is more difficult to improve because the correlation between fertility and production is negative. To achieve a significant increase in response for fertility, its weight would have to increase considerably. Arbitrary multiples of the economic weight for survival and fertility were considered to investigate the impact on responses to selection. The index weight on survival was doubled and fertility quadrupled in the fitness index and the weight on fertility was doubled in the fertility index. Compared to the economic index, the loss in economic response was $1.40 (6%) for the fitness index and $0.21 (1%) for the fertility index. In the economic index, 53% of the total response is due to improvements in production traits, 37% is due to improved survival and 5% is due to fertility, with the remaining 5% for cell count, liveweight, temperament and milking speed. In the fitness index, the proportion of response due to production traits is 44%, survival 41% and fertility 9%. An index with more emphasis on fitness traits may improve the uptake of the APR. If this occurred it would have a favourable effect on overall genetic improvement of Australian dairy cattle.

INTRODUCTION

The Australian Profit Ranking (APR) is a total merit index introduced to the dairy industry by the Australian Dairy Herd Improvement Scheme (ADHIS) in 2001 (Valentine *et al.* 2000). APR includes estimated Australian Breeding Values (ABVs) for nine objective traits, including milk, fat and protein yields, liveweight, somatic cell count, fertility, survival, temperament and milking speed each weighted by its respective index weight. The purpose of this work was to update the farm model to derive economic values to include prices that are relevant in the foreseeable future and include marginal feed costs rather than average costs. Additionally, there are concerns in the dairy industry that insufficient weight is placed on health and fertility traits. This is believed to be one of the reasons why the uptake of this index by farmers is lower than expected. Therefore, the economic responses to selection using alternative index weights were also explored.

MATERIALS AND METHODS

Data. To derive a set of economic values for traits in the breeding objective, a farm model was constructed to reflect an average Australian dairy farm. Within each age group of cows in the model, returns from sales of milk and culls were calculated in addition to feed, rearing, health and fertility costs. The economic value of a trait was the difference in profit (per cow) between the base model and the model with one unit change in the trait of interest.

Historically, an average cost of energy has been used to estimate the cost of additional feed required to produce extra output ($0.012/MJ ME). However, using a marginal value of feed ($0.025/MJ ME) is theoretically more correct, because marginal costs account for the extra costs involved in changing a system to cope with genetic change. The marginal feed cost is generally higher than the average feed cost because feed is either bought in or money is invested on the farm to produce more feed. Economic values for liveweight, survival and milk production traits were sensitive to the value of feed used in the model. This is because liveweight is indicative of maintenance requirements, while high feed costs relative to the value of milk sales would result in the economic optimum being achieved through reducing feed costs by increasing the selection intensity on survival relative to the milk production traits.

Selection index theory was used to predict response to selection of bulls, given phenotypic information on their progeny, updated estimates of genetic parameters, and economic values of objective traits (Hazel 1943). No phenotypic information was assumed to be known for liveweight. Type traits were used to increase ABV accuracy of liveweight, survival and fertility. For other traits ADHIS estimate single trait ABVs, and the correlations between traits were taken into account when deriving optimal index weights. To achieve this, a 35×35 (co)variance matrix of APR and type traits was used (Pryce *et al.* 2008). Since most breeding values are known with high accuracy index weights were close to the economic values.

When constructing a profit based selection index each ABV is weighted to maximize response for overall profit. However, one attribute of a selection index is that close to optimal economic returns may be achieved whilst making subtle changes to the emphasis on some traits. This was explored by comparing selection responses from three alternative indexes: 1) the *economic index*, which uses the set of weights that maximizes profit, i.e. the updated economic values were used to derive weights; 2) a *fitness index* where compared with 1) the weight on survival is doubled and that on fertility is quadrupled and 3) a *fertility index* where compared with 1) the weight on fertility is doubled. By increasing the weights on fertility and survival by the amounts described, re-ranking of bulls is expected.

RESULTS AND DISCUSSION

Updated economic values (EV) along with sets of weights for the current APR index (APR), the *economic index*, the *fertility index* and the *fitness index* are shown in Table 1. The economic value for liveweight is larger than the current APR because maintenance costs were higher due to using marginal feed costs rather than average feed costs. The economic values (EVs) of milk, fat and protein have increased since 2001, because their real prices have increased.

Table 1. Economic values (EV) and index weights for APR, the economic index, fertility index and fitness index

Trait	APR	EV	Economic index	Fitness index	Fertility index
Milk	-0.048	-0.06	-0.09	-0.09	-0.09
Fat	0.9	1.08	1.02	1.02	1.02
Protein	3.8	5.78	6.67	6.67	6.67
Survival	3.9	7.39	6.02	12.04	6.02
Liveweight	-0.26	-1.17	-1.49	-1.49	-1.49
Cell Count	-0.34	-0.26	-0.61	-0.61	-0.61
Fertility	3.0	3.02	2.31	9.26	4.63
Temperament	2.0	2.59	4.13	4.13	4.13
Milking Speed	1.2	1.68	1.83	1.83	1.83

The predicted response of each index (in dollar value per trait) are presented in Table 2, assuming one year of selection results in 0.25 genetic standard deviations of genetic progress. Most of the value of genetic selection on the economic index is protein ($12.29) and survival ($8.09). The total response to selection is $21.80 per year. This is substantially higher than achieved currently (around $10) mainly because of the increase in value of traits. The response to selection in liveweight is positive due to positive correlations between milk, fat and protein yields and liveweight. A positive response in liveweight is unfavourable and the economic response is correspondingly negative.

Table 2. Dollar value of annual responses to selection in APR traits and overall merit when selection is based on the economically optimal index, the fitness index and the fertility index

	Economic index		Fitness index		Fertility index	
	$ response	%	$ response	%	$ response	%
Total merit	21.80		20.40		21.59	
Milk yield	-3.32		-2.40		-3.03	
Protein yield	12.29	53	9.25	44	11.49	51
Fat yield	2.67		2.05		2.53	
Survival	8.09	37	8.29	41	8.07	37
Fertility	0.99	5	2.06	10	1.49	7
SCC	0.37	2	0.37	2	0.37	2
Liveweight	-0.54	-2	-0.38	-2	-0.53	-2
Milking Speed	0.30	1	0.32	1	0.30	1
Temperament	0.94	4	0.86	4	0.89	4

The total economic response to selection for the *Fitness* and *Fertility* indexes was 6% and 1% lower than the economic index, respectively. The percentage of economic response in yield relative to non-production traits was 53%, 44% and 51% for the *Economic*, *Fitness* and *Fertility* indexes respectively. The response in fertility doubled through selection on the *Fitness* index as opposed to the *Economic* index. The correlations between indexes for sires were 0.91 and 0.99 between the *Economic* index and the *Fitness* and *Fertility* indexes respectively. This indicates substantial re-ranking will occur if the *Fitness* index is adopted. The Fertility index will have a much smaller impact on sire rankings.

Figure 1 is constructed by varying the index weights for survival and fertility from none to very high. It shows that selection on the economic index would result in the highest possible economic response and is the highest line of equal profit. This index also has a close to maximal response for fitness (defined as the sum of the response for survival, fertility, cell count, milking speed and temperament). Increasing the weights on fertility and survival results in increased response for these traits, a reduction in response for production, and a relatively small impact on the response in total merit. The ellipse shows that there is a slight positive correlation between production and fitness, which is mainly due to the positive correlation between production and survival.

Of course the proposed 'fitness index' or ‘fertility index’ are only two of the various options to give more emphasis on the fitness traits. Alternatives are doubling or quadrupling the weights for survival and fertility equally. However, there are some good arguments to push fertility more than survival. Firstly, it is more negatively correlated with production traits than survival; hence giving more weight to fertility has more effect on its response. Secondly, survival is positively correlated with production and response to survival is already close to maximal in the economic index. This is because survival has a component which reflects production, as high yielding cows are less likely to be culled. Sometimes survival is corrected for production traits, and redefining it this way

gives a trait uncorrelated to production, reflecting the probability of delaying involuntary culling (due to other reasons than low production). Such a trait, sometimes referred to as "functional" or "residual" survival, may be closer to what many perceive as 'robustness of dairy cows', but practically, a redefinition of trait, based on the same measurement, will not lead to another selection result.

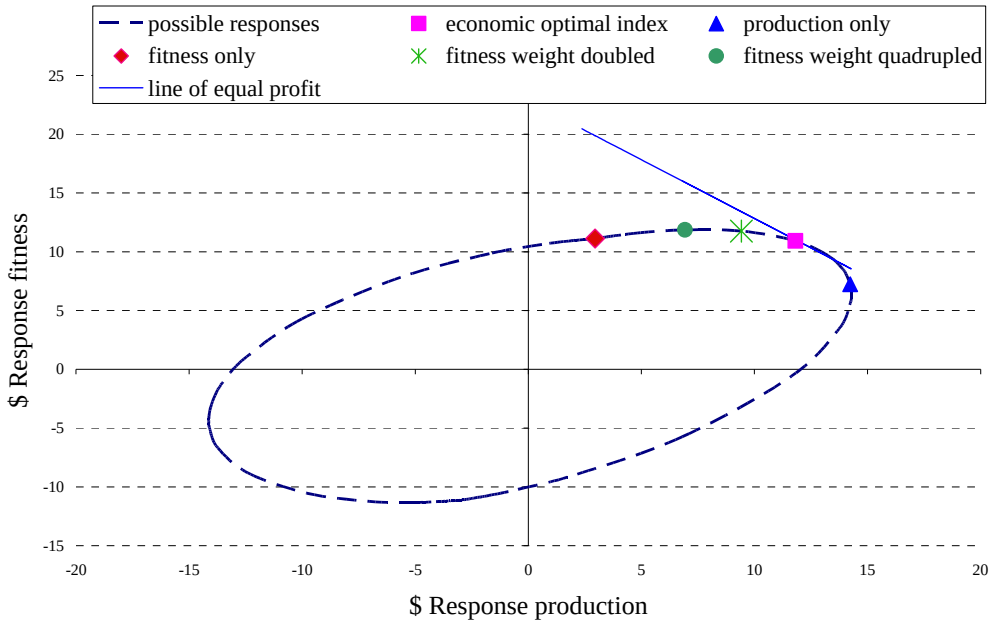


Figure 1. Economic responses in fitness and production of various selection indexes.

CONCLUSIONS

A greater selection response in survival and fertility can be achieved by deviating from the optimal economic index while still 93.6 to 99.0% of economic response to selection. Placing more weight on fertility and survival may be more acceptable to the dairy industry than rigidly accepting the economic optimum.

ACKNOWLEDGEMENTS

We would like to thank ADHIS staff for their assistance in this work and for useful discussions and Dairy Australia for funding this research.

REFERENCE

Hazel, L.N. (1943) *Genetics* **28**:476.

Pryce, J.E., Haile-Mariam, M., van der Werf, J., Malcolm, B. and Goddard, M.E. (2008) A technical manual on deriving economic values for the Australian Profit Ranking (APR) index. ADHIS publication.

Valentine, H., Jones, L., Beard, K.T. and Poole, R. (2000) The development of a total merit index for genetic selection of Australian dairy cattle. ADHIS publication.

Proc. Assoc. Advmt. Anim. Breed. Genet. 18:147-150

GENETIC MARKERS FOR LACTATION PERSISTENCY IN AUSTRALIAN DAIRY COWS

J.E. Pryce, M. Haile-Mariam, K. Verbyla, P.J. Bowman, M.E. Goddard and B.J. Hayes

Department of Primary Industries, 1 Park Drive, Bundoora, Victoria 3083

SUMMARY

A genome wide association study was used to detect quantitative trait loci (QTLs) for lactation persistency in dairy cattle. Persistency was defined as the slope after peak production calculated using test-day solutions for milk yield in Holsteins and Jerseys. As milk yield is correlated to persistency (r= 0.35), persistency was also adjusted for milk yield. Two strategies were used to search for QTLs: a SNP-by-SNP analysis where persistency solutions for sires (adjusted for fixed effects) were regressed on each single nucleotide polymorphism (SNP) in turn and a genomic selection method (BayesA) where all SNPs are fitted simultaneously. In each analysis, the discovery population comprised 743 Holstein bulls proven before 2005 and the validation datasets were 357 Holstein bulls proven after 2005 and 294 Jersey sires. A genomic region located between 21,408 kbp and 23,744 kbp on chromosome 6 had four SNPs that validated in both Holsteins and Jerseys and may indicate the presence of a QTL for persistency. The largest SNP effect from BayesA was in a similar genomic region to a SNP that validated in the single SNP analysis. False discovery rates were higher for persistency (>65%) than milk volume (24%). We hypothesise that there are many mutations that have a small effect on persistency. Genomic selection using a large number of markers appears to be a promising strategy to improve persistency. For sires not included in the prediction of SNP effects, pedigree information alone had a correlation of 0.16 with persistency EBVs, while combining genomic information with pedigree information increased the correlation with persistency EBV to 0.4.

INTRODUCTION

Persistency of lactation in dairy production is defined as the rate of decline in production after peak milk production has been reached (Cole and VanRaden 2006). Persistency may be a trait of economic importance because it can be used to select for extended lactations, which has a beneficial impact on food costs, health and fertility (Dekkers *et al.* 1998).

Dairy cattle are now routinely being genotyped for many thousands of single nucleotide polymorphisms (SNPs). One way in which SNPs are being used is through genomic selection (Meuwissen *et al.* 2001), where the effect on a trait of chromosome segments, defined by SNPs, are estimated simultaneously and used to predict breeding values. An alternative way of using SNPs is to apply a genome wide association study where individual associations between SNPs and a deregressed breeding value or daughter-yield-deviation of a trait of interest are sought. SNPs with strong relationships to traits of interest can be useful as part of a SNP panel to select animals or to improve our understanding of the biological control of a trait.

The aim of this study was to estimate individual SNP effects for persistency in dairy cattle in a reference dataset of Holstein bulls proven before 2005 and validate these SNPs in: 1) Holstein bulls proven after 2005 and 2) Jersey bulls. Secondly, we used a method of calculating a prediction equation for persistency using SNPs called BayesA (Meuwissen *et al.* 2001).

MATERIALS AND METHODS

Data. Genetic markers were obtained for Jersey and Holstein sires from Ilumina using the BovineSNP50 BeadChip (Illumina, San Diego, CA). The SNP data were edited to ensure that the

call rate was greater than 90% and the minor allele frequency was greater than 2%. SNPs that could not be mapped, or that were on the X chromosome were excluded. After editing, the number of SNPs available for analysis in each dataset was 39,048 out of the original 51,386 SNPs. Each SNP was biallelic (e.g. A and G alleles) and recoded to 1 or 2 according to the allele present at each locus. Missing genotypes were imputed using fastPHASE (Scheet and Stephens 2006).

First parity records on milk volume were extracted from the ADHIS database on cows that calved between 1999 and 2007. Only herds with daughters of bulls that had genotype information were kept. The data was edited to include only herd-test-days with a minimum of ten cows and sires with at least 30 daughters. The final datasets included records on 797,025 first lactation daughters of 3,459 sires in Holsteins and 68,230 first lactation daughters of 1,196 sires in Jerseys.

Statistical methods. Sire solutions for 3rd degree Legendre polynomials fitted to first lactation test-day records of milk yield were estimated using a random regression BLUP model in ASReml within breed (Gilmour *et al.* 2006). The (co)variance components and model used in the BLUP analysis were obtained from the study of persistency reported by Haile-Mariam and Goddard (2008). A pedigree was not included in the model because this would increase the probability of SNPs being selected on the basis of relationships between animals, rather than SNPs that are related to persistency. The solutions obtained were equivalent to daughter-yield-deviations.

Persistency (PERS) was calculated as $S_{i,54}$–$S_{i,274}$ where $S_{i,d}$ is sire solution of sire *i* on day *d* of lactation and is the gradient from after peak lactation. A 300 day milk solution (VOL) was calculated as the sum of daily yields. Persistency adjusted for milk yield (PERSadj) was calculated by regressing PERS on VOL. EBVs for PERS and PERSadj were calculated by fitting pedigree.

The Holstein data was split into two according to age of sire, as prediction of genetic merit in younger animals using associations found in older animals is generally more useful for selection. The reference dataset included solutions for Holstein bulls that received their first proof prior to 2005 (n = 743). Holsteins that were first proven after 2005 (until 2007) formed the first validation dataset (n = 357). The second validation dataset comprised 294 Jersey sires.

PERS, PERSadj and VOL were regressed on individual recoded SNPs using ASReml (Gilmour et al., 2006) in the reference and validation datasets separately. The model included the SNP as a fixed effect and sire as a random effect, a pedigree was also included. A subset of SNPs were selected where the F-probability was P<0.005 in the reference dataset and P<0.05 in the validation dataset and the direction of the SNP estimate was the same in both datasets. The false discovery rate (FDR) was calculated as FDR=E(R*P/S, where R is the number of tests, P is the p-value used to in the F-test and S is the number of SNPs with significant F values; e.g. the expectation of the number of false discoveries by chance divided by the actual SNPs significant at this threshold.

The SNPs were also fitted simultaneously in a BayesA model, similar to the model described by Meuwissen et al. (2001), modified to include a polygenic effect. The SNP effects were estimated in the reference population and molecular breeding values (MBV) calculated for the Holstein validation sires by summing the SNP effects multiplied by the allele frequency at each SNP position. The correlation between the persistency EBVs estimated from the data plus pedigree and the corresponding MBV estimated using genomic information in the Holstein validation dataset was calculated. The BayesA SNP solutions were also used to verify SNPs detected using the single SNP approach.

RESULTS AND DISCUSSION

The number of significant SNPs (P<0.005) for persistency in the reference data set were 302 and 262 for PERS and PERSadj respectively, which is a false discovery rate (FDR) of 65% and 75% (Table 1). A low FDR indicates that more SNPs were statistically significant than expected by chance. Of the SNPs that were P<0.005 in the reference dataset, 20 and 22 SNPs were at

P<0.05 in the Holstein validation population and 12 and 9 were in the same direction as the reference population for PERS and PERSadj respectively. Seven and three SNPs validated in Jerseys. FDRs in the Holstein validation dataset were 66% to 69% for PERS and PERSadj and 101% to 164% in Jerseys for PERS and PERSadj.

Table 1. **Number of SNPs found to be significant, in the same direction in Holstein reference and Holstein and Jersey validation datasets and percentage of false discovery rates (FDRs), when SNPs were regressed individually on persistency (PERS), persistency adjusted for milk volume (PERSadj) and 300d milk volume (VOL)**

	PERS	PERSadj	VOL
Holstein reference P<0.005	302	262	738
Holstein validation P<0.05	22	20	152
Holstein same direction	12	9	144
Direction validated	55%	45%	95%
Jersey P<0.05	15	8	77
Jersey same direction	7	3	37
Direction validated	47%	38%	48%
FDR reference Holstein	65%	75%	26%
FDR validation Holstein	69%	66%	24%
FDR validation Jersey	101%	164%	48%

Table 2. SNPs validated at P<0.005 in the reference population and P<0.05 in the validation dataset in Holsteins (H) and Jerseys (J) persistency adjusted (PERSadj), only SNPs that validate in the same direction are shown for Holsteins

Breed	Direction*	SNP name	Chromosome	Position (bp)	F-Prob*
H	-	BTB-01600593	2	16,000,786	0.017
H	-	ARS-BFGL-NGS-112143	4	10,139,426	0.038
H	-	ARS-BFGL-NGS-27962	6	21,593,191	0.017
H	-	UA-IFASA-1756	6	21,620,640	0.023
H	-	BTA-82896-no-rs	8	11,649,044	0.041
H	-	Hapmap59058-rs29016195	12	8,844,600	0.026
H	-	BTA-42074-no-rs	17	11,271,481	0.004
H	+	ARS-BFGL-NGS-38620	18	64,466,895	0.045
H	+	BTB-00920286	26	3,410,026	0.013
J	-	BTB-00780124	1	144,105,011	0.042
J	+	BTA-47105-no-rs	5	113,682,010	0.043
J	-	ARS-BFGL-NGS-60840	6	13,520,548	0.025
J	+	BTB-00245990	6	21,408,490	0.003
J	-	BTB-00245990	6	23,744,743	0.012
J	-	ARS-BFGL-BAC-35623	6	70,432,390	0.022
J	+	ARS-BFGL-NGS-60840	13	37,444,034	0.041
J	+	BTB-00245990	19	33,644,562	0.007

*In validation population

The correlation between PERS and VOL was 0.35 in the Holstein combined reference and validation datasets. Therefore, PERSadj may be better measures of persistency as it is independent of yield. However, the correlation between PERS and PERSadj was 0.94, so they are similar traits.

Although no single SNP validated for PERSadj in both Holsteins and Jerseys (Table 2) there was a genomic region between 21,408 kbp and 23,744 kbp on chromosome 6 that had SNPs validating in both breeds. This is not in the same region as mutations previously found to affect milk production, for example ABCG2 which is located at approximately 37 Mbp (Cohen Zinder *et al.* 2005). The largest effect from the Bayes A analysis (located on chromosome 2 at 15,658 kbp) was close to a SNP that validated in Holsteins (Table 2). There were two SNPs that validated in both Jerseys and Holsteins for milk volume (VOL), both on chromosome 14 and close to the mutation known to exist on chromosome 14 that affects milk yield (DGAT1; Grisart *et al.* 2002).

The correlation between the equivalent PERS EBV and MBV in the Holstein validation dataset (i.e. sires not included in the prediction of SNP effects) using BayesA was 0.40 for PERS and 0.28 for PERSadj, the equivalent correlations between the EBV and parent average EBV was 0.16 and 0.25 for PERS and PERSadj respectively. Thus, genomic selection may be a suitable way to improve persistency.

In the BayesA method all SNPs were fitted simultaneously, while each SNP was fitted in turn in the SNP by SNP analysis. The SNP by SNP analysis is therefore much more likely to select close neighbours, while the BayesA analysis will lead to the selection of the SNP with the largest effect in a groups of neighbours and is therefore more likely to be close to the causative mutation in cases where a number of candidate SNPs have been discovered in the SNP by SNP analysis.

CONCLUSION

Having two validation populations, one being the same breed as the discovery dataset and the other an alternative breed, provides a powerful way to validate SNPs discovered using GWAS and minimise the risk of pursuing incorrect genomic regions when false discovery rates are high. This is especially important for low heritability traits such as persistency. However our results suggest that there are many mutations of small effect on persistency. Genomic selection using a large number of markers appears to be a promising alternative.

ACKNOWLEDGEMENT

We would like to thank the Gardiner Foundation for funding this research.

REFERENCES

Cohen-Zinder, M., Seroussi, E., Larkin, D.M., Loor, J.J. and Everts-van der Wind, A. (2005) *Genome Res*. **15**:936.

Cole, J.B. and VanRaden, P.M. (2006) J. Dairy Sci. **89**:2722.

Dekkers, J.C.M., Ten Hag, J.H. and Weersink, A. (1998) *Livest. Prod. Sci.* **53**:237.

Gilmour, A.R., Gogel, B.J., Cullis, B.R. and Thompson, R. (2006) "ASReml User Guide Release 2.0" VSN International Ltd, Hemel Hempstead, UK.

Grisart, B., Coppieters, W., Farnir, F., Karim, L., Ford, C. Berzi, P., Cambisano, N., Mni, M., Reid, S., Simon, P., Spelman, R., Georges. M. and Snell, R. (2002) *Genome Res*. **12**:222.

Haile-Mariam, M. and Goddard, M.E, (2008) *Animal*, **2**:325.

Meuwissen, T.H.E., Hayes, B.J. and Goddard, M.E. (2001) *Genetics* **157**:1819.

Scheet P. and Stephens, M.A. (2006) *Am J Hum Genet* **78**:629.

Proc. Assoc. Advmt. Anim. Breed. Genet. 18:151-154

GENOME WIDE ASSOCIATION STUDIES IN DAIRY CATTLE USING HIGH DENSITY SNP SCANS

H.W. Raadsma[1], M.S. Khatkar[1], G. Moser[2], M. Hobbs[1], R. Crump[3], J.A.L. Cavanagh[1] and B.Tier[3]

The CRC for Innovative Dairy Products, [1] ReproGen, Faculty of Veterinary Science, The University of Sydney, Camden, Australia, [2] Bellbowrie, QLD, 4070, Australia , [3] Animal Breeding and Genetics Unit, UNE, Armidale, NSW, 2351, Australia

SUMMARY

Use of high density Single Nucleotide Polymorphic (SNP) marker information allows for prediction of genetic merit via genome wide selection and for localization of markers in gene regions of biological interest through Genome Wide Association Studies (GWAS). We report on a replicated GWAS in dairy cattle using 1,945 progeny tested bulls genotyped with three high density SNP panels representing 63,678 informative SNP. Single SNP genotypes were analysed against deregressed EBV for protein percent and fat percent using a mixed linear model accounting for SNP and animal polygenic effects. The 127,356 analyses (63,678 informative SNP by two traits) across the two data sets identified 143 and 87 significant ($P<0.05$, corrected for False Discovery Rate) associations for protein % in data set 1 and 2 respectively, whilst for fat % 102 and 61 significant associations were identified in the two data sets respectively. Outputs from selected SNP analyses are discussed for significance and pleiotropic effects and compared against integrated QTL meta-assembly from public domain studies

INTRODUCTION

Development of high-density large-scale single nucleotide polymorphism (SNP) genotyping platforms (Hardenbol *et al.* 2005) has opened the possibility of large scale genomic investigations. Typically in livestock high density SNP typing has seen the transition from linkage and QTL mapping to the prediction of total genetic merit using genome wide selection or genomic selection approaches (Schaeffer 2006; Raadsma *et al.* 2008) where marker location is not required, or focus on localised SNP in genes to dissect the genetic architecture underlying complex traits through Genome Wide Association Studies (GWAS). Whilst in human over 220 GWA studies have reported on complex disease/performance traits (National Human Genome Research Institute catalogue of published genome-wide association studies at http://www.genome.gov/26525384) relatively few studies describing applications of GWAS in cattle have been reported to date (Daetwyler *et al.* 2008; Kolbehdari *et al.* 2008). In this study we report on a GWA study using two data sets with overlapping SNP data against two commonly recorded milk performance traits.

MATERIALS AND METHODS

Bulls and Genotyping. DNA was extracted from 1,945 Australian progeny-tested Holstein Friesian dairy bulls. All bulls were sourced from Genetics Australia, and represented a cohort of sires used for ongoing commercial use (proven) or rejected (non selected) following progeny testing. The sires were born between 1955 and 2001, with >96% of sires born after 1980. All sires had Australian Breeding Value (ABV) data calculated by the Australian Dairy Herd Improvement Scheme (ADHIS) for traits associated with lactation performance, conformation, reproductive fitness and disease resistance. The bull samples were split into two data sets, the first consisting of 1,309 bulls genotyped with 15,036 SNPs (Khatkar *et al.* 2007) and the second data set of 634 bulls genotyped with 25K (Affymetrix) and/or a 50K SNPs platform (Illumina). Across the two data

sets 8072 SNPs were in common. Locations of the SNPs were determined on the bovine sequence assembly Btau 4.0 (ftp://ftp.hgsc.bcm.tmc.edu/pub/data/Btaurus/fasta/Btau20060815-freeze/).

Traits and Analyses. Deregressed ABV data on protein percent (prot%) and fat percent (fat%) were selected for the GWAS. The deregressed Estimated Breeding Values (EBVs, **y**) were analysed against each SNP by fitting mixed linear models **y** = **Xb** + **Za** + **e**, with SNP as fixed effect (**Xb)** and animal polygenic effect **(Za)** as random using ASReml (Gilmour et al., 2000) providing a nominal significance value and effect size for each SNP. All SNP effects were standardized by dividing the estimated effects by the standard deviation (SD) computed from the test sample for each trait. The effects and probabilities for each SNP were aligned in a genome browser against a spline of a QTL meta-analysis for all published QTLs based on an average score system adapted from Khatkar (2006), allowing for direct comparison with public domain studies. The browser allows selection of subsets of trait/data set combinations for internal comparison. The p-values were scaled to $-\log_{10}$(p-value). The false discovery rate (FDR) for the test of each marker were computed as q-value implemented in R package qvalue (Storey and Tibshirani 2003).

RESULTS

The 127,356 analyses (63,678 informative SNP by two traits) across the two data sets identified 143 and 87 significant P<0.05FDR associations for prot% in data set 1 and 2 respectively, whilst for fat % 102 and 61 significant associations were identified in the two data sets respectively. The highly significant associations occurred across the genome in clusters of SNPs (Figure 1 a and b) some of which were strongly aligned with known QTL locations, whilst others identified novel regions for which no QTL have previously been reported.

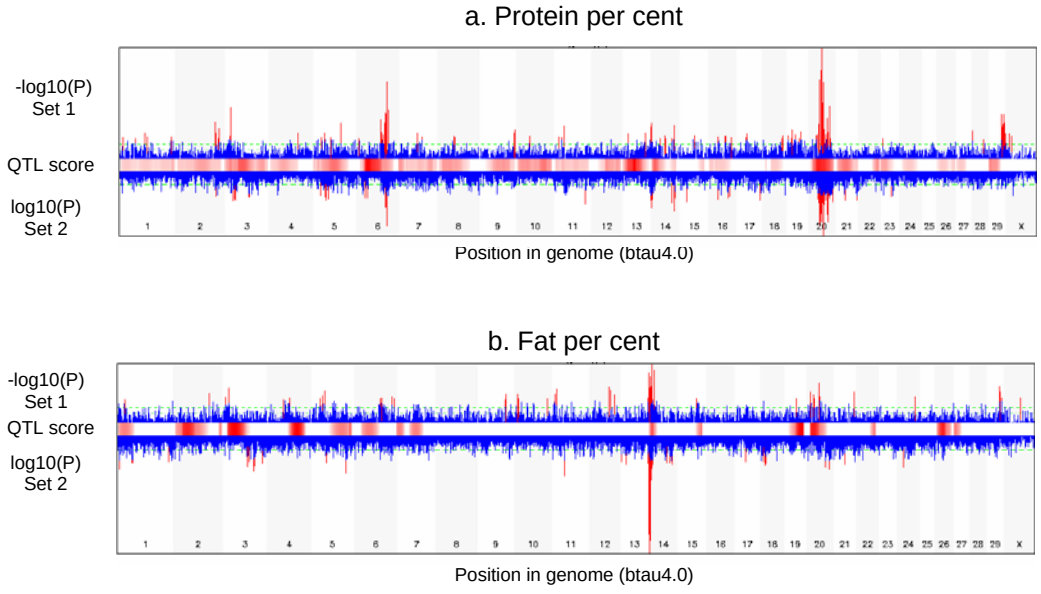


Figure 1. Distribution of SNP association $-\log_{10}$(P) along the genome for 2 data sets of bulls for (a) prot% and (b) fat% as well as alignment against QTL meta score from public domain analyses. SNP association in red exceed FDR significance threshold, green broken line P=0.01, and QTL meta score intensity increases as QTL score increases.

The distribution of SNP effects showed a similar trend for both traits and both data sets and was strongly skewed. Only 10 and 56 SNPs showed an effect of greater than 0.5SD for prot% and fat% respectively. Similarly 10 % and 9.2 % of SNP showed a small (<0.01SD) or non existent effect for prot% and fat% respectively. The relationship of significance values between common SNPs measured in both data sets for the same trait was r= 0.04 and r=0.16 based on Spearman rank correlations, for prot% and fat % respectively showing a high degree of SNP variability in detecting a significant relationships across the two data sets. However, the relationship of effect size between common SNPs measured in both data sets for the same trait was r= 0.28 and r= 0.20 for prot% and fat % respectively showing some degree of SNP repeatability in detecting a similar effect relationship across the two data sets.

For both data sets a strong correlation was observed between SNP showing an association for both prot% and fat%, (r= 0.25 for P-values and 0.60 for effect size). Although the majority of SNP had a positive or negative pleiotophic effect on prot% and fat%, some exceptions were evident with an antagonistic sign of effect (Figure 2a,b) Similarly a strong correlation (r= 0.50) was observed between deregressed EBV for the 1945 sires for prot% with fat % (Figure 2c).

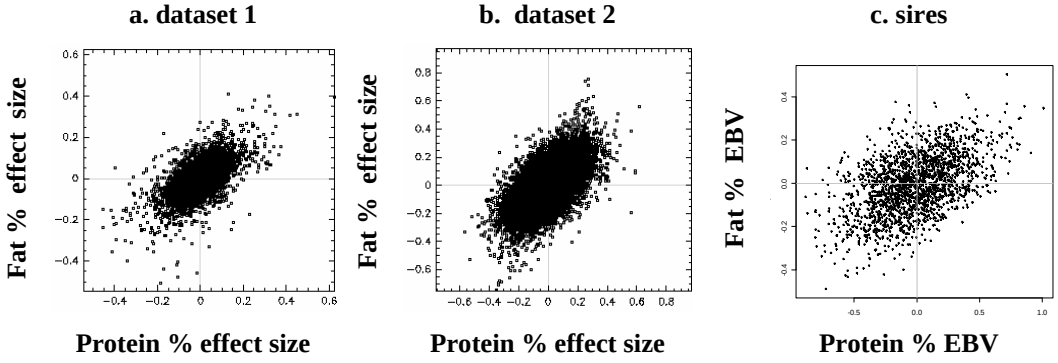


Figure 2. Association between SNP effects for prot% and fat%, data set 1 (a), data set 2 (b) and (c) deregressed EBV solutions for prot% and fat% for the sires used in the association analysis.

DISCUSSION AND CONCLUSIONS

From the large number of single point associations conducted across the two data sets, only a relatively low number of SNP showed a highly significant association which exceeded the threshold for FDR (outlined in red in Figure 1). Furthermore most of the significant SNPs occurred in clusters most likely as a result of LD between closely spaced SNP and the underlying QTL or QTN. Using a replication of SNP association shows a negligible repeatability on significance probability across two data sets for individual SNPs suggesting that some of the significant SNP may still be spurious. Repeatability for effect size attributed to each of the SNP was higher possibly as a scale effect given the broader range for effect compared with P values and the extend of LD over larger segments containing several SNP In line with expectations of likely gene effects, relatively few SNP were associated with very large effects (>0.5SD) as has been shown previously for QTL effects (Hayes and Goddard, 2001). Although it is attractive to select such SNP associations for further investigation, some caution is warranted since these effects may be over estimated and occur by chance alone. Some account of replication is essential.

The most obvious feature of a GWAS is the strong alignment of SNP to QTL which have been independently verified in previous studies, including some cases where an underlying QTN has

been identified. Such clusters were evident on BTA 20 for prot% and BTA 14 for fat%. (http://www.vetsci.usyd.edu.au/reprogen/QTL_Map/). The cluster of SNPs on BTA 6 for prot% shows some departure in location from the widely known QTL cluster on BTA 6 for prot %. The GWAS also showed failure to detect significant SNP in known QTL regions on BTA3 and BTA10 for prot% and BTA 2,3,4,5and 25 for fat% (Figure 1), possible reasons may that no QTL are segregating at those locations in this population, or SNP density is too low to provide LD to underlying QTN although both reasons are unlikely and warrants further examination.

A key problem in applying of GWAS in livestock is the confounding of the estimates of SNP effects with pedigree, and the SNP may partly act as marker for detecting the relationship between individuals without having a large true effect. The BLUP procedure is likely to account for some of the unrelated polygenic effects vs. specific SNP effects, but the confounding is unlikely to be fully compensated for. However, using EBVs from progeny tested sires has the advantage that all polygenic differences between animals are captured with a high degree of accuracy. Finally although it is attractive to search for all large SNP effects independently and use these as marker panel for MAS, it is unlikely to be as efficient as genome wide selection/genomic selection procedures where the effects of all SNP are considered in predicting genetic merit. Such approaches have now shown utility in independent dairy sets (Raadsma *et al.* 2008); (Hayes *et al*, 2009). The down side from such prediction functions is a requirement that all SNPs be genotyped, unless stringent SNP selection procedures are applied which allow for subsets to be used without loss of predictive power. Alignment of SNP identified under GWAS and GS has not yet been reported.

ACKNOWLEDGMENTS

The genotype data and phenotype was provided by the Co-operative Research Centre for Innovative Dairy Products.

REFERENCES

Gilmour, A.R., Thompson, R., Cullis, B.R., and Wellham, S.J. (2000) ASReml Reference Manual.

Daetwyler, H.D., Schenkel, F.S., Sargolzaei, M. and Robinson, J.A. (2008) *J. Dairy Sci*. **91**:3225.

Hardenbol, P., F. Yu, J. Belmont, J. Mackenzie, C. Bruckner, T. Brundage, A. Boudreau, S. Chow, J. Eberle, A. Erbilgin, M. Falkowski, R. Fitzgerald, S. Ghose, O. Iartchouk, M. Jain, G. Karlin-Neumann, X. Lu, X. Miao, B. Moore, M. Moorhead, E. Namsaraev, S. Pasternak, E. Prakash, K. Tran, Z. Wang, H. B. Jones, R. W. Davis, T. D. Willis and R. A. Gibbs, (2005) *Genome Res* **15**: 269.

Hayes, B. and Goddard, M.E. (2001). *Genet. Sel. Evol.* **33:**209.

Hayes, B.J., Bowman, P.J., Chamberlain, A.J. and Goddard, M.E. (2009) *J. Dairy Sci.* **92**:433.

Khatkar, M.S. (2006) PhD Thesis, University of Sydney.

Khatkar, M. S., Zenger, K.R, Hobbs, M., Hawken, R.J., Cavanagh, J.A.L., Barris, W., McClintock, A.E., McClintock, S., Thomson, P.C., Tier, B., F. Nicholas, F.W., and H. W. Raadsma, H.W. (2007). *Genetics* **176:**763.

Kolbehdari, D., Z. Wang, J. R. Grant, B. Murdoch, A. Prasad, Z. Xiu, E. Marques, P. Stothard and S. S. Moore, (2008) *J Dairy Sci* **91**: 2844.

Raadsma, H.W., Moser, G., Crump, R.E., Khatkar, M.S, Zenger, K., Cavanagh, J., and Tier, B. (2008) *Dev. Biologicals. (Basel)* **132:**219.

Schaeffer, L.R. (2006) *J. Anim. Breed. Genet.* **123**:218.

Storey, J.D. and Tibshirani, R. (2003) *Proc. Natl. Acad. Sci.* U S A **100:**9440.

Proc. Assoc. Advmt. Anim. Breed. Genet. 18:155-158

CRITERIA FOR SELECTING AND PREDICTING HERDLIFE IN DAIRY CATTLE

M. Haile-Mariam[1] and M.E. Goddard[1,2]

[1]Biosciences Research Division, DPI Victoria, 1 Park Drive, Bundoora 3083, Australia
[2]Faculty of Land and Food Resources, University of Melbourne, Parkville, Vic 3052 Australia.

SUMMARY

In dairy production, herdlife is a highly desirable trait with a large effect on overall profitability. In Australia, the measure of herdlife is known as survival index and it is currently based on estimated breeding value (EBV) on udder depth, pin set, overall type and likeability and survival scored as survived or culled (dead) at the end of each lactation. The first objective of this study was to compare differences in survival of daughters of sires of high versus low genetic merit for a selected number of alternative selection criteria. The second objective is to report useful indicator traits that can be used to predict survival breeding value for bulls. Data of over one million cows that calved between 1990 and 2002 and EBV of over 6000 sires were used. Difference in survival of daughters of high vs. low genetic merit was the largest if bulls were categorised by their survival index or their Australian Profit Ranking (the economic index in Australia called the APR) than by their EBV for type or workability traits. Cows from sires with the highest EBVs of mammary system, overall type and likeability showed a slightly better early survival (parity 1 to 2). Late survival of cows (parity 2 to 6) from high EBV bulls for udder depth was higher than those from low EBV bulls. This study also showed that there is a potential to improve prediction of the survival index of young sires by using their EBV on type, workability and somatic cell count compared to the current survival index which is based on type and workability traits. In conclusion, farmers who want to have cows that last longer in their herd should use bulls with high survival index or high APR.

INTRODUCTION

In Australia, currently herdlife accounts for 15% of the total economic index called the Australian profit ranking (APR). The Australian Dairy Herd Improvement Scheme (ADHIS) publishes an estimated breeding value for herdlife called the survival index or ABV which is derived from estimated breeding values (EBV) for udder depth, pin set, overall type and likeability and survival scored as survived or culled (dead) at the end of each lactation. Increasing herdlife would increase the proportion of more mature, higher yielding cows in the herd, increase the opportunity for voluntary culling and reduce replacement costs. However, of all economic traits herdlife is the most difficult to measure because cows should be culled before their herdlife is actually known. The heritability of survival is also low so genetic progress is generally slower than for other dairy traits.

Thus, to improve herdlife through selection, genetic evaluation for type traits was introduced in 1980s and for young bulls, selection for herdlife largely depends on indicator traits which are mainly type traits. In most situations, estimated genetic correlations between type traits and longevity which are assumed to be linear are used to predict herdlife indirectly (Boldman et al. 1992; Vollema et al. 2000). However, few traits have a strictly linear relationship with herdlife and the role of traits with intermediate optima or traits that offer "diminishing returns" as scores increase cannot be evaluated properly using genetic correlations (e.g. Berry *et al.* 2005). Alternatively phenotypic records for indicator traits and actual survival score of cows (Berry *et al.* 2005) can be used to predict herdlife EBVs of sires. The practical application of this method in Australia is limited, because only a small proportion of cows are scored for type traits. Another option for developing prediction equations for herdlife of sires is based on EBV of sires for

indicator traits and actual survival data of individual cow (Brotherstone and Hill 1991). In this way data on all sires with reliable EBV for indicator traits and all their progeny with survival score can be used and relations that are not linear can be accommodated.

Because of the historical relationship between type traits and survival, the value attached to some type traits as indirect predictors of herdlife may be greater than is necessary. The first objective of this study is to demonstrate the differences in survival of daughters of sires of high versus low genetic merit for a selected number of indicator traits. The second objective is to present an alternative method for predicting survival EBV (indirect) of bulls from their EBV on different traits and actual survival data of their progeny (cows)**.**

MATERIALS AND METHODS

Data on EBV of Holstein bulls for several traits and actual survival score of their cows were obtained from ADHIS (http://www.adhis.com.au). In the survival data, cows at the end of each lactation were scored as 1 (survived to next lactation) or 0 (culled or died). Cows that calved for the first time between 1990 and 2002 with their milk yield and survival data were included. About 1 million cows which were progeny of ~ 6000 bulls were available for analysis. Bulls selected for these analyses had a reliability of at least 60 % but the average reliability for all bulls was above 90%. The survival data of the cows were merged with the EBV of their sires for all traits.

Table 1. Number cows with survival score and number of sires with high and low genetic merit for type traits, likeability, cell count and survival, APR and ASI of their sires.

Trait	Mean[#] EBV	High genetic merit(EBV or APR)[†]			Low genetic merit(EBV or APR)[‡]		
		No. cows	No. sires	Mean EBV	No. cows	No. sires	Mean EBV
APR	12	31395	96	83	154216	1277	-64
Survival index	100	86743	274	105	272099	1060	96
ASI	9	23782	99	73	140298	1211	-55
Cell count	101	113991	1105	140	93365	991	62
Likeability	100	102098	206	103	166996	1092	96
Udder Depth	100	70128	428	113	153454	535	82
Overall type	101	100427	182	110	103419	1037	90
Mammary system	100	129202	218	109	109183	1068	90

[†]Sire and their progeny in the top 10 % for traits of interest are categorised in the high group;
[‡]Sire and their progeny in the bottom 10 % for traits of interest are categorised in the low group;
[#]Mean for all Holstein bulls in the ADHIS database of Feb. 2008 genetic evaluation.

Survival of cows from low versus high genetic merit. To demonstrate differences in survival ability between progeny of extreme group of sires, cows from the top 10% of the sires were classified as high EBV or APR (referred hereafter as high genetic merit) group and cows from the bottom 10% of sires were classified as low EBVs or APR (referred as low genetic merit) for traits of interest. The number of cows and their sires (with mean) for high versus low genetic merit group is shown in Table 1. Survival of cows in the first 5 parities from sires with high versus low genetic merit for type traits, likeability, cell count, survival index, Australian Selection Index (ASI, calculated from EBV on protein, fat and milk yield) and APR was determined. The y-variable was survival score of cows recorded as 1 or 0 at the end of each parity and the model used to determine difference between the high and low genetic merit included herd-year-season of calving (HYS), age at calving (AFC) and month of calving of the cows. As the objective of these analyses was to

point to farmers useful selection criteria for herdlife, the y-variable was survival unadjusted for milk yield.

Indirect predictors of survival. The second set of analyses were performed to derive prediction equations for survival from EBVs for type, workability and cell count and the actual survival record of their progeny. This involved several regression analyses fitting EBV of bulls for predictor traits (as covariate) and HYS, AFC and month of calving of cow. Non-linear relationships (i.e. regressions) were tested by fitting linear and quadratic component of the indicator traits. As the objective when selecting for survival was to reduce involuntarily culling, the y-variable (survival to next lactation) was adjusted for relative production of cows by fitting production index as a covariate in the same model. Finally all EBVs of the predictor traits which had significant ($P<0.05$) effect on survival when fitted individually were included in the model to get the best possible equation that predicts early survival (from first to second) and late survival (second to sixth parity). As the results of these analyses, two prediction equations one for early and another for late survival were developed.

RESULTS AND DISCUSSION

Survival of cows from low versus high genetic merit. As expected difference in survival score at the end of each lactation between progeny from high and low genetic merit was the largest when sires were categorised based on survival index (Table 2) followed by APR. The survival scores of cows with high sire EBV for mammary system were not statistically different from those with low sire EBV (results not tabulated). In the case of overall type, the difference was significant ($P<0.05$) in survival from first to second parity (Table 2) only. In the case of likeability, differences were only significant in the first 3 parities. All other differences between low and high genetic merit sires shown in Table 2 were statistically significant ($P<0.05$).

Table 2. Survival of cows at the end of each parity (%) from sires with high and low genetic merit for survival index, cell count, likeability and udder depth.

Par.	APR		Survival index		Cell count		Likeability		Udder depth		Overall type	
	Top	Last	Top	Last	Top	Last	Top	Last	Top	Last	Top	Last
1	85	80	86	79	82	80	85	80	83	81	85	83
2	84	79	85	78	82	79	83	80	82	80	83	83
3	81	77	83	76	80	77	80	78	81	78	81	80
4	79	73	79	72	77	72	76	75	76	73	76	76
5	76	67	75	66	71	67	70	70	71	67	74	72

Difference between survival score of cows from high vs. low EBV sires increased with increase in parity when sires were categorised by their EBV for udder depth (Table 2). In the case of likeability, ASI (results not tabulated) and overall type, for which voluntary culling is practiced, differences between progeny of low and high EBV sires were larger in the first parity (early survival) than in later parities. Among the currently used predictors in survival index, udder depth followed by likeability were the most important traits associated with differences in survival of cows from high EBV versus low EBV sires were observed (Table 2). The results show the consequence of selecting the top or bottom bulls for a number of traits on the longevity of their progeny. Type traits like overall type and mammary system which are highly valued by some farmers are not necessarily good at predicting survival of cows.

Indirect predictors of survival. EBV on udder depth, overall type, likeability, somatic cell count, pin set and milking speed had positive effect on early survival (parity 1 to 2) whereas the effect of angularity and foot angle was negative. These traits can be used as indirect predictors of early survival. Survival in later parities were positively affected by EBV on udder depth, overall type, likeability, somatic cell count, pin set and teat length and negatively affected by rear attachment height and angularity. In addition in both early and late survival, the relationship between body depth EBV and actual survival was non-linear where cows from intermediate sires have a better survival than cows from high or low EBV bulls and these effects were also considered when deriving the two prediction equations, one for early and one for late survival. These two equations can be combined to predicted survival BV for a bull.

The simple correlation between the official survival index from ADHIS and the indirectly predicted survival value calculated using the equations from the current study for bulls that had over 100 progeny that calved after 2002 for the first time was 0.76 and 0.81, when the equation from early and late survival were used for prediction, respectively. The corresponding correlation value when the ADHIS equation based on overall type, udder depth, pin set and likeability was applied was slightly lower at 0.61.This shows that there is a potential to improve the prediction of the survival index of young bulls by using their EBV on type, workability and somatic cell count. However, selection against angularity may not be popular so farmers need to be convinced about the usefulness of the set of predictors suggested,

In conclusion, the largest difference in survival of daughters was observed if bulls were categorised by their survival index or their APR into low and high genetic merit group than by their EBV for type or workability traits. The value of individual type traits such as overall type and mammary system in predicting survival is rather small. The study also showed that there is a potential to improve prediction of survival of bulls by using separate equations based on combination of type, workability and somatic cell for predicting early and late survival. However, before these prediction equations are recommended for wider use extensive validation and then extension work to convince the industry on the value of selecting against some type traits such as angularity to improve herdlife is necessary.

ACKNOWLEDGEMENTS

This research is funded by Dairy Australia and Victorian Department of Primary Industries. We thank ADHIS for providing the data and Dr Kevin Beard for data extraction.

REFERENCES

Berry, D. P., Harris, B. L., Winkelman, A. M. and Montgomerie, W. (2005) J. Dairy Sci. **88**:2962.
Boldman, K. G., Freeman, A E., Harris, B. L. and Kuck, A. L. (1992) J. Dairy Sci. **75**:552.
Brotherstone, S. and Hill, W.G. (1991) Anim. Prod. **53**:279.
Vollema, A. R, Van Der Beek, S., Harbers, A.G.F. and De Jong, G. (2000) J. Dairy Sci. **83**:2629.

Proc. Assoc. Advmt. Anim. Breed. Genet. 18:159-162

THE IMPACT OF GENETIC MARKERS FOR TENDERNESS ON STEER CARCASS AND FEEDLOT EXIT AND HEIFER PUBERTY TRAITS IN BRAHMAN CATTLE

M.L. Wolcott and D.J. Johnston

Animal Genetics and Breeding Unit*, University of New England, Armidale, NSW 2351, Australia

SUMMARY

Four genetic markers (T1, T2, T3 and T4) have been shown to have a significant effect on tenderness (measured as shear force) in Brahman cattle. This study examined the relationship between tenderness markers and steer (N = 940) feedlot exit and carcass, and heifer (N = 973) puberty traits. For most traits assessed, tenderness markers had no significant effects. Differences in T1 genotype, however, significantly ($P < 0.05$) affected P8 fat depth measured in steers at the end of finishing and/or on the carcass, as did T3 and T4. Regression coefficients for these relationships were consistently negative, showing that steers with more copies of the favourable alleles were leaner. When protocols were applied to combat false discovery in multiple testing analyses, the relationship between T3 genotype and feedlot exit P8 fat depth remained significant. T1 also influenced age at puberty in Brahman heifers, with the trait significantly reduced (regression coefficient = -14.54) for animals which had more copies of the favourable allele. The exploitation of tenderness markers in Brahman cattle will need to be undertaken with an awareness of the impact of selection for these markers on other production and puberty traits.

INTRODUCTION

Tenderness is the key factor influencing consumer satisfaction reported for cooked beef products (Egan *et al.* 2001). As tenderness can only be measured on slaughter stock late in the production cycle, and measurement is time consuming and expensive, it has been among the first traits in beef cattle to be targeted for genetic marker analyses. The recently released SmartGENE for Beef report (Johnston and Graser 2008) showed that genetic markers for tenderness explained a significant amount of variation for the trait (measured as kg shear force) in Brahman cattle. If genetic marker results are to be incorporated into the multi-trait BREEDPLAN evaluation, there is a need to determine whether there is any association between marker genotypes and other important carcass and production traits. This study aimed to examine the relationship between tenderness marker genotypes and steer finishing and meat quality, and heifer puberty traits, to determine the degree to which selection based on tenderness markers may impact these key productivity traits.

MATERIALS AND METHODS

Animals. The experiment involved Brahman steers and heifers representing 53 sires. The cattle were bred on 5 co-operating properties in Queensland and the Northern Territory with calving taking place over 4 years from 1999 to 2003. After weaning, steers (N = 853) were relocated to one of five backgrounding properties, where they were grown out to feedlot entry, at a mean target group liveweight of 400kg. Steers were implanted with a hormonal growth promotant (COMPUDOSE®) at an average age of 10 months, with implants replaced, according to manufacturer recommendations, until feedlot exit. Steers were slaughtered after an average of 119 days in the feedlot (Wolcott, *et al.* 2009). After weaning, heifers (N = 907) were transported to one of 3 locations which characterized the range of beef production systems in northern Australia.

*AGBU is a joint venture of NSW Department of Primary Industries and University of New England

Heifers were mated for the first time at approximately 27 months. At each location, heifers of the same year of birth were managed as a single contemporary group (defined as a cohort). See Johnston *et al.* (2009) and Barwick, *et al.* (2009) for a complete description of heifer and steer allocation and management respectively.

Measurements. Eleven feedlot exit and 11 carcass traits (described by Barwick *et al.* 2009 and Wolcott *et al.* 2009 respectively) were analyzed. Feedlot exit measurements included liveweight, growth rate from feedlot entry to feedlot exit, feed intake and efficiency, hip height and condition score, and ultrasound scanned measurements of P8 fat depth (P8X), rib fat depth (RIBX), and eye muscle area (EMAX). Blood samples were collected at feedlot exit for insulin like growth factor I (IGFX) analysis. Carcass measurements included carcass weight, hot P8 fat depth (HP8), cold P8 (CP8) and rib fat depth, eye muscle area, retail beef yield as well as Meat Standards Australia grading of hump height (HUMP), USDA marbling score (MS) and ossification score. At slaughter, right sides were hung by the Achilles tendon. Sides were chilled overnight, and at approximately 24 hours post mortem, a 15cm sample of the *M. Longissimus thoracis et lumborum* (LD) muscle was collected caudal to the $12/13^{th}$ rib, and frozen for later shear force (SF) measurements. Samples were thawed, cooked in a water bath (70^0C for 60min) and chilled overnight, prior to SF measurement. A second sub-sample of the LD was used to measure percent intramuscular fat (CIMF) using a near infra-red spectrophotometry method. See Perry *et al.* (2001) for a complete description of methodologies regarding meat sample preparation and trait measurement.

Heifer puberty traits are described by Johnston *et al.* (2009). Briefly, prior to the onset of puberty (at a cohort mean liveweight of 200kg), a regime of ultrasound scanning at approximately monthly intervals was initiated to identify the presence of the first corpus luteum (CL), which, when identified, was interpreted as signaling the onset of puberty. When the first CL was observed, measurements of age (AGECL), liveweight, P8 fat depth, and condition score were recorded. Binary scores were also generated to identify heifers which had a CL identified at any time prior to the start of joining (CLPRIOR), and of these animals, a subset were identified as displaying a CL at the scanning closest to joining (CLJOIN).

Data for four tenderness markers (T1, T2, T3 and T4) were analysed for this experiment. Marker genotypes were generated by Catapult Genetics®, and reported as expressing 0, 1 or 2 copies of the favourable allele for each marker. See Johnston and Graser (2008) for details of phenotypic and tenderness marker data.

Statistical analysis. The significance and effect of tenderness markers on feedlot exit, carcass and heifer puberty traits were analyzed by individually including marker genotype (0, 1 or 2) in models for each trait, containing significant fixed effects, and with sire fitted as random, using PROC GLM in SAS (SAS Institute Inc.: Cary, NC). Wolcott *et al.* (2009), Barwick *et al.* (2009) and Johnston *et al.* (2009) provided details of fixed effect modeling for carcass and meat quality, steer feedlot exit, and heifer puberty traits respectively. Marker genotype was fitted as a continuous variable. Significance levels were re-estimated applying the principles described by Benjamini and Hochberg (1995) to account for the potential errors associated with multiple testing, and accepting a 5% false discovery rate. For each marker, animals with missing genotypes were excluded from the analysis.

RESULTS AND DISCUSSION

Table 1 presents the number of animals analysed for each marker, and the genotype frequencies for tenderness markers. Genotype frequencies did not differ significantly for steers and heifers (Johnston and Graser, 2008) and the results presented in Table 1 are therefore pooled across sexes. T2 was virtually fixed at the unfavourable homozygous genotype, with the

favourable allele occurring in only 3.6% of animals. For T3, the favourable allele was also present at a low frequency (15.9%), though it was only the T2 allele frequency which deviated significantly from the expectations of Hardy-Weinberg equilibrium ($P < 0.001$).

Table 1. Number of animals (pooled across sexes) displaying 0, 1 or 2 copies of the favourable allele for GeneSTAR tenderness markers T1 – T4

Genotype	T1	T2	T3	T4
0	203	1277	1248	554
1	831	76	463	845
2	726	11	49	361

Table 2 presents significant regression coefficients, their standard errors, number of observations analysed, and significance levels for the relationship between marker genotypes and steer feedlot exit and carcass traits, and heifer puberty measurements. For the majority of traits measured, there was no significant ($P > 0.05$) effect of the markers. There was, however, a significant effect of the T1, T3 and T4 markers on fat depth or fat related traits in steers carcass and feedlot exit measurements. The regression coefficients of T1, T3 and T4 marker genotypes for fat depths (HP8 and P8X) were consistently negative, showing that steers which had more copies of the favourable tenderness alleles were significantly leaner. The consistency of the direction of the effect across markers and fat measurements provides some confidence that the effect is real. The T4 marker also had a significant and negative relationship with HUMP, which was positively genetically related to both SF and P8 fat depth for these animals ($r_g = 0.19$ and 0.31 respectively, see Wolcott *et al*. 2009). The significant relationship between T4 and IGFX in steers is also likely to reflect a strong positive genetic correlation between IGF-I measured at feedlot exit and shear force ($r_g = 0.59$), as reported by Wolcott *et al*. (2009).

Table 2. Number of observations (N), means and standard deviations (SD), regression coefficients (b) and standard errors (SE), and P-values describing significant relationships between tenderness markers (T1 – T4) and steer feedlot exit and carcass, and heifer puberty traits in Brahman cattle

Marker	Trait	Units	N	Mean	SD	b	S.E.	P - value
Steer carcass and feedlot exit traits								
T1	HP8	mm	853	13.6	4.0	-0.49	0.22	0.0260
T3	P8X	mm	787	11.9	3.4	-0.71	0.22	0.0014*
T3	CIMF	%	708	2.2	0.8	0.11	0.05	0.0359
T3	MS	Score	842	264.8	64.3	-9.10	4.27	0.0334
T4	HUMP	mm	786	166.8	34.8	-4.45	1.81	0.0144
T4	IGFX	ng/ml	636	584.2	132.9	16.35	6.96	0.0192
T4	HP8	mm	855	13.6	4.0	-0.48	0.21	0.0242
Heifer puberty traits								
T1	AGECL	Days	907	746.5	141.7	-14.54	5.85	0.0130
T1	CLPRIOR	1=yes, 0=no	885	0.43	0.49	0.049	0.02	0.0221

* Relationship maintained significance under multiple testing protocols at $P < 0.0018$.

The significant relationship between the T3 marker and marbling measurements is difficult to interpret, though differences in numbers of measurements available for CIMF and MS may have

contributed to the results. Genetically CIMF and MS are highly correlated (r_g = 0.95), though at the phenotypic level the relationship was weaker (r_p = 0.50) (Wolcott *et al.* 2009). That CIMF and MS should have regression coefficients in opposite directions is likely to reflect the marginal significance levels of T3 genotype as predictors of the traits (P = 0.036 and 0.033 for CIMF and MS respectively). These results must therefore be considered with the caveat that further testing would be desirable, before conclusions are drawn.

The significant relationships between T1 genotype, and AGECL and CLPRIOR, suggested that heifers with more copies of the favourable T1 allele, were more likely to reached puberty earlier (b = -14.54), and display a CL prior to joining (b = 0.049). Results presented by Johnston *et al*. (2009) suggested that there was little evidence of genetic antagonism between tenderness and heifer puberty traits, with genetic correlations between shear force and AGECL and CLPRIOR of -0.16 and 0.11 respectively. The directions of these correlations were in contrast to the regression coefficients presented for the current study, but must be considered in association with the report of Johnston and Graser (2008) which demonstrated marginal significance (P = 0.068) of T1 genotype as a predictor of tenderness in the Brahman cattle used for this experiment.

Applying the multiple testing principles of Benjamini and Hochberg (1995), the regression of T3 genotype on P8X was significant at a re-estimated threshold of 0.0018, while P-values for the remaining relationships did not meet these more stringent requirements. This suggests that the effect of improved T3 genotype on fatness in BRAH steers at feedlot exit was real, and that this effect must be considered when Brahman breeders select animals based on the T3 marker.

CONCLUSIONS

Genetic markers have been shown to be significantly related to tenderness in Brahman cattle. This study has demonstrated that, for most of the feedlot exit, carcass and heifer puberty traits examined, tenderness marker genotype had no significant effect on phenotypic performance. For some fat depth and marbling traits measured in steers, and age at puberty in heifers, however, there was a significant (P < 0.05) association with T1, T3 or T4 markers. This suggests that if BREEDPLAN is to incorporate tenderness marker results into the Brahman multi-trait evaluation, provision will need to be made for their relationship with steer fatness and marbling, and heifer age at puberty. More generally, as the technology develops and larger panels of markers are released, it will be important to estimate the relationships of these markers with other traits, before they are incorporated into the BREEDPLAN multi-trait evaluation. Future work in this area will expand the markers tested to include those for marbling and feed efficiency, as well as fitting combinations of markers to test for possible epistatic effects.

REFERENCES

Barwick, S.A., Wolcott, M.L., Johnston, D.J., Burrow, H.M. and Sullivan, M.T. (2009) *Anim. Prod. Sci*. **49**:351.

Benjamini, Y. and Hochberg, Y. (1995) J. R. Statist. Soc. B. **57**:289.

Egan, A.F., Ferguson, D.M. and Thompson, J.M. (2001) *Aust. J. Agric. Res*. **41**:855.

Johnston, D.J., Barwick, S.A., Holroyd, R.G., Fordyce, G., Wolcott, M.L. and Burrow, H.M. (2009) *Anim. Prod. Sci.* **49**:399.

Johnston, D.J. and Graser, H.U. (2008) SmartGene for Beef: Effect of marker genotype on phenotypic performance. http://agbu.une.edu.au/Smartgene Final AGBU Report2.pdf.

Perry, D., Thompson, J.M., Hwang, I.H., Butchers, A. and Egan, A.F. (2001) *Aust. J. Exp. Agric.* **41**:981.

Wolcott, M.L., Johnston, D.J., Barwick, S.A., Iker, C.L., Thompson, J.M. and Burrow, H.M. (2009) *Anim. Prod. Sci.* **49**:383.

Proc. Assoc. Advmt. Anim. Breed. Genet. 18:163-166

ASSOCIATION BETWEEN MYOSTATIN DNA MARKERS AND MUSCULARITY IN ANGUS CATTLE

B.A. O'Rourke[1], P.L. Greenwood[2], P.F. Arthur[1] and M.E. Goddard[3]

Cooperative Research Centre for Beef Genetic Technologies, Armidale NSW 2351, Australia
[1]NSW Department of Primary Industries, Elizabeth Macarthur Agricultural Institute, Camden NSW 2570, Australia
[2]NSW Department of Primary Industries, Beef Industry Centre, Armidale NSW 2351, Australia
[3]Victorian Department of Primary Industries, Victorian Agribiosciences Centre, Bundoora Vic 3083, Australia

SUMMARY

Myostatin (MSTN) is a potent negative regulator of skeletal muscle development. High genetic variability has been observed in *bovine MSTN* which includes 6 specific disruptive mutations responsible for extreme muscular hypertrophy in cattle. In this study the effect of non-disruptive *MSTN* polymorphisms on muscularity was examined in a population of 594 Angus cattle. Six tag SNP (single nucleotide polymorphism), which included 5 non-disruptive SNP and the disruptive 821 del11 mutation, were genotyped in each animal and haplotype phase was inferred. Eleven haplotypes were found in the Angus population and the muscular hypertrophy marker (821del11) was confined to 2 haplotypes (8 and 11). Association between *MSTN* haplotypes and eye muscle area (EMA) at weaning age was tested using multiple linear regression. In the regression analysis comprising all cattle, haplotypes 1, 2, 4, 7 and 9 had significant regression coefficients ($P<0.05$) relative to haplotype 5. These haplotype associations were confirmed in a second analysis that contained only cattle without the 821del11 muscular hypertrophy marker ($n=528$). These results indicate that other *MSTN* DNA markers, when assessed as haplotypes, are associated with variation in EMA and therefore contribute to differences in muscularity.

INTRODUCTION

Myostatin (MSTN), a secreted protein, is a member of the transforming growth factor – β superfamily. Loss of MSTN function causes large increases in muscle mass and hence, MSTN is regarded as a potent negative regulator of skeletal muscle mass (McPherron *et al.* 1997). This large increase in muscle mass is termed muscular hypertrophy or double muscling. Naturally occurring mutations in the *MSTN* gene that are implicated in double muscling have been reported in cattle (Grobet *et al.* 1998), sheep (Clop *et al.* 2006), humans (Schuelke *et al.* 2004) and dogs (Mosher *et al.* 2007).

Analysis of the bovine *MSTN* sequence prompted by the genetically heterogeneous nature of double muscling, has also uncovered a series of non-disruptive polymorphisms (Grobet *et al.* 1998; Crisa *et al.* 2003; Dunner *et al.* 2003; O'Rourke *et al.* 2009). These studies show considerable genetic diversity within *MSTN*, which may also contribute to variation in muscle mass.

In a previous study, 18 *MSTN* DNA markers, including the 821del11 double muscling marker, were identified in a sub-group of Angus cattle (O'Rourke *et al.* 2007). The purpose of this study was to examine the effects of non-disruptive *MSTN* polymorphisms on muscularity in a larger population of Angus cattle. We tested the null hypothesis that only the double muscling marker 821del11 was contributing to variation in muscularity.

MATERIALS AND METHODS

Data and DNA samples for 594 Angus cattle born between 1998 and 2006 from a NSW Department of Primary Industries Research Herd at Glen Innes, Australia were used in this study. This herd was established for research purposes in 1988 and comprised high and low muscle selection lines. Selection was based on muscle score (McKiernan, 1990) as assessed at weaning. Of the 594 cattle used, 324 (198 female and 126 male) were classified as high muscle and 270 (154 female and 116 male) were from the low muscle line. Measurements for eye muscle area (EMA) were taken at weaning age (approximately 9 months of age) by experienced and/or accredited technicians using real-time ultrasound 3.5 MHz/180-mm linear array animal science probe (Esoate Pie Medical, Maastricht, Netherlands).

A tag SNP genotyping approach was employed to determine genotypes at 6 *MSTN* polymorphic sites. The tag SNP included 5 non-disruptive polymorphisms (2 promoter polymorphisms, 1 in intron 1, and 2 in the 3' untranslated region) that did not alter the length of the *MSTN* coding region and 1 disruptive mutation (821 del11) in Exon 3; a frameshift mutation, which introduces a premature stop codon. All animals with the 821del11 mutation were heterozygous at this site; no 821 del11 homozygotes were included in the study. A polymerase chain reaction/restricted fragment length polymorphism method was used for genotyping the promoter and 3' untranslated region polymorphisms and a primer extension methodology was used for the intron 1 polymorphism. Genotypes at the 821 del11 site were determined by real time PCR (O'Rourke *et al.* 2009). At each polymorphic site the allele differing from the GenBank reference database sequences AF320998 and AF348479 was designated as the mutated allele.

Haplotype phase was inferred from the genotypic data using PHASE v2.1.1 (Stephens et al. 2001; Stephens and Scheet 2005). Ambiguous genotypes were also inferred using PHASE v2.1.1. The association between each haplotype and eye muscle area (EMA) at weaning age was tested using multiple linear regression (SAS 9.1.3; SAS Institute). The statistical model accounted for the main effects of sex, muscling selection line and birth year. Interaction effects were not fitted due to low sub-class observations. Sire was included as a random term and weight at the time of measurement was used as a covariate. Haplotype 5, which does not contain the disruptive 821del11 mutation was used as the reference haplotype (regression coefficient = 0) in the analysis. The association analysis initially included all animals in the cattle population that had either 0 (n=528) or 1 (n=66) copy of the 821 del11 mutation. A second analysis was performed which excluded the 821 del11 heterozygotes.

RESULTS AND DISCUSSION

In this study an Angus cattle population was genotyped at 6 *MSTN* polymorphic sites, which included the 821 del11 disruptive mutation in Exon 3, historically associated with double muscling in Belgian Blue cattle (Grobet *et al.* 1997). The 6 sites were selected using a tag SNP approach from a total of 18 *MSTN* sites previously found in a sub-group of this cattle population (O'Rourke *et al*. 2007). Haplotype phase was inferred for each animal and 11 haplotypes were identified (Table 1). The disruptive 821 del11 mutation was confined to haplotypes 8 and 11. The low frequency of haplotype 8 prompted confirmation of genotypes for this animal. Parent genotypes for the 821del11 mutation indicated that the mutation had been inherited from the sire (heterozygous for haplotype 11) and parentage was confirmed by DNA testing. We have therefore deduced that haplotype 8 has arisen from recombination of the paternal gamete indicating that haplotype 11 is the ancestral double muscling haplotype in this cattle population. Haplotype 10, also in low frequency, was confirmed as the most likely allele combination by pedigree analysis. Moderate to high frequency was observed for the other haplotypes with haplotype 7 the most prevalent.

Table 1. Haplotype diversity in the *myostatin* gene for 594 Angus cattle

Haplotype	***N***	tag SNP					
		1	2	3	4	5	6
1	162			+			
2	57						+
3	75					+	
4	136	+		+			
5	68	+					+
6	32	+				+	+
7	555	+				+	
8	1	+			+		
9	35	+	+				+
10	2	+	+			+	
11	65	+	+		+		

+ indicates the presence of the mutant allele with respect to the reference sequences AF320998 and AF348479. Tag SNP 1 and 2 are the promoter polymorphisms, site 3 is in intron 1, site 4 is the 821 del11 mutation in exon 3, and 5 and 6 are in the 3' untranslated region. *N* = haplotype observations (2/animal)

The association between *MSTN* haplotype and ultrasound measurements for eye muscle area was determined by multiple linear regression (Table 2). The alleles were assessed collectively as haplotypes offering greater power for association studies, particularly for complex traits where many markers of small effect and few with large effect may be linked to phenotypic variation (Hayes and Goddard 2001). Initially, the association between haplotype and EMA was tested in the entire cattle population. Relative to haplotype 5, all haplotypes except for haplotype 11 had a negative regression coefficient, suggesting haplotype 5 is associated with the second largest EMA. The positive regression coefficient for haplotype 11 and its association with the largest EMA was expected since this group contains all but one of the 821 del1l heterozygotes, and therefore adds a quality control aspect to the analysis. Haplotypes 9, 2, 1, 4 and 7 ranked in order of their regression coefficients (*b*), showed significantly less EMA compared to haplotype 5 ($P<0.05$).

Table 2. Haplotype association with eye muscle area (EMA, cm^2) at weaning age in Angus cattle

Haplotype	All cattle			Cattle without 821 del11 marker		
	N	***b*** ± s.e.	***P***	***N***	***b*** ± s.e.	***P***
1	162	-2.68 ± 0.85	0.0018	154	-3.01 ± 0.89	0.0008
2	57	-3.74 ± 1.37	0.0066	57	-4.13 ± 1.37	0.0027
3	75	-0.63 ± 0.96	0.5131	69	-0.84 ± 1.01	0.4085
4	136	-1.89 ± 0.89	0.0336	133	-2.12 ± 0.92	0.0215
5	68	0	-	64	0	-
6	32	-0.82 ± 1.16	0.4797	31	-0.96 ± 1.20	0.4259
7	555	-1.79 ± 0.76	0.0184	516	-1.81 ± 0.79	0.0227
8	1	-1.05 ± 5.14	0.8386	0	-	-
9	35	-4.14 ± 1.16	0.0004	30	-4.80 ± 1.26	0.0001
10	2	-2.53 ± 3.67	0.4901	2	-2.43 ± 3.73	0.5149
11	65	1.43 ± 1.08	0.1865	0	-	-

[a]All haplotypes are relative to haplotype 5. *N*, number of haplotype observations (2/animal); *b*, EMA regression coefficient.

The double muscled phenotype occurs in animals homozygous for a disruptive *MSTN* mutation. Double muscled cattle can have up to a 20% increase in muscle mass indicating that these disruptive mutations have a large effect on muscularity (Grobet *et al.* 1997). The partially recessive mode of inheritance for these mutations means that in the heterozygous form, significant increases in muscling are also observed (Gill *et al.* 2008; O'Rourke *et al.* 2009). In this study, the initial association analysis included cattle heterozygous for the 821 del11 mutation. To determine if the original haplotype associations were confounded by the inclusion of a double muscling marker, the analysis was repeated in a reduced population which excluded all 821 del11 heterozygotes (haplotypes 8 and 11; Table 2). The results again showed that haplotype 5 was associated with the largest EMA for each the 'functional' haplotypes, and haplotypes 9, 2, 1, 4 and 7 remained significant ($P<0.05$) and their ranking relative to haplotype 5 had not altered.

In conclusion, the results presented in this study indicate that other *MSTN* polymorphisms that have not been implicated in double muscling are associated with variation in muscularity.

ACKNOWLEDGEMENTS

This study was funded by Meat and Livestock Australia, the Cooperative Research Centre for Beef Genetic Technologies, NSW Department of Primary Industries and the University of Melbourne.

REFERENCES

Clop, A., Marcq, F., Takeda, H., Pirottin, D., Tordoir, X., Bibe, B., Bouix, J., Caiment, F., Elsen, J.M., Eychenne, F., Larzul, C., Laville, E., Meish, F., Milenkovic, D., Tobin, J., Charlier, C. and Georges, M. (2006) *Nat. Genet.* **38**:813.

Crisa, A., Marchitelli, C., Savarese, M.C. and Valentini, A. (2003) *Cytogenet. Genome Res.* **102**:48.

Dunner, S., Miranda, M.E., Amigues, Y., Canon, J., Georges, M., Hanset, R., Williams, J. and Menissier, F. (2003) *Genet. Sel. Evol.* **35**:103.

Gill, J.L., Bishop, S.C., McCorquodale, C., Williams, J.L. and Wiener, P. (2008) *Anim. Genet* **40**:100.

Grobet, L., Martin, L.J., Poncelet, D., Pirottin, D., Brouwers, B., Riquet, J., Schoeberlein, A., Dunner, S., Menissier, F., Massabanda, J., Fries, R., Hanset, R. and Georges, M. (1997) *Nat. Genet.* **17**:71.

Grobet, L., Poncelet, D., Royo, L.J., Brouwers, B., Pirottin, D., Michaux, C., Menissier, F., Zanotti, M., Dunner, S. and Georges, M. (1998) *Mamm. Genome* **9**:210.

Hayes, B.J. and Goddard, M.E. (2001) *Genet. Sel. Evol.* **33**:209.

Kijas, J., McCulloch, R., Edwards, J., Oddy, V.H., Lee, S. and vander Wer,f J. (2007) *BMC Genet.* **8**:80.

McKiernan, W.A. (1990) *Proc. Ninth Conf. AAABG.* **9**:447.

McPherron, A.C., Lawler, A.M. and Lee,S.J. (1997) *Nature* **387:**83.

Mosher, D.S., Quignon, P., Bustamante, C.D., Sutter, N.B., Mellersh, C.S., Parker, H.G. and Ostrander, E.A. (2007) *PLoS Genet.* **3**:79.

O'Rourke, B.A., Dennis, J.A., Healy, P.J., McKiernan, W.A., Greenwood, P.L., Café, L.M., Perry, D., Walker, K.H., Marsh, I., Parnell, P.F. and Arthu,r P.F. (2009) *Anim. Prod. Sci.* **49**:297.

O'Rourke, B.A., Hayes, B.J., Greenwood, P.L., Arthur, P.F. and Goddard ,M.E. (2007) *Proc. Seventeenth Conf. AAABG* **17**:135.

Saunders ,M.A., Good, J.M., Lawrence, E.C., Ferrell ,R.E., Li W.H. and Nachman, M.W. (2006) *Am. J. Hum.Genet.* **79**:1089.

Schuelke, M., Wagner, K.R., Stolz, L.E., Hubner, C., Riebel, T., Komen, W., Braun, T., Tobin, J.F. and Lee, S.J. (2004) *New Engl. J. Med.* **350**:2682.

Proc. Assoc. Advmt. Anim. Breed. Genet. 18:167-170

MUSCLE SPECIFIC EXPRESSION OF REGULATORY FACTORS IN CATTLE SELECTED FOR HIGH AND LOW MUSCLING

G. Parnell[1], Y. Chen[2], G.S. Nattrass[3], P.L. and Greenwood[2]

Cooperative Research Centre for Beef Genetic Technologies. [1] University of New England, Armidale, NSW 2351. [2] NSW Department of Primary Industries, Beef Industry Centre, UNE, Armidale, NSW 2351. [3] SARDI - Livestock and Farming Systems, Roseworthy, SA 5371

SUMMARY

This paper reports an investigation of the allele-specific expression of myostatin in the *semitendinosus* and *longissimus dorsi* muscles in a population of animals selected for high or low muscling, including animals heterozygous for the *nt821(del11)* loss of function myostatin polymorphism. In addition, expression of follistatin, myogenin, and *MYOD*, genes that also affect muscle growth and development were studied. Animals that were heterozygous for the *nt(821)del11* loss of function polymorphism expressed higher amounts of total myostatin but lower amounts of the functional (wild-type) allele compared to homozygous wild-type animals. The level of *MYOD* expression was greater in the wild-type high muscling line compared to the wild-type low muscling line. These findings demonstrate an up-regulation of total myostatin expression in cattle heterozygous for a non-functional myostatin allele compared to homozygous wild-type cattle, presumably due to the role of negative feedback in these cattle which express less wild-type myostatin than their homozygous wild-type counterparts. The findings also show that selection for divergence in muscling score can influence expression levels of other muscle regulatory genes such as *MYOD*.

INTRODUCTION

There is continuing interest in the use of the double muscled phenotype in cattle as a method of increasing the amount of saleable beef, or retail beef yield from carcases (O'Rourke *et al.* 2006, 2009). However, the double muscled phenotype of cattle is also associated with negative attributes such as decreased fertility and calving difficulties (Arthur 1995), hence less extreme phenotypes with increased muscling but without fertility and calving difficulties are more desirable for use in the beef industry.

Several mutations have been identified within the *bovine* myostatin gene that disrupt the function of the myostatin protein and, therefore, give rise to the double muscling phenotype. One such mutation, common in the Belgian Blue breed, is caused by an 11-bp deletion of nucleotides 821-831 inclusive and is referred to as *nt821(del11)*. The *nt821(del11)* polymorphism results in a frame shift and subsequent premature stop codon in the bioactive carboxy-terminal domain of the gene, a domain which is highly conserved in the TGF-β superfamily (Grobet *et al*. 1997).

The objective of this project was to assess if there is any differences in the level of gene expression of myostatin and other genes that also affect muscle growth and development in animals selected for differences in their levels of muscling, including a genotype heterozygous for the *nt821(del11)* mutation .

MATERIALS AND METHODS

The experimental animals were obtained from the NSW Department of Primary Industries low and high muscling selection lines. These lines originated at the Elizabeth Macarthur Agricultural Institute, Camden, NSW in 1998 from 140 females selected from 260 Angus x Hereford F1 female progeny based on muscle score (O'Rourke *et al*. 2006, 2009). Allocation to the high and low lines was based on individual muscle scores. Further divergence of the selection lines was achieved by

mating the high muscle line females with high muscle score bulls and the low muscle line females with low muscle score bulls. Forty-four 2003-born steers were used to study the effects of selection for muscling and of the polymorphisms in the myostatin gene. The steers were divided into 3 separate groups for the analyses: high muscling line with wild-type myostatin (n=14), high muscling line heterozygous for the *nt821(del11)* myostatin polymorphism (n=11), and low muscling line with wild-type myostatin (n=19). The steers were slaughtered at 25 months of age and live weight, carcass, and yield characteristics measured, as reported previously (Cafe *et al*. 2006). At slaughter, *semitendinosus* and *longissimus dorsi* muscle samples were collected, snap frozen in liquid nitrogen, and stored at - 80°C.

Total RNA extraction from muscle was carried out by homogenising 100mg of tissue in Tri-reagent and isolating the RNA as per the manufacturer's instructions (Ambion Inc., USA). RNA cleanup and an on-column DNase treatment of extracted RNA was carried out to remove residual genomic DNA and other non-RNA impurities with RNesay Mini columns (Qiagen, Germany). First-strand cDNA synthesis was carried out with the Omniscript cDNA synthesis kit (Qiagen, Germany) with the addition of an 18SrRNA gene-specific primer (5'-ACACGCTGAGCCAGTCAGT-3'). Quantification of gene expression was carried out on a Rotorgene 3000 (Corbett Research, Australia). Each sample was assayed in triplicate. Real-time PCR reactions were carried out using a reagent containing 0.5 units Amplitaq Gold DNA polymerase (Applied Biosystems, USA), 1 x PCR Gold buffer (Applied Biosystems, USA), 200nM dNTPs (Invitrogen, USA) and 1 x SYTO9 fluorescent dye (Invitrogen, USA). Quantification of gene expression was carried out for total myostatin, *nt821(del11)* myostatin, wild-type myostatin, myogenin, *MYOD*, and follistatin. Reference genes assayed were *RPL19* and *18SrRNA*. Table 1 shows the primer sequences, the length of the amplified products and the Genbank accession numbers for each of the primers used.

Table 1. Forward and reverse primers for real-time PCR assays

Gene	Forward primer (5' – 3')	Reverse primer (5' – 3')	Length (bp)	Genbank accession no.
Total Myostatin	accttcccagaaccaggagaa	tcacaatcaagcccaaaatctct	101	AF019622
Myostatin wild-type allele	tcttgctgtaaccttcccagaac	acagcatcgagattctgtggagt	124	AF019622
Myostatin *nt821(del 11)* allele	tcttgctgtaaccttcccagaac	acgacagcatcgagattctgtca	121	AF019622
Myogenin	ggcagcgcactggagttt	ccgctgggagcagatgat	52	AF433651
MYOD	aactgttccgacggcatgat	gacaccgcagcgctcttc	128	X62102
Follistatin	gggcagatctattggattgg	cctctgccaaccttgaagtc	114	BC133637
18SrRNA	cggtcggcgtcccccaactt	gcgtgcagccccggacatctaa	103	M10098
RPL19	caactcccgccagcagat	ccgggaatggacagtcaca	76	AY158223

For each gene, cycle thresholds (Cts) were determined for the cDNA samples and a standard curve generated from 7 consecutive 2-fold dilutions of pooled cDNA. Normalised relative quantitation was carried out using qBase (Hellemans *et al*. 2007). One-way Analysis of Variance (ANOVA) of the normalised real-time PCR data was carried out for the 2 muscle types and the 3 muscling lines within muscle type, using the statistical package R (R Foundation for Statistical Computing, Austria). Tukey multiple comparison tests were carried out if a significant ANOVA result was observed for a particular muscle type.

RESULTS AND DISCUSSION

Mean normalised relative expression of follisatin, myogenin, *MYOD*, total myostatin, myostatin wild-type, *nt821(del11)* myostatin are presented in Table 2. There were no significant differences between muscles or between the muscling lines for the reference genes.

Table 2. Mean (S.E.) normalised relative expression of genes for the three muscling lines

Gene	High muscling wild-type (n=14)	High muscling Heterozygote (n=11)	Low muscling wild-type (n=19)
	M. longissimus dorsi		
Follistatin	3.13 (0.260)	3.10 (0.270)	3.21 (0.200)
Myogenin	6.62 (0.584)	6.41 (0.667)	6.82 (0.460)
MYOD	3.88 (0.363)	4.21 (0.604)	2.94 (0.260)
Total Myostatin	3.90 (0.350)	4.48 (0.509)	3.20 (0.352)
Wild-type Myostatin	3.87 (0.366) a	2.25 (0.239) b	3.17 (0.359) ab
nt821(del11) Myostatin	na	2.26 (0.199)	na
	M. semitendinosus		
Follistatin	3.32 (0.338) a	2.25 (0.222) b	3.07 (0.226) ab
Myogenin	4.19 (0.313)	3.26 (0.495)	4.05 (0.388)
MYOD	3.97 (0.264) a	3.09 (0.382) ab	2.77 (0.254) b
Total Myostatin	7.50 (0.521) a	9.82 (0.654) b	7.64 (0.562) a
Wild-type Myostatin	7.17 (0.555) a	4.86 (0.496) b	7.58 (0.578) a
nt821(del11) Myostatin	na	4.79 (0.198)	na

Within rows, mean values with different letters differ significantly ($P<0.05$). na = not applicable.

Myostatin. Expression of total myostatin, myostatin wild-type, and myostatin *nt821(del11)* was higher in *M. semitendinosus* than in *M. longissimus dorsi* ($P<0.0001$) in all 3 muscling lines. This finding is consistent with previous studies that have shown greater myostatin expression in muscle tissue with a higher proportion of glycolytic fibres (Bass *et al.* 1999). There was no significant difference in expression of the myostatin wild-type allele between the high muscling wild-type and low muscling wild-type groups for both *M. semitendinosus* ($P= 0.85$) and *M. longissimus dorsi* ($P=0.30$). Total myostatin expression was higher in heterozygotes than homozygous wild-type animals. Despite the differences in the range of expression levels there was no significant difference in mean expression of the myostatin *nt821(del11)* and wild-type alleles in either the *M. semitendinosus* ($P=0.89$) or *M. longissimus dorsi* ($P=0.98$) from the heterozygote animals. However, heterozygote animals expressed lower levels of the functional wild-type allele than the homozygous normal animals. Since the myostatin *nt821(del11)* heterozygote line has only one copy of the wild-type allele, it would be expected that cattle from this line would express this allele at half the level as that of the wild-type lines. However, the level of expression of the wild-type allele in the myostatin *nt821(del11)* heterozygote cattle was greater than half that of the high muscling wild-type and the low muscling wild-type lines. This finding was more pronounced for *M. semitendinosus* than *M. longissimus dorsi.*

MYOD. There was no significant difference in the expression of *MYOD* between the two muscle types. There was however, a difference within the three muscling lines in expression in *M. semitendinosus*. The expression of *MYOD* in *M. semitendinosus* was significantly higher in the

high muscling wild type line than the low muscling wild type line (*P*=0.007). *MYOD* is a marker of satellite cell proliferation (Grounds *et al*. 1992) and increased *MYOD* expression appears to have been indirectly impacted by selection for high muscling. Hence, this may indicate that the low muscling wild-type line had less satellite cell activity than the high muscling wild-type line at the time of sampling.

Myogenin. Myogenin expression was higher in *M. longissimus dorsi* than *M. semitendinosus* (P<0.0001), but there were no significance differences between the three muscling lines.

Follistatin. Follistatin expression tended to be lower in the heterozygote animals compared to the high muscling wild-type animals, although the only significant difference was due to higher expression in the high muscling wild-type line compared with the high muscling heterozygote line for the *M. semitendinosus* (Table 2, *P*=0.049). Along with the findings for myostatin, this suggests a possible negative feedback loop whereby the lack of functional myostatin transcript in the heterozygotes results in an up-regulation of transcription of both the functional and non-functional myostatin alleles. The regulatory trigger for expression of myostatin may be myostatin itself, as found for transforming growth factor-β (TGF-β) by Kim *et al*. (1990). Alternatively, an intracellular/autocrine feedback loop may operate, as suggested for interactions between insulin-like growth factors and TGF-β (Bosche *et al*. 1995). It is possible that an endocrine growth factor stimulated by myostatin may then regulate expression of myostatin mRNA (Oldham *et al*. 2001).

CONCLUSIONS

There is an up-regulation of total myostatin expression in cattle heterozygous for a non-functional myostatin allele compared to homozygous wild-type cattle. Selection for divergence in muscling score can influence expression levels of other muscle regulatory genes such as *MYOD*.

REFERENCES

Arthur, P.F. (1995) *Aust. J. Agric. Res*. **46**:1493.

Bass, J., Oldham, J., Sharma, M. and Kambadur, R. (1999). *Anim. Endocrinol*. **17**:191.

Bosche, W.J., Ewton, D.Z., Florini, J.R. **(**1995) *J. Cell Physiol.* **164**:324.

Cafe, L., O'Rourke, B., McKiernan, W. and Greenwood, P. (2006) *Proc. Aust. Soc. Anim. Prod.* **26**: SC

Grobet, L., Martin, L.J., Poncelet, D., Pirottin, D., Brouwers, B., Riquet, J., Schoeberlein, A., Dunner, S., Menissier, F., Massabanda, J., Fries, R., Hanset, R. and Georges, M. (1997) *Nat. Genet.* **17**:71.

Grounds, M.D., Garrett, K.L., Lai, M.C., Wright, W.E. and Beilharz, M.W. (1992) *Cell Tiss. Res.* **267**: 99.

Hellemans, J., Mortier, G., De Paepe, A., Speleman, F. and Vandesompele, J. (2007) *Genome Biol.* **8**:R19.

Kim, S.J., Angel, P., Lafyatis, R., Hattori, K., Kim, K.Y., Sporn, M.B., Karin, M. and Roberts, A.B. (1990) *Mol. Cell Biol.* **10**:1492.

Oldham, J. M., Martyn, J. A. K., Sharma, M., Jeanplong, F., Kambadur, R. and Bass, J. J. (2001) *Am. J. Physiol (Regul. Integr. Comp. Physiol.)* **280**:R1488.

O'Rourke, B.A., Dennis, J.A., McKiernan, W.A., Greenwood, P.L., Cafe L.M., Perry D., Arthur P.F., Parnell P.F. and Walker K.H. (2006) Myostatin alleles as production markers, Final report to Meat and Livestock Australia.

O'Rourke, B.A., Dennis, J.A., Healy, P.J., McKiernan, W.A., Greenwood, P.L., Cafe, L.M., Perry, D., Walker, K.H., Marsh, I., Parnell, P.F. and Arthur, P.F. (2009) *Anim. Prod. Sci.* **49**: (In Press)

Proc. Assoc. Advmt. Anim. Breed. Genet. 18:171-174

GLOBAL GENE EXPRESSION PROFILING OF ANGUS CATTLE SELECTED FOR LOW AND HIGH NET FEED INTAKE

Y. Chen[1], C. Gondro[2], K. Quinn[1], B. Vanselow[1], P.F. Parnell[1] and R.M. Herd[1]

Cooperative Research Centre for Beef Genetic Technologies
[1] NSW Department of Primary Industries, Beef Industry Centre, Armidale, NSW 2351
[2] The Institute for Genetics and Bioinformatics, University of New England, Armidale, NSW 2351

SUMMARY

Feed efficiency measured as net feed intake (NFI) is the difference between the actual feed intake by an animal over a test period and its expected feed intake based on its size and growth rate. The experiment reported here aimed to identify differentially expressed genes between animals with low and high NFI, and pathways which contribute to the phenotype, by global gene expression profiling using a 24K bovine long-oligo array. Liver tissue biopsies were taken from the top 30 and bottom 30 NFI-ranked bulls following their NFI test. The bulls were from lines of Angus cattle divergently selected for low or high NFI, and 44 animals from the sixty sampled were chosen for the microarray experiment. One hundred and eighty-one probes were identified as differentially expressed between liver samples of low and high NFI animals by microarray data analysis with a cut-off threshold of $P<0.01$. Gene ontology analysis revealed that 86% of the up-regulated genes were involved in known biological processes, 92% have a known molecular function assigned and 84% related to cellular components. Among the down-regulated genes, 83% have a known molecular function, 80% are involved in known biological processes and 78% are components.

INTRODUCTION

Net feed intake (NFI) has been adopted in Australia for measurement of feed efficiency in beef cattle for the purpose of genetic improvement. It is the difference between an animal's actual feed intake over a test period and its expected feed intake based on its size and growth rate. As NFI is phenotypically independent of production level and metabolic weight it may reflect differences in efficiency in basic metabolic processes (Archer *et al*. 1999). Although several QTL (quantitative trait loci) have been identified by traditional linkage mapping and fine mapping, and 160 candidate SNPs (single nucleotide polymorphisms) identified by whole-genome association studies (Barendse *et al*. 2007; Nkrumah *et al*. 2007; Sherman *et al*. 2008), the key physiological systems important to variation in NFI are not well understood, and the genes involved in these processes and their functions in relation to NFI are largely unknown.

The objectives of this study was to examine the global gene expression pattern in cattle from low and high NFI selection lines in order to identify genes and critical pathways affecting feed efficiency as determined using NFI. This information will be used to inform the search for candidate genes for marker-assisted selection for NFI and avenues for alternative non-genetic methods which might be used to manipulate metabolism and net feed efficiency of cattle.

MATERIALS AND METHODS

Animals. Angus cattle selection lines for low and high NFI were established in 1993 at the Agricultural Research Centre, Trangie, NSW, Australia. Ninety bulls born in 2005, after approximately three generations of selection, were reared on pasture before starting a post-weaning NFI test at approximately 300kg live weight. Feed intake was measured for each animal using an automated recording system over a standard 70-day NFI test at the Beef CRC "Tullimba"

Research Feedlot. Based on their test NFI, 30 animals with the lowest NFI and 30 animals with the highest NFI were selected for collection of liver biopsies.

Biopsy sampling and total RNA extraction. This experiment was approved by the University of New England Animal Ethics Committee (AEC 06/123) and followed the University of New England code of conduct for research with Animals. Liver biopsy was performed according to the protocol of Davies and Jebbettt (1981). Fresh liver tissues were quickly immersed in 2.5mL RNAlater solution (Ambion, Applied Biosystem). Total RNA was isolated using TRI reagent (Ambion, Applied Biosystems) according to the manufacturer's instructions.

Bovine long-oligonucleotide array and experiment design. The 24,000 long-oligonucleotide array developed by the Bovine Oligo Microarray Consortium (BOMC) consists of: 16,846 probes designed from expressed sequence tags (ESTs) that were aligned to homologous vertebrate proteins and to the 6X bovine genome assembly (BGA); 703 probes from predicted RefSeq genes (Taylor *et al*. 2007); 5,943 probes from reproductive tissue ESTs with a BGA but no protein alignment; and 504 positive and negative controls . The microarray slides were printed by the University of Missouri Microarray Facility.

Animals were ranked based on their test NFI and the 22 with the lowest NFI and 22 with highest NFI were selected for the microarray hybridisation. Animals were matched by difference in NFI in descending order and the RNA from each pair (the highest with the lowest) was co-hybridized with a dye swap. Forty-four microarray slides were produced.

Microarray Data Analysis. Quality control measures, pre-processing, and analyses were performed using the statistical computing language R (R Development Core Team 2008) and Bioconductor (Gentleman *et al*. 2004). All microarray images and quality control measurements were within recommended limits. Prior to testing for differential expression, the data were filtered to remove control and empty spots (n=1,200), spots flagged as bad (n=752), and spots with less than two good reads in either contrast (n=351), thus 22,897 features to be tested. Differential transcription was tested using a moderated t-test in limma (Smyth, 2004) and features with an adjusted *P*-value of <0.01 were considered to be differentially expressed.

RESULTS AND DISCUSSION

One hundred eighty-one probes were identified as differentially expressed between the low and high NFI animals. Among them, 161 unique genes were identified by blast search of the bovine genome assembly (Btau4.0), reference sequence databank (NCBI), GeneBank and the EST database (NCBI). Four probes were not matched to known genes from GeneBank. Some genes had more than one probe that showed differential expression. Thirteen differentially expressed genes revealed by microarray data analysis were chosen to validate the microarray results by quantitative real-time reverse PCR. Ten of these genes had highly significant differential expression, and confirmed the microarray data (Figure 1).

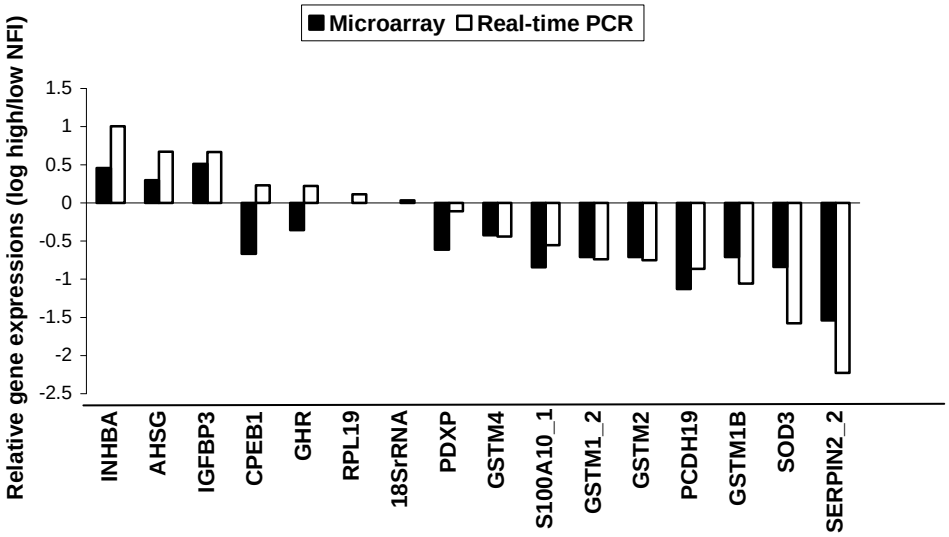


Figure 1. **Relative gene expression between low and high NFI animals determined by microarray and real-time PCR. Values are log-fold changes in expression of high over low NFI animals. RPL19 and 18SrRNA were the reference genes selected for real-time PCR**.

Among 161 unique differentially expressed genes, 85 genes were up-regulated and 76 were down-regulated in the low NFI (high efficiency) animals. The differentially expressed genes were annotated using the Database for Annotation, Visualization and Integrated Discovery (DAVID) software (SUPPLIER DETAILS) . Gene ontology analysis revealed that 86% of the up-regulated genes were involved in known biological processes, 92% have a known molecular function and 84% were related to cellular components. For down-regulated genes, 83% have known molecular function, 80% are involved in known biological processes and 78% are cellular components.

Up-regulated and down-regulated genes of the high efficiency animals were grouped based on functional clustering. The enriched GO-term functional cluster in up-regulated genes in high efficiency animals were biological processes involved in organism development and extracellular region. The KEGG pathway related to this biological process is the ECM-receptor interaction with 5 genes up-regulated in the high efficiency animals. For the down-regulated genes, the most enriched functional cluster is glutathione transferase activity in the molecular function that involves two KEGG pathways, metabolism of xenobiotics by cytochrome P450 and glutathione metabolism.

Differentially expressed genes were further analysed with Pathway Analysis software (Ingenuity systems, Mountain View, CA; http://www.ingenuity.com). Seven highly significant gene networks from the differentially expressed genes involving functions of cellular growth and proliferation, protein synthesis, lipid metabolism, carbohydrate metabolism, cancer, drug metabolism and small molecular biochemistry were identified.

CONCLUSION

Feed efficiency is a complex trait and the metabolic factors that contribute to variation are largely unknown. The global gene expression profiling of liver samples revealed 161 differentially expressed genes between animals with low and high NFI. Using gene ontology and pathway analysis, we identified that most of those genes have a known molecular function and revealed some important biological pathways that related to differences in NFI. This is the first report of differentially expressed genes between animals with high or low NFI using global gene expression. Further study of differentially expressed genes will add to our knowledge of the biological processes important for differences in efficiency of untilisation of nutrients. The differentially expressed genes provide evidence that will assist in the search for commercial genetic markers for feed efficiency in beef cattle.

ACKNOWLEDGEMENTS

We are grateful to Bill John, Reg Woodgate, Stu McClelland of BIC, NSW DPI and Reid Geddes of "Tullimba" Research Feedlot, and Dr. Steve Atkinson for helping the collection of the biospys.

REFERENCES

Archer, J.A., Richardson, E.C., Herd, R.M. and Arthur, P.F. (1999) *Aust. J. Agric. Res.* **50:**147.

Barendse, W., Reverter, A., Bunch, R.J., Harrison, B.E., Barris, W. and Thomas, M.B. (2007) *Genetics*. **176:**1893.

Dennis, G., Sherman, B., Hosack, D., Yang, J., Gao, W., Lane, H.C. and Lempicki, R. (2003) *Genome Biol.* **4:**P3.

Davies, D.C. and Jebbett, I.H. (1981) *In Pract.* 3:14.

Gentleman, R.C., Carey, V.J., Bates, D.M., Bolstad, B., Dettling, M., Dudoit, S., Ellis, B., Gautier, L., Ge, Y., Gentry, J., Hornik, K., Hothorn, T., Huber, W., Iacus, S., Irizarry, R., Leisch, F., Li, C., Maechler, M., Rossini, A.J., Sawitzki, G., Smith, C., Smyth, G., Tierney, L., Yang, J.Y., Zhang, J., 2004. *Genome Biol.* 5, R80.

Huang, D.W., Sherman, B.T. and Lempicki, R.A. (2009) *Nat. Protocols*. **4:**44.

Nkrumah, J.D., Sherman, E.L., Li, C., Marques, E., Crews, D.H., Jr., Bartusiak, R., Murdoch, B., Wang, Z., Basarab, J.A. and Moore, S.S. (2007) *J. Anim. Sci.* **85:**3170.

R Development Core Team, (2008) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org.

Sherman, E.L., Nkrumah, J.D., Murdoch, B.M. and Moore, S.S. (2008) *Animal Genetics*. **39:**225.

Smyth, G.K., 2004. *Stat Appl Genet Mol Biol.* 3:3.

Taylor, J.F., Elsik, C.G., Antoniou, E., Fahrenkrug, S.C., Reecy, J.M. and Wolfinger, R.D. (2007) *Plant & Animal Genomes XV Conference, January 13-17, 2007, CA*. W105.

Proc. Assoc. Advmt. Anim. Breed. Genet. 18:175-178

EXPRESSION OF MITOCHONDRIAL RESPIRATORY COMPLEX GENES IN LIVER TISSUE OF CATTLE WITH DIFFERENT FEED EFFICIENCY PHENOTYPES

K.J. Kochan, R.N. Vaughn, T.S. Amen, C.A. Abbey, J.O. Sanders, D.K. Lunt, A.D. Herring, J.E. Sawyer, C.A. Gill and P.K. Riggs

Department of Animal Science, Texas A&M University, College Station, Texas, USA 77843-2471

SUMMARY

Feed efficiency is an economically important trait that is likely influenced by complex molecular mechanisms. We utilized the NRC Beef Cattle Model to predict feed intake based on observed gain (and gain based on observed intake), where the NRC model also accounts for breed type, sex and season. The difference between this NRC-predicted feed intake and observed intake is termed model predicted residual consumption (MPRC). Associations between feed efficiency and mitochondrial respiration have been previously reported in the literature. From a study of 177 animals, RNA was extracted from liver samples from 18 animals at each extreme of the MPRC tails (36 samples). Following microarray analysis, quantitative realtime RT-PCR (qPCR) was used to examine expression of several respiratory complex genes including mitochondrial genes *COX1*, *COX2*, *COX3* (Complex IV) and *CYTB* (Complex III), and nuclear genes *COX4*, *COX6A1*, *COX7A2*, *COX7B*, *COX7C* (Complex IV), *SHDB* (Complex II) and *NQO2* (Complex I). Although expression for some genes was influenced by sire and family, no relationship between expression of any of these genes was found to be associated with feed efficiency phenotype.

INTRODUCTION

Cattle producers may derive economic benefit from selection of feed efficient animals, but molecular mechanisms that affect feed efficiency phenotype have not been clearly elucidated. Previously, Mukherjee *et al*. (1970) reported an association between feed efficiency and mitochondrial respiration in chickens. A number of studies confirmed the link (reviewed by Bottje *et al*. 2004), but the results of those experiments were influenced by differences in breed and/or diet until Bottje *et al*. (2002) observed the same association in male chickens of a single broiler strain fed the same diet. Several parameters of respiration were measured in mitochondria from leg and breast muscle of high and low feed efficiency (FE; the ratio of gain-to-feed) birds, and a correlation between FE and efficient coupling of electron transport was hypothesized. Subsequent studies (Bottje *et al*. 2004; Iqbal *et al*. 2004, 2005; Ojano-Dirain *et al*. 2004, 2005) supported that hypothesis, suggesting that activity of all 5 respiratory complexes was decreased in low FE chickens.

More recently, Kolath *et al*. (2006) isolated mitochondria from the longissimus muscle of Angus steers with low or high residual feed intake (RFI). Mitochondria from steers with low RFI (that is, the more feed-efficient animals) exhibited higher respiration rates than those from steers with high RFI, but mitochondrial function was not different between the two groups.

In the current study, we compared expression of respiratory complex genes by measuring mRNA quantity in liver from steers with feed efficient or inefficient phenotypes. Liver tissue was chosen for this study because of the physiological role of the liver in metabolic processes.

MATERIALS AND METHODS

Animals. Data were collected on 177 Nellore-Angus F_2 steers produced by embryo transfer from 10 Nellore-Angus F_1 donor females and 4 Nellore-Angus F_1 sires (Amen 2007). Individual feed intake was evaluated as previously described by Amen (2007). Briefly, steers were fed and intake

was measured by use of a Calan gate system, beginning at an average age of 11 to 13 months until slaughter at age 17-18 months. Liver samples were collected at time of harvest and snap-frozen in liquid nitrogen, then stored at -70°C until processed. As described by Amen (2007) and based on the NRC (2000) model, daily feed intake was predicted based on observed weight gain for each animal and standardized input for animal type, age, sex, condition, and breed. The model predicted dry matter intake (MDMI) was subtracted from observed DMI and the difference defined as MPRC, such that those animals that consumed less than predicted (and thus, were more efficient) had negative MPRC. This method was used instead of traditional RFI so that data from multiple contemporary groups could be used simultaneously. Liver RNA samples from the 18 steers at each extreme of MPRC were used in the microarray analysis (n=36). MPRC values for these steers ranged from 1.2 to 2.8 standard deviations from the mean of the group as a whole.

Table 1. List of real-time PCR assays

Gene	Complex	Genome	Accession #	Primers
COX1	IV	mitochondrial	DQ124400	F: 5'-gggaatagtttgggctataatgtc
				R: 5'-gatgtgaagtaggctcgtgtgt
COX2	IV	mitochondrial	DQ124400	F: 5'-tcgtcccgtccaggctta
				R: 5'-aactgtggtttgacccgca
COX3	IV	mitochondrial	DQ124400	F: 5'-ccaccacttcggctttgaag
				R: 5'-ggaaaagtcagactacgtctacgaaa
COX4	IV	nuclear	NM_001001439	F: 5'-atcccgcacacctttga
				R: 5'-ttccactcgttcttgtcgtag
COX6a1	IV	nuclear	NM_001077831	F: 5'-ccctattccataaccctcatgtg
				R: 5'-tccaggttctctttattcgtcttca
COX7a2	IV	nuclear	NM_175807	F: 5'-cggttggtgggtagtaactg
				R: 5'-atggtcctcttagcaatctgac
COX7b	IV	nuclear	NM_175795	F: 5'-tttatgttcaacctcaggatgtttc
				R: 5'-atctgccttgccactgctt
COX7c	IV	nuclear	NM_175831	F: 5'-tgcagccgccatttcttc
				R: 5'-tagcgctgttggacgctcta
CYTB	III	mitochondrial	EF693798	F: 5'-catccgacacaacaacagcatt
				R: 5'-gctccgtttgcgtgtatgtatc
SDHB	II	nuclear	NM_001040483	F: 5'-tactggtggaacggagacaag
				R: 5'-gtgtggcagcggtagaga
NQO2	I	nuclear	NM_001034323	F: 5'-gtgacatcattgaggagcagaaga
				R: 5'-cgggcacgctgaacca
RPS20	control	nuclear	BC103289	F: 5'-accagccgcaacgtgaa
				R: 5'-ccttcgcgcctctgatca

RNA. Liver tissue (about 200 mg/sample) was pulverized under liquid nitrogen, transferred to 2 ml TRI Reagent® (Molecular Resource Center, Cincinnati, OH, USA) and homogenized through an 18 ga needle. RNA was extracted from homogenized tissue with TRI Reagent® and 1-bromo-3-chloropropane (BCP, Molecular Resource Center). The manufacturer's recommended protocol was modified to include additional extractions with 2:1 TRI:BCP and BCP alone. RNA was precipitated in 1 ml isopropanol, washed consecutively with 70%, 95% and 100% ethanol, and resuspended in 100 ul nuclease-free water (Invitrogen, Carlsbad, CA, USA). RNA quality was assessed by capillary electrophoresis through RNA 6000 NanoChips on a 2100 Bioanalyzer (Agilent Technologies, Palo Alto, CA, USA) according to the manufacturer's instructions. RNA was purified through RNeasy mini columns (Qiagen, Valencia, CA) using the manufacturer's RNA cleanup protocol, DNase-treated with the DNA-*free*™ kit (Ambion, Austin, TX, USA)

according to the kit instructions, and quantified on a NanoDrop™ 1000 spectrophotometer (Thermo Scientific, Waltham, MA, USA).

Real-time PCR. Total RNA (800 ng) was reverse transcribed (RT) in a 40 µl reaction with the High Capacity cDNA Reverse Transcription Kit (Applied Biosystems, Foster City, CA, USA). One reaction containing template, but no enzyme, was included as a control. The cDNA was diluted 1:4 in 25 ng/ul yeast tRNA (Sigma-Aldrich, St. Louis, MO, USA). Real-time PCR was performed in a 20 µl reaction containing 2 µl cDNA, 1X SYBR® GreenER™ PCR master mix (Invitrogen) and 300 nM primers. Primer pairs (Table 1) were designed with Oligo 6 software (Molecular Biology Insights, Inc., Cascade, CO). Amplification was carried out in 96-well plates in a 7900 HT real-time thermal cycler (Applied Biosystems, Foster City, CA, USA) with the thermal profile: 95°C for 10 minutes followed by 40 cycles of 95°C for 15 seconds and 60°C for 1 minute. Amplification data were analyzed with SDS software v.2.2.2 (Applied Biosystems). Amplification efficiency was validated for all primer pairs.

Data were normalized to *RPS20* expression (a gene whose expression was consistent across samples in this study), and relative expression was calculated (Livak and Schmittgen, 2001). The mean of all 36 samples was used as the calibrator value. Expression data were compared by analysis of covariance (SAS® 9.2, PROC GLM; SAS Institute, Inc., Cary, NC, USA). Sire, family, contemporary group, and efficiency phenotype (efficient or inefficient) were modeled as fixed effects with MPRC as a covariate.

RESULTS AND DISCUSSION

This study was conducted to identify whether expression of mitochondrial respiratory complex genes in liver was correlated with efficiency phenotype. Previously published results indicated some association of mitochondrial function with feed efficiency (reviewed in Bottje *et al.* 2006; Bottje and Carstens, 2009). Preliminary microarray experiments (Bovine Oligo Consortium arrays; http://bovineoligo.org) also indicated a possible association between *COX3* and *CYTB*

Table 2. Respiratory complex gene expression, in liver from steers with high and low MPRC values. Data are presented as arbitrary units (AU), relative to mean expression of all samples. Expression is not different between groups.

Gene	Relative gene expression	
	Low MPRC	High MPRC
COX1	0.985 ± 0.225	1.052 ± 0.379
COX2	1.122 ± 0.249	0.919 ± 0.315
COX3	1.137 ± 0.359	0.959 ± 0.381
COX4	0.996 ± 0.220	1.038 ± 0.220
COX6A1	1.008 ± 0.190	1.000 ± 0.155
COX7A2	1.019 ± 0.185	1.008 ± 0.235
COX7B	1.026 ± 0.285	1.001 ± 0.186
COX7C	1.027 ± 0.233	0.983 ± 0.141
CYTB	1.097 ± 0.302	0.969 ± 0.387
SHDB	0.985 ± 0.225	0.985 ± 0.225
NQO2	0.985 ± 0.225	0.985 ± 0.225

(Riggs, 2008). We conducted additional experiments with commercial bovine microarrays and used qPCR to analyze genes representing subunits from all four complexes of the mitochondrial electron transport chain (ETC). We could not validate *COX3* and *CYTB* as differentially expressed

between feed efficiency phenotypes. The initial arrays resulted in weak hybridization signals and may have been affected by the presence of mitochondrial DNA (not shown). Additional DNase treatment was performed prior to qPCR analysis to eliminate potential mitochondrial DNA contamination. Statistical analysis also indicated that expression of *COX2* and *NQO2* appeared to be influenced by sire, and expression of *COX7B* and *NQO2* was affected by family ($p < 0.05$).

CONCLUSIONS

While variation in feed efficiency has been attributed to differences in mitochondrial respiratory function or electron leak due to electron transport defects, differential expression of ETC genes in liver tissue was not associated with differences in feed efficiency in this study. Genetic variation in expression of some of the subunit genes (*COX2*, *COX7B*, and *NQO2*) does appear to exist in the study population, but is not related to feed efficiency. The relationship between mitochondrial functional activity and the quantity of mRNA or protein for individual ETC complex subunits is not clear (e.g. Bottje *et al*. 2004; Ojano-Dirain *et al*. 2005; Garrabou *et al*. 2007). If mitochondrial function is critical for feed efficiency phenotype, other genes that regulate activity may be involved, since differences in expression of mitochondrial genes *COX1*, *COX2*, *COX3* (Complex IV) and *CYTB* (Complex III), and nuclear genes *COX4*, *COX6A1*, *COX7A2*, *COX7B*, *COX7C* (Complex IV), *SHDB* (Complex II) and *NQO2* (Complex I) were not observed between steers with high (inefficient) or low (efficient) MPRC. Expression of these genes was examined only in liver tissue from animals at the tails of the MPRC distribution and not the population as a whole.

ACKNOWLEDGMENTS

We thank Jordan E. Shields and Jessica Alonzo for technical assistance. This project was funded in part by America's beef and veal producers through the $1 per head check off, and by Texas AgriLife Research.

REFERENCES

Amen, T.S. (2007) PhD Thesis, Texas A&M University.

Bottje, W. and Carstens, G.E. (2009). J Anim. Sci. **87**:E48.

Bottje, W., Iqbal, M., Tang, Z.X., Cawthon, D., Okimoto, R., Wing, T. and Cooper, M. (2002) *Poult. Sci.* **81**:546.

Bottje, W., Iqbal, M., Pumford, N.R., Ojano-Dirain, C. and Lassiter, K. (2004) *J. Appl. Poult. Res.* **13**:94.

Bottje, W., Pumford, N.R., Ojano-Dirain, C., Iqbal, M., and Lassiter, K. (2006) *Poul. Sci.* **85**:8.

Garrabou, G., Soriano, A., López, S., Guallar, J.P., Giralt, M., Villarroya, F., Martínez, J.A., Casademont, J., Cardellach, F., Mensa, J. and Miró, O. (2007) *Antimicrob. Agents Chemother.* **51**:962.

Iqbal, M., Pumford, N.R., Tang, Z.X., Lassiter, K., Ojano-Dirain, C., Wing, T., Cooper, M. and Bottje, W. (2004) *Poult. Sci.* **83**:474.

Iqbal, M., Pumford, N.R., Tang, Z.X., Lassiter, K., Ojano-Dirain, C., Wing, T., Cooper, M. and Bottje, W. (2005) *Poult. Sci.* **84**:933.

Kolath, W.H., Kerley, M.S., Golden, J.W. and Keisler, D.H. (2006) *J. Anim. Sci.* **84**:861.

Livak, K.J. and Schmittgen T.D. (2001) *Methods* **25**:402.

Mukhurjee, T.K., Stevens, R.W.C. and Hoogendoorn, M.P. (1970) *Poult. Sci.* **49**:1130.

NRC. (2000) "Nutrient Requirements of Beef Cattle." 7th ed., revised. National Academy Press, Washington, D.C.

Ojano-Dirain, C.P., Iqbal, M., Cawthon, D., Swonger, S., Wing, T., Cooper, M. and Bottje, W. (2004) *Poult. Sci.* **83**:1394.

Ojano-Dirain, C.P., Pumford, N.R., Iqbal, M., Wing, T., Cooper, M. and Bottje, W.G. (2005) *Poult. Sci.* **84**:1926.

Riggs, P.K. (2008) *Beef Cattle Research in Texas. In press.*

Proc. Assoc. Advmt. Anim. Breed. Genet. 18:179-182

IDENTIFICATION OF DIFFERENTIALLY EXPRESSED TRANSCRIPTS IN M. *LONGISSIMUS DORSI* WITH DIVERGENT MARBLING PHENOTYPES IN HANWOO (KOREAN NATIVE CATTLE)

S. H. Lee[1,2,3], C. Gondro[3,4], J. H. J. van der Werf[3], N. K. Kim[1], D. J. Lim[1], Y. H. Shin[1], J. P. Gibson[2,3] and J. M. Thompson[2,3]

[1]Animal Genomics & Bioinformatics Division, National Institute of Animal Science, RDA, Suwon 441-706, Korea. [2]The Cooperative Research Centre for Beef Genetic Technologies, University of New England, Armidale, NSW 2351, Australia. [3]School of Environmental and Rural Science, University of New England, Armidale, NSW 2351, Australia. 4 The Institute for Genetics and Bioinformatics, University of New England, Armidale, NSW 2351, Australia

SUMMARY

Marbling is a major trait in determining profit in the Korean beef industry. However, the underlying biology of muscles with divergent marbling phenotypes is still poorly understood in cattle. In this study, we attempted to detect differentially expressed genes in *M. longissimus dorsi* in Hanwoo steers with high and low estimated breeding values for marbling score using an Affymetrix bovine gene expression array. Three data-processing methods (MAS5.0, GCRMA and RMA) were implemented to test for differential expression (DE). Analysis identified 21 transcripts exhibiting significant DE in at least two data-processing methods ($P<0.001$). Squalene epoxidase and Cytochrome P450 gene, thought to play a role in biotransformation of steroids, were expressed more in highly marbled muscle whereas gene for Arginyl-tRNA synthetase, Ribosomal protein S6 kinase, Thimetoligopeptidase 1, Proteosome activator subunit 4, ATP binding protein and CDC-like kinase gene, which are all involved in protein synthesis and cell division, were down-regulated in highly marbled muscle. These results suggest that down-regulation of genes involved in cell division, growth regulation and protein synthesis may lead to decreased muscle mass and increased adiposity within muscle.

INTRODUCTION

Intramuscular fat deposition within the musculature starts to become visible at the age of 12 months and the rate increases from 15 months to 24 months (Nishimura *et al.* 1999). The initial formation of visible intramuscular fat seems to be a result of the development of adipocytes (Pethick *et al.* 2006). Hocquette *et al.* (1998) reported that metabolic differences to balance triacylgriceride (TAG) storage within muscle such as fatty acid trafficking and oxidation of fatty acid in mitochondria are known as an important biological contributor in determining marbling levels in the later stages of finishing cattle.

Kokta *et al.* (2004) reviewed the interaction between myogenic cells and adipocytes to determine the rate and extent of myogenesis and adipogenesis during animal growth. Fat and muscle development are regulated by a number of complex biological pathways such as adenoreceptor signaling (Fruhbeck *et al*. 2001), cytokine signaling pathway (Shin *et al*. 2003) and a wide range of hormonal and transcriptional factors. Therefore, marbling differences might be expressed by metabolic differences resulting from a complex mechanism of communication among cells (Sorisky *et al*. 1999).

This study identified differentially expressed genes in *M. longissimus dorsi* of animals with divergent marbling phenotypes. We selected high and low marbling animals and then looked at between-group differential expression of genes using the bovine genome array (Affymetrix Inc, USA).

MATERIALS AND METHODS

Animals. From a group of 90 steers, the 5 highest and the 5 lowest were selected based on the marbling score. Carcass measurements and intramuscular fat percentage (IMF % as defined by AOAC, 1990) of the muscle sample were measured. Summary statistics for the animals and muscle samples used in this study are shown in Table 1.

Table 1. Summary statistics of tissue sample for gene expression analysis

Groups	Animal	*EBV	Age (Month)	Marbling score (1-7)	IMF (%)
Low	509	0.37	26	2++	7.11
	537	0.2	27	2++	6.02
	554	0.4	27	3	4.88
	670	0.31	28	3	7.36
	691	0.2	28	3	12.04
High	527	1.02	26	7++	24.35
	547	1.015	27	7++	32.49
	586	0.7	31	7++	16.56
	589	0.69	28	7++	26.24
	632	0.415	28	7++	18.81

Target preparation and high-density array hybridization. Double stranded cDNA was synthesized from 3 μg mRNA using a Genechip Expression 3'-Amplification One Cycle Synthesis kit (Affymetrix Inc. USA). After the cDNA was purified, Biotin-labeled cRNA was synthesized in vitro using the Gene chip Expression 3'-Amplification reagents in the IVT labeling kit (Affymetrix Inc.). A hybridization cocktail (200 μl) containing 15 μg fragmented cRNA was injected into the Genechip Bovine Gene expression Array (Affymetrix Inc). The array was placed in a 45 hybridization oven at 60 rpm for 16 hours. The array was scanned with a GeneChipScanner 3000 (Affymetrix Inc.).

Data pre-processing. Data quality control and background correction were carried out using the statistical computing language R (http://www.R-project.org). All slides were deemed to be within normal quality standards. Expression intensities on a log2 scale were obtained from the probe level data using the R *affy* package (Gautier *et al.* 2004) for MAS5.0 (Affymetrix 2002), RMA (Irizarry *et al.* 2003) and GCRMA (Wu *et al.* 2003) methods. In MAS5.0, each probe was adjusted using a weighted average. All arrays were scaled to the same mean value for normalization (200) and were summarized by an adjusted log2 scale average using 1-step Tukey biweight. For RMA, the background was corrected by convolution. The data were quantile normalized and summarized by median polish. GCRMA background correction used an affinity measure model based on probe sequences and mismatch probe intensities. The data were filtered to remove control probe (n=131) and probes detected as marginal (M) and absent (A) in all arrays using MAS5 presence calls.

Statistical analysis. Genes differentially expressed between the high- and low marbling groups were detected using a moderated t-test in *limma* (Smyth 2004).

RESULTS AND DISCUSSION

Differentially expressed genes. A total of 136 differentially expressed genes (DEGs) were detected in 3 data-processing methods; MAS5.0 (65 transcripts), RMA (37 transcripts) and GCRMA (34 transcripts) (Figure 1). Of 136 DEGs, 21 were shown to be significant in at least 2 of the summarization methods (Figure 1). Of the 21 differentially expressed genes listed in Table 2, 8

DEGs were identified as up-regulated in the high marbling group and the remaining 13 DEGs were down-regulated in the low marling group. Based on the gene identities and associated function, 2 up-regulated genes are involved in steroid biosynthesis (squalene epoxidase and cytochrome P450). Six down-regulated genes (Arginyl-tRNA synthetase, Ribosomal protein S6 kinase, Thimetoligopeptidase 1, Proteosome activator subunit 4, ATP binding protein and CDC-like protein) belong to functional classes involved in DNA replication, protein synthesis and cell division. However, 7 DEGs are yet unidentified hypothetical proteins (single EST clones). Of these 7 DEGs, three hypothetical proteins (LOC788205, LOC509649 and LOC777601) and 2 transcribed loci (Bt.19107.2.A1_at and Bt.19107.1.S1) were identified as differentially expressed. These probes can be considered candidate genes for biochemical indicators of IMF accretion.

Gene Ontology (GO) analysis. GO terms were annotated onto the GO database (http://www.geneontology.org). The GO biological process is assigned to 10 categories at level 3 (Figure 1.B). Gene Ontology (GO) analysis shows that the 21 DEGs were mainly involved in primary metabolic process and cellular metabolic process in the biological process category, and more specifically in oxidation reduction and regulation of the development process.

Table 2. Differentially expressed genes in *M. longissimus dorsi* of high and low marbled Hanwoo steers

Probe ID	Gene Names	Fold Change	Significance in MAS, RMA and GCRMA
Bt.5323.1.S1_at	SH3 domain YSC-like 1	0.818	* * *
Bt.15675.1.S1_at	ADAM metallopeptidase with thrombospondin type 1	0.953	* * *
Bt.21021.1.S1_at	TBC1 domain family, member 7	0.712	* * *
Bt.2933.1.S1_at	Hypothetical protein LOC788205	0.668	* * *
Bt.9767.1.S1_a_at	Squalene epoxidase	0.867	* * *
Bt.621.1.S1_at	Cytochrome P450, family 51, subfamily A	0.525	* * ns
Bt.23903.1.A1_at	Unknown	-0.53	* * ns
Bt.22362.1.S1_at	Similar to SH3-domain kinase binding protein 1	-0.96	* * ns
Bt.16752.1.A1_at	ATP binding protein	-0.693	ns * *
Bt.1020.1.S1_at	Similar to CDC-like kinase 1	-0.408	ns * *
Bt.19107.2.A1_at	Transcribed locus	-0.548	ns * *
Bt.28011.1.S1_at	Unknown	-1.066	ns * *
Bt.22718.1.A1_at	Proteasome (prosome, macropain) activator subunit 4	-0.326	ns * *
Bt.19107.1.S1_at	Transcribed locus	-0.642	ns * *
Bt.25102.1.S1_a_at	Hypothetical LOC509649	-0.496	ns * *
Bt.22038.1.S1_a_at	Similar to Arginyl-tRNAsynthetase	-0.215	ns * *
Bt.21268.1.S2_at	Ribosomal protein S6 kinase, 70kDa, polypeptide 1	0.459	ns * *
Bt.13342.1.S1_at	Similar to Src-associated protein SAW	-0.383	ns * *
Bt.344.1.S1_at	Major histocompatibility complex, class II, DM alpha	-0.595	ns * *
Bt.21827.2.S1_at	Thimetoligopeptidase 1	-0.818	* * ns
Bt.21794.1.S1_at	Hypothetical protein LOC777601	1.1243	* * ns

CONCLUSIONS

The analysis detected 2 genes which were upregulated in highly marbled cattle (Squalene epoxidase and Cytochrome P450) and thought to play a role in the fatty acid oxidation/reduction pathway. In addition there were 6 downregulated genes (Arginyl-tRNA synthetase, Ribosomal protein S6 kinase, Thimetoligopeptidase 1, Proteosome activator subunit 4, ATP binding protein

and CDC-like protein) which are involved in protein synthesis. This indicates that increased marbling in cattle was associated with down-regulation of genes involved in the cell division cycle and protein synthesis and upregulation of genes associated with adiposity within the muscle.

(A) **(B)**

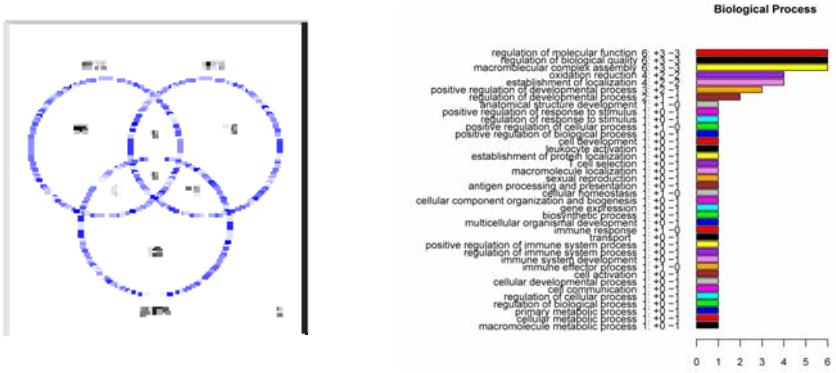


Figure 1. (A) Distribution of the 136 differentially expressed genes on intramuscular fat (IMF) in three data-processing methods (MAS5.0, RMA and GCRMA). (B) GO annotation (biological process term) for 21 differentially expressed genes.

ACKNOWLEDGMENTS

This study was supported by the International Collaborative research fund (Grant No: 200712A01032083) between the Rural Development Administration (RDA) in Korea and Co-operative Research Centre for Beef Genetic Technologies in Australia. Mr Seung Hwan Lee held an International Postgraduate Research Scholarship (IPRS).

REFERENCES

Affymetrix (2002) Statistical algorithms description document

AOAC (1990) Association of Official Analytical Chemists, Washington, DC, USA

Fruhbeck, G., Gomez-Ambrosi, J and Burrell, M.A (2001) Am J Physiol. **280**:E827.

Gautier, L., Cope, L., Bolstad, B.M., Irizarry, R.A (2004) Bioinformatics **20**:307.

Hocquette, J.F., Ortigues-Marty, I, Pethick, D and Fernandez, X (1998) Livest Prod Sci. **56**:115.

Kokta, T.A, Dodson, M.V., Gertler, A and Hill, R.A (2004) Domes Anim Endocrinol **27**:303.

Irizarry, R.A., Bolstad, B.M., Collin, F., Cope, L.M., Speed, T.P (2003) Nucleic Acids Research **31:** e15.

Nishimura, T., A. Hattori and K. Takahashi (1999) J Anim Sci, **77**:93.

Pethick, D.W., Harper, G.S., Hocquette, J.F and Wang, Y.H (2006) Aust beef-the leader conference, 103.

R Development Core Team (2008). R: A language and environment for statistical computing.

Shin, S.M et al (2003) FEBS Lett. **543**:25.

Smyth, G.K (2004) Statistical Applications in Genetics and Molecular Biology **3:**3.

Sorisky, A (1999) Crit Rev Clin Lab Sci. **36**:1.

Zhijin, Wu et al (2003) Journal of the American Statistical Association **99:** 909.

Proc. Assoc. Advmt. Anim. Breed. Genet. 18:183-186

QTL MAPPING FOR FEED CONVERSION EFFICIENCY ON PORCINE CHROMOSOME 10 IN AN AUSTRALIAN COMMERCIAL POPULATION

Y. Chen[1], Y. Zhang[2], I. MacLeod[3], R. Kerr[2], K.L. Bunter[2], B. Hayes[3], B. Tier[2], H.-U. Graser[2], B.G. Luxford[4], M. Goddard[3] and C. Moran[1]

[1]Centre for Advanced Technologies in Animal Genetics and Reproduction, University of Sydney, NSW 2006
[2] Animal Genetics and Breeding Unit*, University of New England, Armidale, NSW 2351
[3]Department of Primary Industries, PIRVic, Attwood, VIC 3049
[4]QAF Meat Industries, Corowa, NSW 2646

SUMMARY

A genetic linkage map of 23 markers on porcine chromosome 10 was constructed with a resource pedigree based on an Australian commercial pig population. Six new markers (UMNP885, UMNP1049, UMNP875, UMNP925, UMNP876 and UMNP519) were linkage mapped to porcine chromosome 10 for the first time. Phenotypes were available for juvenile IGF-I along with feed intake, average daily gain and feed conversion ratio (FCR) recorded over a 6 week performance test period; associations between QTL and residuals for these traits were investigated. A significant QTL for FCR was found between marker UMNP875 and UMNP925 by single family and across family analysis with maximum likelihood using composite interval mapping, confirmed by linkage disequilibrum and linkage analysis (LDLA). In single family analyses, a significant QTL for average daily gain was found in sire family 896TS and an IGF-I QTL was found in sire family 52103. The study provided strong justification for further fine mapping and positional cloning of causative genes for FCR on chromosome 10 for marker assisted selection in pig breeding.

INTRODUCTION

In the Australian pig industry, feed costs account for ~60% of the costs of production (Henman 2003). All other things being constant, the less feed that it takes to bring an animal to market weight, the more profitable the enterprise will be. Many pig breeders would like to include feed efficiency in their breeding programmes. However, feed efficiency measurement is difficult and expensive, since individual feed intake must be recorded. Gene markers for feed efficiency would be very useful as alternative methods for improving feed efficiency without the need to measure feed intake individually. Initially QTL (quantitative trait locus) mapping in pigs used crosses between divergent breeds to find chromosome regions that affect particular traits. Since the first publication of a QTL detected with a cross of European Wild Boar and Large White (Andersson *et al*. 1994), 1831 QTLs have been reported in the PigQTL database from 113 publications representing 317 different pig traits (http://www.animalgenome.org/cgi-bin/QTLdb/SS/summary) on 28 March 2009. Eight QTLs for feed conversion ratio have been reported on chromosomes 3, 4, 5, 6, 8, 13 18 X (Geldermann *et al*. 2003; Lee *et al*. 2003; Stratil *et al*. 2006). This paper reports the first QTL for feed conversion ratio mapped in an Australian commercial pig population from an extensive QTL mapping project funded by Australian Pig Research and Corporation (unpublished).

* AGBU is a joint venture of the NSW Department of Primary Industries and the University of New England

MATERIALS AND METHODS

Data collection. An Australian resource population consisting of 430 progeny of eight sires was bred for QTL mapping at QAF Meat Industries, Corowa, NSW, Australia, between 1999 and 2001. The animals were from two closed lines of Large White and Landrace origin. Daily feed intake was recorded during the 6 week performance test period (from 18-24 weeks). Animals were single penned and fed *ad-libitum*. Weight of the animal was recorded at the beginning of the testing period and shortly before slaughter at 24 weeks. Average daily gain (ADG2), average daily feed intake (ADFI) and feed conversion ratio (FCR) were calculated from the performance test data. Blood was collected 3-5 days after weaning and juvenile pigs for concentrations of insulin-like growth factor-I (IGF-I) were carried out by Primegro (see Bunter *et al.* 2005). Residuals, or phenotypes corrected for non-genetic systematic effects, were used for the association study.

Markers and map construction. A total of 26 markers were genotyped on pig chromosome 10 for the resource pedigree (SW830, SWR136, SW249, SW767, SW1894, SW2491, UMNP885, UMNP1049, UMNP875, UMNP925, SW2195, SWC19, SW173, KS115, S0070, UMNP876, ACO1, SWR1849, UMNP519, SW1041, SW1405, SW1991, SW1626, UMNP104, UMNP599 and UMNP238)**.** The linkage map was constructed using CRIMAP (Green *et al.* 1990) with options FIXED, FLIP and CHROMPIC.

QTL analysis with maximum likelihood using composite interval mapping. A segregating QTL in a sire family causes a phenotype contrast between progeny inheriting alternative QTL alleles Q and q when the sire is heterozygous for the QTL (Qq). At a given map position, genotypes of two flanking markers were used to calculate prior probabilities for progeny having inherited the Q or q allele. As the linkage phases between markers and QTL cannot be considered consistent across families, QTL analyses were performed separately for each sire family. Maximum likelihood estimates were obtained by using the expectation/conditional maximization (ECM) (Zeng 1994). The QTL was tested at every 2 cM along the chromosome with chromosome-wide critical value obtained by the empirical threshold determined by 1000 permutations at each point.

QTL analysis with linkage disequilibrum and linkage analysis (LDLA). Marker haplotypes of sires and dams were reconstructed from progeny genotypes. For the midpoint of each marker interval (putative QTL locations), a matrix of the probability of identity by descent (IBD matrix) was constructed among the base haplotypes using the LD method of Meuwissen and Goddard (2001). This method requires an assumption to be made about the effective size of the population The average value of the effective population size was calculated from chromosomal segment homozygosity (CSH) which is defined as the probability that two gametes drawn at random from the population carry homologous chromosome segments descended from the same common ancestor. The IBD probability was then used in a covariance structure were you fitted random QTL effects.

RESULTS AND DISCUSSION

Genetic Linkage map. Most markers showed significant linkages to other markers by pair-wise linkage analysis. Markers UMNP104, UMNP599 and UMNP238 were excluded from the final genetic linkage maps because there were insufficient informative genotypes to determine their map positions (Table 1). The genetic linkage map of chromosome 10 derived from this data is

significantly bigger than the USDA-MARC map (v2). However, the order of the markers in our study was consistent with the map of USDA-MARC map (v2). Six new markers (UMNP885, UMNP1049, UMNP875, UMNP925, UMNP876 and UMNP519) were linkage mapped for the first time and were consistent with their positions in the physical map of chromosome 10.

Table 1. Linkage map of genotyped markers on porcine chromosome 10

Marker	Position	USDA(2.0) §	Marker	Position	USDA(2.0) §
SW830	0	0.0	SW173	147.7	56.1
SWR136	27.2	7.6	KS115	154.1	58.4
SW249	29.3	17.3	S0070	165.2	62.3
SW767	48.5	20.4	UMNP876*	170.5	63.3
SW1894	48.6	23.2	AC01	175.8	64.3
SW2491	66.4	43.0	SWR1849	179.9	65.1
UMNP885	80.1	43.3*	UMNP519*	189.9	66.5
UMNP1049	97.9	43.6*	SW1041	196.1	67.5
UMNP875	109.1	43.8*	SW1405	196.2	67.5
UMNP925	121.4	44.0*	SW1991	223.5	79.4
SW2195	121.5	44.0	SW1626	269.3	108.0
SWC19	135.3	50.5			

§ map position on MARC-Map; *no linkage map position is available in the MARC-Map; the indicative position was derived from the high resolution IMpRH physical map.

Single family and across family analyses with maximum likelihood using composite interval mapping. The maximum likelihood analysis revealed chromosome-wide significance ($p<0.01$) QTL for FCR at position 112 cM in family 80496 (Table 2). A point-wise significant QTL ($p<0.01$) for FCR was also found in sire family 896TS at position 136. A significant QTL for ADG2 was found at 110 cM in sire family 896TS. Evidence from ADFI, FCR and ADG suggested that a QTL at 110 cM was segregating in this sire family. A QTL for IGF-I levels was found in sire family 52103 at position 122 with chromosome-wide significance. Multi-family analysis revealed QTL for FCR at chromosome-wide significance ($p<0.01$) at position 118 cM (between SW2195 and SWC19). An IGF-I QTL was found at this position with point-wise significance ($p<0.01$).

QTL analysis with linkage disequilibrum and linkage analysis (LDLA). A very significant FCR QTL was found at position 115 ($p<0.01$) between markers UMNP875 and UMNP925, corresponding to USDA-MARC map position 43.9 cM. It was consistent with the finding of single and across family analyses with the maximum likelihood interval mapping. There was a suggestive QTL ($p<0.05$) for ADG2 at position 73.5 between markers SW2491and UMNP875. No significant QTL was found for ADFI. This is the first reported QTL for FCR found in a commercial pig population and no FCR-QTL have previously been reported on porcine chromosome 10. The position and effect were consistent by single family and across family analysis with maximum likelihood using composite interval mapping linkage disequilibrum and linkage analysis (LDLA). It provides strong justification for further fine mapping and positional cloning of causative genes for marker assisted selection in pig breeding.

Table 2. QTL for FCR, ADFI, ADG2 and IGF-I from single trait, single family Maximum Likelihood analyses of chromosome 10

Trait	Family	n	cM	Effect	LRT[1]
ADFI (kg/day)	896TS	45	34	0.25	4.97*
	896TS	45	110	0.21	4.48*
	80475	33	98	0.54	8.59**
	R7292	56	224	0.20	4.92*
ADG2 (g/day)	80393	52	0	88.9	6.63*
	80496	61	66	113	5.48*
	896TS	45	110	127	8.1***
FCR (kg/kg)	80393	51	18	0.17	4.64*
	52103	56	66	0.15	5.15*
	896TS	43	136	0.18	7.42**
	80496	59	112	0.28	21.4****
IGF-I (ng/ml)	80475	33	264	1.61	4.48*
	52103	63	122	2.15	8.46***

n number of progeny in the family;[1] * 5% point-wise significant; ** 1% point-wise significant; *** 5% chromosome-wide significant; **** 1% chromosome wide significant; see text for trait abbreviations

ACKNOWLEDGEMENT

This work was financially supported by Australian Pork Limited (previously Pig Research and Development Corporation) under Grant US43. The cooperation and support from staff of QAF for the project is greatly appreciated.

REFERENCE

Andersson, L., C. S. Haley, H., Ellegren, S. A., Knott, M., Johansson, K., Andersson, L.. Anderssoneklund, I., Edforslilja, M., Fredholm, I., Hansson, J., Hakansson and Lundstrom, K. (1994). *Science* **263**:1771.

Bunter, K. L., S. Hermesch, B. G., Luxford, H. U., Graser H. and Crump, R.E. (2005). *Aust.J.Exp.Agri*.**45**:783.

Geldermann, H., E. Muller, G., Moser, G., Reiner, H., Bartenschlager, S., Cepica, A., Stratil, J., Kuryl, C., Moran, R., Davoli and Brunsch, C. (2003). *J.Anim.Breed.Genet.***120**:363.

Henman, D. (2003) In "Perspectives in pig science" 11-132. editors J. Wiseman, M. A. Varley and B. Kemp

Lee, S. S., Chen, Y., Moran, C., Stratil, A., Reiner, G., Bartenschlager, H., Moser G. and Geldermann, H. (2003). *J..Anim.Breed. Genet.***120**:38.

Stratil, A., Van Poucke, M., Bartenschlager, H., Knoll, A., Yerle, M., Peelman,L.J., M. Kopecny M. and Geldermann, H. (2006). *Anim. Genet.* **37**: 415.

Zeng, Z. B. (1994). *Genetics* **136**: 1457.

Proc. Assoc. Advmt. Anim. Breed. Genet. 18:187-190

A GENOME-WIDE ASSOCIATION ANALYSIS IDENTIFYING SNPS FOR PRRS TOLERANCE ON A COMMERCIAL PIG FARM

C.R.G. Lewis[1,3], M. Torremorell[2], L Galina-Pantoja[2], N Deeb[2], M.A. Mellencamp[2,4], A.L. Archibald[1] and S.C. Bishop[1]

[1]The Roslin Institute & R(D)SVS, University of Edinburgh, Roslin, Midlothian EH25 9PS, UK
[2]Genus/PIC, 100 Bluegrass Commons Blvd, suite 2200, Hendersonville, TN 37075, USA
[3]Current Address: Animal Genetics and Breeding Unit, University of New England, Armidale, NSW 2351, Australia
[4]Current Address: Ralco Nutrition, Inc., 1600 Hahn Road, Marshall, MN 56258, USA

SUMMARY

Porcine reproductive and respiratory syndrome (PRRS) is currently the most economically important viral disease affecting pig production outside Australia. This study utilised commercial data to perform a genome wide association study looking for single nucleotide polymorphism (SNP) markers associated with PRRS resistance or tolerance, as assessed indirectly from reproductive traits. In total, phenotypes were measured on 1,545 sows, with the data split according to whether the trait was measured during a healthy PRRS-free phase on the farm (4,378 litters from 1,019 sows) or a diseased phase (1,977 litters from 1,526 sows). All animals were genotyped using the Illumina porcine 7k SNP chip. Associations between each individual SNP and reproductive outcomes were assessed using the residual values from a linear mixed (animal) model analysis of each trait in a series of single SNP analyses. Significant SNP associations were only observed for reproductive traits recorded during the disease phase, implying specificity of identified SNPs to a PRRS active phase. After correction for false positives, six significant SNP markers were identified for piglets born alive, piglets born dead and mummified piglets born per litter. SNP effects were then re-estimated from linear mixed model analyses of the data in which the significant SNPs were fitted as additional fixed effects. The SNPs are generally additive in their mode of action and on average the significant SNPs explain 4.7% of the genetic variation for born alive, 1.6% for born dead and 2.2% for born mummified. After further validation and research, these SNPs may allow breeding of pigs that are more robust in the face of PRRSV infection.

INTRODUCTION

Porcine reproductive and respiratory syndrome (PRRS) is by far the most economically important viral problem to the global pig industry. Although vaccines and control measures exist, the disease is persistently endemic and causes problems in both the growing and breeding herds of infected farms. The associated losses from PRRS virus (PRRSV) infection are considerable due to the impact of the virus on the respiratory system, thus affecting overall productivity in the grower/finisher pigs. It also has effects on the reproductive success of the host. Between-breed genetic variation in response to PRRSV infection has been previously demonstrated several times, however it is only recently that evidence for within-breed host genetic variation has been identified, as assessed by impacts on reproductive performance traits (Lewis *et al.* 2009).

This study seeks to dissect the host genetic variation highlighted by Lewis *et al.* (2009) and identify single nucleotide polymorphisms (SNPs) associated with PRRSV tolerance. This study was done using a porcine 7k SNP chip that was developed by the Roslin Institute, the University of Aarhus (Denmark) and the Sanger Centre (UK) on an Illumina platform. The aim was to

identify significant SNPs for traits of relevance to PRRS, which could potentially be utilized in a breeding program to make commercial lines more robust to PRRSV infection.

MATERIALS AND METHODS

The data were collected from a multi-line multiplication herd that had continual recording of the main herd descriptors and production traits. This dataset has been previously described by Lewis *et al*. (2009). DNA was extracted by PIC/Genus, and all SNP genotyping was done at the Wellcome Trust Clinical Research Facility of the Western General Hospital, University of Edinburgh. Individual litters in the data were then assigned to two groups (baseline vs. diseased) using the methodology described by Lewis *et al*. (2009). There were 821 sows that had records common to both data sets.

The SNP genotype data were subjected to quality control (QC) measures. Genotyping was done in two batches of 864 sows and 768 sows respectively for 6,523 SNP markers. The first level of QC removed all of the SNPs that failed to call (1,072), then the SNP markers with a minor allele frequency (MAF) less than 0.01 were removed. Individual sows that had more than 40% of their SNPs uncalled were also removed from the data (~40 individuals). This left 1,545 individual sows recorded for reproductive performance, genotyped with 4,595 SNPs.

The SNP association analyses were done separately on the baseline and disease data for all traits: number of services, total services to conception, gestation length, total piglets *in-utero*, born alive, born dead (mummified + stillborn), mummified piglets (log+1 transformed as it was a non-normally distributed trait), stillborn piglets, total weaned piglets and lactation length. The approach used was that described by Aulchenko *et al.* (2007), i.e. regressing residuals obtained from a mixed model analysis of each trait on the SNP genotypes, performing the regression one SNP at a time. The residuals were obtained from the estimation of the genetic parameters presented in Lewis *et al.* (2009) fitting an animal model including all pedigree and the fixed effects of line (11 levels including European e.g. Large White, Landrace etc. and Asian breeds e.g. Meishan) and parity (seasonal effects were explored and found non-significant). The SNP association analysis was done utilizing the GenABEL package in R. To correct for multiple testing, p-values were corrected using permutation (10,000 iterations) to identify the genome-wide significant SNPs (corrected p-value < 0.05). SNPs that were significant from the genome-wide association analysis were further explored in mixed model single and multi-SNP analyses. These analyses were used to estimate additive and dominance effects for each SNP, to test for SNP interactions (pair-wise combinations to test independence) and also SNP by line interactions. The significant SNPs were then added individually to the model to determine their direction and size of effect. Finally, the proportion of genetic variation explained by each significant SNP was calculated. The additive genetic variance due to the SNP was calculated as $2pq[a+d(q-p)]^2$ (Falconer and MacKay, 1996) and the total additive genetic variance was obtained from the animal model analyses of the same trait, in a model ignoring SNP effects.

RESULTS

A full description of the data is found in Lewis *et al*. (2009). Briefly, the mean and phenotypic standard deviation values of the baseline data for all traits were well within the bounds of what is expected on a commercial farm. The impact of disease on the farm was marked. The PRRS outbreak had a significant ($P<0.001$) impact on the numbers of piglets born alive (10.3 in baseline vs 9.0 during disease), piglets mummified (0.04 in baseline vs 1.13 during disease), and piglets born dead (0.59 in baseline vs 2.15 during disease). The losses highlighted due to PRRSV infection hint at the costs associated with the disease. Indeed, piglets weaned per litter is a common measure of whole farm performance and this also decreased significantly ($P<0.05$) in the

presence of the disease (9.57 in baseline vs 8.39 in the disease data). Further, variability (i.e. the standard deviation) increased in most traits in the presence of the disease.

Details of the significant SNPs (from the single SNP analysis) are shown in Table 1. After permutation testing there were no significant SNPs for any trait measured during the baseline phase; all significant SNPs were for traits affected by PRRS. In total six significant SNP were found from genome wide association with four SNPs affecting more than one trait. Significant SNPs were for: born alive (1 SNP), born dead (6 SNPs) and mummified piglets (4 SNPs), all during the disease phase. There were no significant SNP effects for all other traits investigated.

Upon examination of the two-way SNP interactions, in each case the SNP effects were statistically significant, i.e. they had independent effects on the trait of interest, and none of the interactions between SNP and sow lines were significant. The results indicate that the fixed effects of line were successfully removed as there were no interactions between the SNP and line. The proportion of additive genetic variance explained by each SNP for each trait is also shown in Table 1. The SNPs identified in this study explain a relatively small proportion of the additive genetic variance (across lines). It should be noted that the size of effect and variance explained by the significant SNPs may decrease in validation populations.

Table 1. The estimated effects of significant SNP markers identified in the disease phase data, along with the proportion of genetic variation explained by each of the six significant correction for genome wide significance using permutation

Trait - Total Born:	SNP	P-value raw	P-value*	Minor Allele Freq.	Additive effect	Dominance effect	Proportion of genetic variance explained
Alive	3393	$3.30E^{-05}$	0.08	0.332	0.285	0.575	0.047
Dead	3393	$1.16E^{-10}$	<.001	0.332	-0.055	0.005	0.012
	297	$3.10E^{-07}$	<.001	0.026	-0.210	0.050	0.018
	2589	$6.82E^{-06}$	0.01	0.031	-0.115	0.055	0.006
	870	$9.67E^{-06}$	0.02	0.229	-0.080	-0.010	0.029
	382	$2.13E^{-05}$	0.04	0.039	0.025	-0.015	0.001
	1479	$4.78E^{-05}$	0.09	0.228	-0.050	-0.070	0.030
Mummies	3393	$9.71E^{-07}$	0.001	0.332	-0.070	-0.050	0.029
	382	$1.33E^{-06}$	0.001	0.039	-0.200	-0.030	0.009
	2589	$4.75E^{-06}$	0.009	0.031	-0.150	0.000	0.022
	297	$4.31E^{-06}$	0.08	0.026	0.085	-0.055	0.027

* P-values after permutation for genome wide significance, using 10,000 iterations.

DISCUSSION

This study is the first genome wide SNP association study for PRRS resistance or tolerance in pigs. The study found that significant SNPs can be identified that are associated with disease tolerance using data collected from an outbreak in a commercial herd. The quantity of data needed for a SNP association study of this type, or indeed to estimate genetic parameters for any type of disease resistance or tolerance is considerable. Large numbers of animals with detailed phenotypic records and well managed DNA collections are all needed for a successful study. Indeed, studies of this type are still in their infancy and a major recommendation from this study is that high quality phenotypic data is needed to dissect the traits of interest. The difficulty of detecting SNPs that meet genome-wide significant thresholds is highlighted by the finding that no significant

SNPs were detected for performance traits measured during the PRRS-free period. The data quality issue for this and other studies also extends to the level of the SNP chip. The 7k SNP chip was utilized within this study as at the time it was the best available, however, one could argue that this chip is inadequate as there are large areas of the genome that are not covered and it is likely that there are mutations contributing to the trait variation that are not in LD with any of the SNPs on the chip. A denser chip would be required for a true genome-wide association analysis and likewise for genomic selection.

The choice of dependant variable, i.e. using the residual, was proposed by Aulchenko *et al*. (2007) and it was utilized because all fixed effects (in this case line and parity) are removed, the background polygenic variation is removed (by fitting the pedigree), and genetic covariances between related individuals are accounted for (thus reducing false positives due to population structure). Effectively, associations between the SNP and the within-family Mendelian segregation term are being investigated. This enabled a robust association scan for SNPs associated with the traits of interest across the whole population removing, wherever possible, SNPs associated within specific lines or groups of animals. Another major advantage from utilizing this methodology is that it is very fast computationally, thus saving considerable time in the analysis of each trait. A major issue with the data presented here is that we did not find any SNPs associated with any of the underlying traits (in the absence of disease).

SNP 3393 had significant effects on piglets born alive, born dead, and mummies, but was not significant for total born. Since there is no evidence of PRRSV affecting ovulation rate then the SNP effects for born alive and born dead must be opposite for the SNP effect on total born to be non significant. This can be observed to be the case from the results in Table 1, providing some support for the estimated SNP effects to be biologically meaningful, given that PRRS infected sows have abnormally high numbers of mummies and therefore lower born alive.

Now that SNP markers for PRRS tolerance have been identified in this data, the next step is to validate these in another dataset. Ideally, this would be a PRRS outbreak on another commercial farm. It would also be beneficial to repeat this process using a denser SNP array. Once any significant SNPs are validated, the effects of these SNPs on other traits important to production (e.g. FCR or growth rate) need to be examined to determine whether selection for these markers would have detrimental effects on performance traits. If all of these safeguards are cleared successfully then marker assisted selection could be used to create lines that are more robust to PRRSV infection. Further research is needed; however identification of causal genes may also shed light on the mechanisms of host resistance to PRRS.

ACKNOWLEDGEMENTS

This work was financially supported by the BBSRC (grant EGA16307) and Genus plc. The authors would also like to thank Dr. Kim Bunter for critical revision and constructive comments.

REFERENCES

Aulchenko, Y.S., De Koning, D.J. and Haley, C.S. (2007) *Genet*. **177**:577.

Falconer, D.S. and MacKay, T.C. (1996) "Introduction to Quantitative Genetics" 4th ed. Pearson Education Ltd., Essex, UK.

Lewis, C.R.G., Torremorell, M., Galina-Pantoja, L. and Bishop, S.C. (2009) *J. Anim. Sci*. **87**:876.

Proc. Assoc. Advmt. Anim. Breed. Genet. 18:191-194

EFFECT OF THE HALOTHANE GENOTYPE ON GROWTH PERFORMANCES, CARCASE AND MEAT QUALITY TRAITS IN THE PIETRAIN BREED OF THE FRENCH NATIONAL PIG BREEDING PROGRAM

I. Mérour[1], S. Hermesch [2], S. Schwob [1,3] and T. Tribout [3]

[1] IFIP: French Institute for pig and pork industry, BP 35104 – 35651 Le Rheu Cedex, FRANCE
[2] AGBU: Animal Genetics and Breeding Unit*, University of New England, Armidale, NSW 2351
[3] INRA: Animal Genetics and Integrated Biology Research Unit – 78352 Jouy en Josas, FRANCE

SUMMARY

The halothane allele (n) is segregating in the French national Pietrain breed. Records from the three French central test stations were available for 1,557 Pietrain pigs of known halothane gene status (128 NN, 334 Nn and 1,095 nn). Production traits, carcase composition and meat quality measurements were studied to compare the three genotypes and assess the allele effects. Water holding capacity was the trait most affected by the halothane allele (-0.76 phenotypic standard deviations, sd) followed by the length of the carcase (-0.62 sd) and by carcase traits related to leanness and fatness: dressing percentage, fat and muscle depth and weights of back leg, loin, and fat and rind above loin. The magnitude of the effect of the halothane allele varied from 0.21 to 0.47 phenotypic standard deviations for these carcase traits. In addition, performance of the heterozygous genotype was more similar to the homozygous (NN) genotype. Significant differences between the three genotypes were found for ultimate pH but not for colour. Colour (L*-value) was the only trait for which the heterozygous genotype was more similar to the superior homozygous genotype. In comparison to carcase and meat quality traits, the halothane allele effect was lower for production traits (from -0.12 to 0.02 sd).

INTRODUCTION

Following the new grading criteria and payment scheme in France implemented in 2006, the use of Pietrain pigs as terminal boars has increased at the expense of terminal hybrid boars. In 2007, 69% of semen doses sold to production farms were from the Pietrain breed. The halothane positive allele (n) is segregating in this breed at a high frequency. The effect of the halothane allele in pigs has been studied for nearly thirty years and it is well established that this allele influences carcase and meat quality traits (Aalhus *et al.* 1991; Guéblez *et al.* 1995; Hanset *et al.* 1995 and Larzul *et al.* 1997). However, these past studies were based on data from crossbred populations and the magnitude of the effect of the halothane allele may have been affected by selection. The accuracy of genetic evaluation can be improved if genetic evaluation models are adapted to account for major gene effects explicitly (Tier and Bunter 2003). Thus, the aim of this study was to investigate and quantify the effect of the halothane allele on growth, carcase and meat quality traits in the French national Pietrain breed.

MATERIALS AND METHODS

Data were recorded on female pigs between 2002 and 2008 in three French test stations located in Argentré, Le Rheu and Mauron. The purebred Pietrain animals were sourced from seven herdbook farms participating in the French national breeding scheme. Each herd had at least two halothane genotypes represented. Pigs from each herd were tested in no less than two stations. Pigs arriving at each station at the same time formed a batch, which consisted of a minimum of two herds. Animals from the same herd were housed in groups of 12 animals. The halothane genotype

* AGBU is a joint venture of NSW Department of Primary Industries and University of New England

(NN, homozygous halothane-negative; Nn, heterozygous halothane-negative; nn, halothane-positive) was deduced from the genotypes of their parents or, if this was not possible, was determined using a DNA test (Fujii *et al.* 1991). A total of 1,557 pigs (128 NN, 334 Nn and 1,095 nn) descending from 399 sires were tested within stations. No data recorded on farm were used in this study.

Pigs arrived at stations at the maximum age of five weeks and given *ad libitum* access to feed. However, performance testing was from 35 to 105 kg body weight. Daily feed intake (**DFI**), feed conversion ratio (**FCR)** and average daily gain (**ADG**) were recorded for this test period. Empty body weight was recorded after a fasting period of 16 hours to derive dressing percentage in the abattoir. Pre-slaughter management of pigs and transport time (around 35 minutes) were similar for the three stations. Pigs were slaughtered in two commercial abattoirs. Fat and muscle depth (**FD** and **MD**) between the third and fourth last ribs were recorded using a fat and lean sensor (Sydel CGM – reflectance measurements). Carcases were allowed to chill for approximately 24 hours at 4ºC. Dressing percentage was defined as the ratio of cold carcase weight to empty body weight. Carcase length was measured from the atlas to the anterior edge of the pubic symphysis. The right half of each carcase was submitted to a normalised cutting procedure (Métayer and Daumas 1998) and weights of the back leg, the loin with the skin and fat trimmed, the shoulder, the belly and the fat and rind above the loin (**backfat**) were recorded. Meat quality measurements were taken 24 h *post mortem*. Ultimate pH (**pHu**) was recorded on the *semimembranosus* muscle using a Sydel or Knick pH meter. Meat colour (**L***) was assessed on the *gluteus medius* (GM) using a Minolta CR-300 photocolorimeter (a lower value of L* is associated with darker meat). Water holding capacity (**WHC**) scores, measured at the freshly cut surface of the GM muscle, were determined by the time (in tenths of seconds) required to wet completely one cm^2 of filter paper; a larger value is associated with better WHC.

Contemporary group (61 levels) defined as a concatenation of year, batch and station along with genotype (three levels) were fitted as class effects for all traits. Cold carcase weight was included as a linear covariable in the model used for all carcase traits except for dressing percentage. A Log Likelihood ratio test was used to evaluate the significance of random effects. Effect of the herd in which a piglet was born proved to be significant for all traits except feed conversion ratio, weight of shoulder and meat quality traits. So, herd was included as a random effect. As there were few animals per genotype and per slaughter day nested within batch, slaughter day was included as a random effect in the model for the three meat quality traits. An animal model did not converge and sire linked with pedigree information was fitted as a random effect to account for additive genetic effects. On average, there were 1.2 animals per litter tested in station and litter effect was not included in the model. The ASReml program (Gilmour *et al*. 2006) was used to predict significant differences between genotypes and to estimate the additive (a) and the dominance (d) effects of the n allele along with the estimates of their standard errors, which were used to derive the 95% confidence limits for each estimate. A number of alternative models were evaluated but showed minimal differences between estimates.

RESULTS AND DISCUSSION

Production traits. Heterozygous pigs grew faster than homozygous NN pigs, whereas homozygous nn pigs had significantly better feed conversion ratio than the other two genotypes (Table 1). Hanset *et al.* (1995) also found that heterozygous pigs had the highest growth rate. In comparison, other studies reported no significant differences between the three genotypes (Guéblez *et al*. 1995) or a better growth rate for NN pigs (McPhee *et al*. 1994). Similar to results from this study, the nn genotype had the lowest feed intake and feed conversion ratio in the studies by McPhee *et al*. (1994) and Guéblez *et al*. (1995).

Table 1. Phenotypic standard deviations (σ_p), predicted values with standard errors for each halothane genotype, allele effect (a=((nn-NN)/2)/σ_p) and dominance (d=(Nn-5*(nn+NN))/σ_p), both expressed in σ_p units

Traits[1]	σ_p	NN	Nn	nn	a	d
ADG (g/d)	87.0	822.1 ± 15.7[a]	843.0 ± 13.38 [b]	834.6± 12.83 [b]	0.07	0.17*
FCR (kg/kg)	0.17	2.53 ± 0.02[a]	2.52± 0.01[a]	2.49± 0.008 [b]	-0.12*	0.06
DFI (kg/d)	0.20	2.08± 0.04[ab]	2.12 ± 0.03 [a]	2.07± 0.03 [b]	-0.02	0.22*
Dressing %	1.13	81.4 ± 0.15 [a]	81.7 ± 0.11 [b]	82.4 ± 0.10 [c]	0.47*	-0.13
Length (mm)	25.3	971.0 ± 3.67 [a]	962.3 ± 2.71 [b]	939.6 ± 2.44 [c]	-0.62*	0.27*
Back leg wt (kg)[2]	0.44	11.12 ± 0.06 [a]	11.21 ± 0.04 [b]	11.43 ± 0.04 [c]	0.36*	-0.14
Belly wt (kg) [2]	0.37	4.52 ± 0.05 [ab]	4.58 ± 0.04 [a]	4.47 ± 0.03 [b]	-0.06	0.21*
Shoulder wt(kg)[2]	0.36	9.11 ± 0.04	9.09 ± 0.02	9.13 ± 0.02	0.03	-0.07
Loin wt (kg) [2]	0.54	11.78 ± 0.07 [a]	11.83 ± 0.05 [a]	12.00 ± 0.04 [b]	0.21*	-0.10
Backfat wt (kg) [2]	0.33	2.29 ± 0.05 [a]	2.20 ± 0.04 [b]	1.98 ± 0.04 [c]	-0.46*	0.20*
FD (mm)	1.70	11.88 ± 0.26 [a]	11.31 ± 0.19 [b]	10.44 ± 0.18 [c]	-0.42*	0.10
MD (mm)	4.93	64.38 ± 0.63 [a]	65.87 ± 0.40 [b]	67.76 ± 0.34 [c]	0.35*	-0.05
pHu	0.14	5.59 ± 0.01 [a]	5.62 ± 0.009 [b]	5.64 ± 0.007 [c]	0.18*	0.06
WHC (scores)	3.22	6.73 ± 0.32 [a]	3.44± 0.19 [b]	1.85± 0.12[c]	-0.76*	-0.26*
L* value	3.46	51.04 ± 0.39 [a]	51.50 ± 0.24 [a]	53.47 ± 0.19 [b]	0.36*	-0.21*

Within row, predictions with different superscript letters were significantly different ($P < 0.05$). *Additive and dominance effects of n allele were significant ($P < 0.05$). [1]Abbreviations for traits: ADG: Average Daily Gain; FCR: Feed Conversion Ratio; DFI: Daily Feed Intake; FD: Fat Depth; MD: Muscle Depth; WHC: Water Holding Capacity; pHu: ultimate pH. [2]Weight based on right half of carcase only.

Carcase traits. The three halothane genotypes differed significantly for dressing percentage, length, weight of back leg and backfat, fat depth and muscle depth. The magnitude of the allele effect varied from 0.35 to 0.62 phenotypic standard deviations for these traits which correspond well with the effects reported by Larzul *et al*. (1997) for dressing percentage, carcase length and fat depth measurements. In contrast, larger differences were found between NN and nn pigs for ham and loin weight in the studies by Guéblez *et al*. (1995) and Hanset *et al*. (1995). Shoulder weight was not affected by the halothane allele (Table 1). For all carcase traits affected by the halothane allele, heterozygotes were closer to the inferior NN than to the superior nn genotype, which was also observed by Larzul *et al.* (1997) for comparable traits.

Meat quality traits. As expected, halothane sensitivity was accompanied by worse overall meat quality. Relative to other traits in this study, the largest effect of the n allele was on water holding capacity (-0.76 sd). In comparison, the magnitude of the allele effect on water holding capacity was substantially lower (-0.37 sd) in the study by Larzul *et al*. (1997). A scoring system was used for WHC in this study which superseded the paper wetting time method used in Larzul et al. (1997). The heterozygous genotype was closer to nn carcases for water holding capacity, which was also reported by Larzul *et al*. (1997) and Guéblez *et al*. (1995). Significant differences between the three genotypes for ultimate pH were found in this study confirming results from a recent meta-analysis (Salmi *et al*. 2009). In contrast, other studies (Fàbrega *et al*. 2004; Larzul *et al*. 1997) reported that the n allele had no significant effect on ultimate pH. With respect to the L*-values, nn carcases had paler meat colour in the *gluteus medius* than both Nn and NN carcases, which did not differ significantly from each other. Among all the traits investigated, this meat colour measurement was the only trait for which the heterozygous carcases were closer to the superior homozygous genotype.

CONCLUSIONS

The halothane allele affected significantly most of the traits considered in this study. The magnitude of the allele was highest for water holding capacity and carcase traits related to leanness and fatness measurements. Performance in these carcase traits of the heterozygous genotype was more similar to the inferior homogeneous (NN) genotype for these carcase traits. The current study presents results based on purebred data for a wider range of traits, improving knowledge from historical analyses conducted more than ten years ago on crossbred data. The estimates from this study should be used in genetic evaluations that incorporate the halothane allele effect explicitly (Tier and Bunter 2003).

ACKNOWLEDGMENTS

This paper was prepared while Isabelle Mérour was on sabbatical leave from IFIP to AGBU. The authors thank staff at test stations of Argentré, Le Rheu and Mauron for diligent data recording. The analyses at AGBU were funded by Australian Pork Limited under project APL2133.

REFERENCES

Aalhus, J.L., Jones, S.D.M., Robertson, W.M., Tong, A.K.W. and Sather, A.P. (1991) *Anim. Prod.* **52**:347.

Fàbrega, E., Manteca, X., Font, J., Gispert, M., Carrión, D., Velarde, A., Ruiz-de-la-Torre, J.L. and Diestre, A. (2004) *Meat Sci.* **66**:777.

Gilmour, A.R., Gogle, B.J., Cullis, B.R. and Thompson, R. (2006) "ASReml User Guide Release 2.0" VSN International Ltd, Hemel Hempstead, HP1 1ES, UK.

Guéblez, R., Paboeuf, F., Sellier, P., Bouffaud, M., Boulard, J., Brault, D., Le Tiran, M.H. and Petit, G. (1995). *J. Rech. Porcine Fr*. **27**:155.

Hanset, R., Dasnois, C., Scalais, S., Michaux, C and Grobet, L. (1995) *Genet. Sel. Evol*. **27**:63.

Fujii, J., Otsu, K., Zorzato, F., De Leon, S., Khanna, V.K., Weiler, J.E., O'Brien, P.J., and Maclennan D.H. (1991) Science **253:**448.

Larzul, C., Le Roy, P., Guéblez, R., Talmant, A., Gogué, J. and Sellier, P. (1997) *J. Anim. Breed. Genet*. **114**:309.

McPhee, C.P., Daniels, L.J., Kramer, H.L., Macbeth, G.M and Noble, J.W. (1994) *Livest. Prod. Sci*. **38**:117.

Métayer, A. and Daumas, G. (1998) *J. Rech. Porcine Fr*. **30**:7.

Salmi, B., Bidanel, J.P. and Larzul, C (2009) *J. Rech. Porcine Fr*. **41**:19.

Tier, B. and Bunter, K. (2003) *Proc. Assoc. Advmt. Anim. Breed. Genet*. **15**:214.

Proc. Assoc. Advmt. Anim. Breed. Genet. 18:195-198

GENETIC CORRELATIONS BETWEEN CARCASE LENGTH, FAT AND MUSCLE DEPTHS AND PRIMAL CUT WEIGHTS IN THE FRENCH LARGE WHITE SIRE LINE

I. Mérour[1], S. Hermesch[2], R.M. Jones[2] and T. Tribout[3]

[1] IFIP: French Institute for pig and pork industry, BP 35104 – 35651 Le Rheu Cedex, France
[2] AGBU: Animal Genetics and Breeding Unit[*], University of New England, Armidale, NSW 2351
[3] INRA: Animal Genetics and Integrated Biology Research Unit – 78352 Jouy en Josas, France

SUMMARY

Records for the French national Large White sire line were collected between 1999 and 2008 for 65,082 pigs on farm and for 2,429 carcases of siblings measured in three test stations. Ultrasonic measures of backfat and muscle depth were recorded *in vivo* on farm. In addition, fat and muscle depth as well as length were recorded on carcases of littermates. Weights of primal cuts included back leg, loin with the skin and fat trimmed as well as shoulder and belly weights. Heritability estimates were moderate to high for the four primal cuts, ranging from 0.21 for shoulder to 0.46 for back leg weight. Estimates of genetic correlations (ra) showed back leg weight was genetically independent from loin (ra: 0.06) and shoulder (ra: 0.01) weights. Belly weight was negatively correlated with the weight of the three other cuts (ra: -0.57 to -0.33). The heritability of carcase length was high (0.64) but the genetic correlations between this trait and weights of primal cuts were low (ra: 0.05 to 0.26) limiting its use as a potential selection criterion for these traits. Estimates of genetic correlations between fat and muscle depth with primal cut weights were generally higher for fat and muscle depth measured on the carcase in contrast to *in vivo* measures.

INTRODUCTION

In pig production, genetic selection has greatly improved body composition towards a higher lean meat content and a lower fat deposition. Breeding goals are subject to change and may be directed towards yield of primal cuts to better reflect the market value of the carcase. Weights of primal cuts vary for carcases with a fixed weight and fat depth. This variation in weights of primal cuts resulted in a higher return per pig of $7 (farm gate level) and $21 (wholesale/retail level) for the top 10% of carcases in comparison to the group average (Mérour and Hermesch 2008). Currently, pig producers would receive a similar price for these carcases in France and Australia. However, new technologies are being developed in both countries to better quantify variation in weights and quality of primal cuts. Selection for weight of primal cuts requires genetic parameters for these traits as well as other traits that may be used as selection criteria. Few studies about genetic parameters for back fat and muscle depth measured *in vivo* and on the carcase, carcase length and primal cuts weights have been found in the literature. Estimates of genetic parameters for these traits are presented in this study.

MATERIALS AND METHODS

Data for this study were based on the French national Large White sire line (LWM) developed in France since 1998 based on the French national Large White dam line. The aim was to create a new breed complementary to the Pietrain breed to produce a terminal boar with not only good conformation but also high growth rate. Since 2000, there has been no gene transfer between the French national Large White sire and dam lines.

Performance recording of the national breeding program in France is based on on-farm testing combined with central testing. Siblings of on-farm tested animal candidates are sent to test stations

[*] AGBU is a joint venture of NSW Department of Primary Industries and University of New England

for post-slaughter recording. Data for this study consisted of 65,082 records from ten herds and 2,429 records from three stations collected between 1999 and 2008. Males and females tested on farm were housed in pens of 10 to 15 pigs and fed *ad libitum*. All farms practiced batch management with a batch consisting of at least 18 animals from the same sex and minimal variation in age (maximum range of 15 days). Ultrasonic backfat thickness (**USBF**) and muscle depth (**USMD**) were measured four cm right and left from the chordal spine between the third and forth last ribs at an average live weight of 97.2 kg (±10.7 kg). For each of the latter two traits, the right and left measurements were averaged for analysis. Litter mates (only castrates) of pigs tested on farm were sent to the three test stations. Pigs from each herd were tested in at least two stations. Pigs arriving at the station over a period of two weeks formed a batch, which consisted of at least two herds. Pigs were a maximum age of five weeks at arrival. Animals from the same herd were housed in groups of 12 animals and fed *ad libitum* with individual electronic feeders. When pigs reached the target weight of 105 kg (108.7±6.13), they were slaughtered in one of two commercial abattoirs. Fat and muscle depths (**CBF** and C**MD**) between the third and fourth last ribs were recorded on the carcases using a fat and lean sensor (Sydel CGM – reflectance measurements). Carcase length (**LEN**) was measured 24 h *post mortem* from the atlas to the anterior edge of the pubic symphysis. The right half of each carcase was submitted to a normalised cutting procedure (Métayer and Daumas 1998), and weights of primal cuts including back leg (**LEG**), loin with the skin and fat trimmed (**LOIN**), shoulder (**SHLDR**) and belly (**BELLY**) were recorded.

The GLM (SAS 1999) procedure was used to derive the fixed effect model for each trait. Variance and covariance components were estimated in univariate and bivariate analyses using ASReml (Gilmour *et al*., 2006). On-farm contemporary group (1455 levels) combining herd, year, batch and sex was fitted for ultrasonic measurements (USBF and USMD) along with the linear covariable of live weight and the direct additive genetic and permanent environment of the litter effects. Contemporary groups (99 levels) in station tested animals were based on year, batch and station and was fitted for primal cut weights and linear carcase measurements (LEN, CBF and CMD) along with the linear covariable of cold carcase weight (83.2±5.26 kg). Random effects were the direct additive genetic and common litter effects for on-farm traits, whereas the litter effect was not significant for traits recorded in station.

RESULTS AND DISCUSSION

Heritabilities. Moderate to high heritability estimates were obtained for carcase traits (Table 1). Primal cuts with higher lean to fat ratio (LOIN and LEG) were more heritable (0.43 and 0.46) than belly and shoulder (0.35 and 0.23). Definition of primal cuts differed between studies and it is difficult to directly compare these heritability estimates with literature values. Nevertheless, Johansson *et al*. (1987) reported heritabilities of 0.43, 0.40 and 0.35 for ham, loin and shoulder percentages for a Yorkshire breed, whereas van Wijk *et al*. (2005) estimated heritabilities of 0.40 and 0.29 for leg and loin weights. The heritability estimate of carcase length (0.64) was slightly higher than literature values, ranging from 0.44 to 0.62 (Engellandt *et al***.** 1997; Johansson *et al*. 1987). Other studies (Johnson and Nugent III 2003; Nakavisut *et al*. 2006) reported lower heritabilities of 0.12 to 0.41 for body length. Differences in the measurement position may have contributed to variation in estimates. The heritability of backfat depth measured *in vivo* was lower than the one measured on carcase (0.37 versus 0.59), whereas estimates were the same for both *in vivo* and *post mortem* muscle depth measures (0.30). The litter effect estimates were 0.04±0.003 and 0.04±0.004 for ultrasonic backfat and muscle depth recorded on farm.

Table 1. Number of records, means and coefficients of variation (CV), heritability estimates (h^2) with standard errors (s.e.) and phenotypic variance (σ^2_p) for carcase traits and *in vivo* depth

Trait (unit, abbreviation)	N	Mean	CV	h^2	s.e	σ^2_p
Leg weight (kg, LEG)	2,422	9.8	7.1 %	0.46	0.06	0.15
Loin with skin and fat trimmed (kg, LOIN)	2,426	10.9	8.8 %	0.43	0.06	0.28
Belly weight (kg, BELLY)	2,424	4.7	11.8 %	0.35	0.06	0.14
Shoulder (kg, SHLDR)	2,422	9.3	6.8 %	0.23	0.05	0.12
Carcase length (mm, LEN)	2,418	984	3.0 %	0.64	0.06	668
Fat depth abattoir (mm, CBF)	2,396	15.0	20.5 %	0.59	0.06	8.11
Muscle depth abattoir (mm, CMD)	2,400	54.4	10.1 %	0.30	0.05	20.3
Ultrasound *in vivo* backfat (mm, USBF)	65,082	9.1	19.2 %	0.37	0.01	1.45
Ultrasound *in vivo* muscle depth (mm, USMD)	55,036	52.7	10.1 %	0.30	0.01	12.5

Genetic correlations. Leg weight was genetically uncorrelated with loin weight, suggesting that selection for high leg weight does not result in a high loin weight (Table 2). In comparison, van Wijk *et al.* (2005) found a genetic correlation of 0.31 between leg and loin weight. The weights of the two valuable cuts, leg and loin, had negative genetic correlations with belly (-0.49 and -0.57) and fat depth measurements (range from -0.58 to -0.35). Loin and belly weights had stronger genetic correlations with carcase muscle depth (0.55 and -0.30) than live muscle depth (0.23 and 0.03). Both muscle depth traits had no genetic association with leg weight. These genetic correlations were considerably lower than comparable estimates published by Hermesch *et al*. (2000) and van Wijk *et al.* (2005) for trait combinations involving weights of loin and leg along with fat and muscle depth traits. Due to the cost of recording these traits, the sizes of data sets were limited and there was a large range in standard errors of genetic correlations from 0.05 to a maximum of 0.38.

Shorter carcases were genetically associated with heavier legs (-0.24) and lighter loins (0.26), both adjusted for carcase weight. Therefore, selection for shorter length would lead to muscular pigs with proportionally larger ham weights similar to the characteristics of the Pietrain breed. However, these aspects of conformation of pigs were not necessarily reflected in genetic correlations between carcase length and the four fat and muscle depth traits. These estimates were all negative, ranging from -0.30 to -0.20 for backfat and -0.42 to -0.39 for muscle depth. Nakavisut *et al*. (2006) also reported negative genetic relationships between body length and backfat or muscle depth while Engellandt *et al*. (1997) found negative genetic correlations between carcase length and lean meat percentage (-0.38) or *longissimus dorsi* area (-0.37).

CONCLUSIONS

Estimates of heritabilities of primal cut weights were moderate to high in the Large White sire line, implying that the weight of primal cuts and ultimately market value of the carcase can be improved via selection. Given the higher genetic correlations with primal cuts weight along with similar or larger heritabilities and variation, measurements of fat and muscle depth on the carcase may be better selection criteria for loin and back leg weights than *in vivo* measurements. However, *in vivo* measurements can be recorded prior to selection often on more animals. Index calculations are required to evaluate implications of these genetic parameters for pig breeding programs. Genetic correlations between primal cut weights and carcase length were of lower magnitude than estimates between fat or muscle depth and primal cut weights.

Table 2. Genetic correlations (above diagonal) along with environmental (r_e) and phenotypic correlations (r_p; both below diagonal) between carcase traits[1]

Traits[2]		LEG	LOIN	BELLY	SHLDR	LEN	CBF	CMD	USBF	USMD
LEG			0.06	-0.49	0.01	-0.24	-0.40	0.04	-0.35	-0.05
LOIN	r_e	-0.10		-0.57	-0.22	0.26	-0.58	0.55	-0.47	0.23
	r_p	-0.03								
BELLY	r_e	-0.12	-0.23		-0.33	0.05	0.37	-0.28	0.39	0.03
	r_p	-0.27	-0.36							
SHLDR	r_e	-0.08	-0.15	-0.15		-0.08	-0.16	-0.30	-0.07	-0.25
	r_p	-0.05	-0.20	-0.20						
LEN	r_e	-0.08	0.27	-0.06	0.02		-0.30	-0.42	-0.20	-0.39
	r_p	-0.16	0.26	-0.01	-0.02					
CBF	r_e	-0.22	-0.45	0.14	0.001	-0.24		0.04	0.73	0.19
	r_p	-0.31	-0.51	0.24	-0.06	-0.28				
CMD	r_e	0.09	0.05	-0.05	-0.01	-0.07	-0.21		-0.17	0.70
	r_p	0.07	0.23	-0.12	-0.09	-0.22	-0.09			

[1] Standard errors of the genetic correlations presented in this table ranged from 0.05 to 0.13.
[2] For traits abbreviations see Table 1.

ACKNOWLEDGMENTS

This paper was prepared while Isabelle Mérour was on sabbatical leave from IFIP to AGBU. The authors thank staff at test stations of Argentré, Le Rheu and Mauron for diligent data recording. The analyses at AGBU were funded by Australian Pork Limited under project APL2133.

REFERENCES

Engellandt, T., Reinsch, N., Reinecke, S. and Kalm, E. (1997) *Züchtungskunde* **69**:39.

Gilmour, A.R., Gogle, B.J., Cullis, B.R. and Thompson, R. (2006) "ASReml User Guide Release 2.0" VSN International Ltd, Hemel Hempstead, HP1 1ES, UK.

Hermesch, S., Luxford, B.G. and Graser, H.-U. (2000) *Livest. Prod. Sci*. **65**:249.

Mérour, I. and Hermesch, S. (2008) In *"Pig Genetics Workshop Notes"*, p.83, AGBU, UNE, Armidale, NSW, Australia.

Johansson, K., Andersson, K. and Sigvardsson, J. (1987) *Acta Agric. Scand*. **37**:120.

Johnson, Z.B. and Nugent III, R.A. (2003) *J. Anim. Sci*. **81**:1943.

Métayer, A. and Daumas, G. (1998) *J. Rech. Porcine Fr*. **30**:7.

Nakavisut, S., Crump, R.E. and Graser, H.-U. (2006) In *"Pig Genetics Workshop Notes"*, p.25, AGBU, UNE, Armidale, NSW, Australia.

SAS (1999). Enterprise Miner, Release 9.1. SAS Institute, Cary, NC, USA.

van Wijk, H.J., Arts, D.J.G., Matthews, J.O., Webster, M., Ducor, J.B. and Knol E.F. (2005). *J. Anim. Sci*., **83**:324.

Proc. Assoc. Advmt. Anim. Breed. Genet. 18:199-202

EVALUATION OF PIG FLIGHT TIME, AVERAGE DAILY GAIN AND BACKFAT USING RANDOM EFFECT MODELS INCLUDING GROWER GROUP

R.M. Jones, S. Hermesch and R.E. Crump

Animal Genetics and Breeding Unit*, University of New England, Armidale, NSW 2351

SUMMARY

Records for flight time, average daily gain and backfat were analysed from 9,638 pigs measured between April 2004 and November 2007. Genetic parameters were estimated for the three traits using alternative animal models, fitting combinations of additive genetic, common litter and grower group effects fitted as random terms. Log likelihood ratio tests showed that all three random effects were significant for the three traits analysed. Heritability estimates were 0.15±0.02 for flight time, 0.22±0.03 for average daily gain and 0.40±0.04 for backfat. Estimates for common litter (c^2) and grower group (g^2) effects were small (0.02 to 0.05) for flight time and backfat. Conversely, for growth rate estimates of c^2 and g^2 were 0.08±0.01 and 0.17±0.01, respectively. It is recommended that grower group be recorded for further investigations of social genetic effects. Flight time was genetically correlated with backfat (0.21±0.09) but had no significant genetic relationship with growth rate (0.14±0.11). Overall, little correlated response is expected in flight time resulting from selection for higher growth and lower backfat.

INTRODUCTION

Flight time, an objective measure of temperament has been analysed and found heritable in pigs (Hansson *et al.* 2005). It is likely that selection for calmer pigs should be beneficial to animal welfare and possibly also ease stock handling, lowering occupational health and safety risks. In addition, reducing pig stress levels could be beneficial to pork quality by reducing the incidence of pale soft exudative and dark firm dry pork (Guardia *et al*. 2005).

It is likely that the temperament of an individual pig influences the performance of other pigs housed in the same group. In one pig population Bergsma *et al*. (2008) found that social effects contributed the vast majority of heritable variance for growth rate and feed intake. Bergsma *et al*. (2008) used the variance of true breeding values (TBVs) among individuals as the definition of heritable variance. Bijma *et al*. (2007) defined TBVs as the sum of an individual's direct additive variance and n-1 times (n=the number of animals within the group) the individuals social additive variance. Fitting group as an additional random effect without fitting social genetic effects explicitly accounts for heritable and environmental social effects (Bijma, pers. comm.). It was the aim of this study primarily to evaluate various random effect models for flight time, average daily gain and backfat and secondly to estimate genetic correlations between flight time and growth rate or backfat.

MATERIALS AND METHODS

Data from Belmont, a farrow to finish commercial piggery located in Queensland, Australia, were recorded between April 2004 and November 2007 on three purebred breeds of grower pigs (6,072 Large White, 2,795 Landrace, 771 Duroc). Pigs were recorded for backfat (**BF**) at the P2 site using ultrasound, average daily gain (**ADG**) and flight time (**FT**) at 103 (±9.4) kg liveweight (Table 1). In pigs, flight time is the time taken to clear a one metre distance between light sensitive start and stop diodes set 0.25 and 1.25 meters from a weigh scale exit (Crump *et al*. 2005). Pigs were also scored from one to five (by six staff) for the assistance required to move them past the

* AGBU is a joint venture of NSW Department of Primary Industries and University of New England

stop diode. Pigs were housed in one shed and grown in one of 16, 8m by 3m pens. Pens one to eight predominantly housed boars and were fed a higher energy diet (14.5-14.7 MJ DE/kg) than pigs in the other pens (14.0-14.2 MJ DE/kg) which predominantly housed gilts. Animals in the same pen did not always complete their testing at the same time. Grower groups (n=353) were constructed from the date of test (n=128) with animals tested from the same pen within a 15 day period combined into single grower groups. Post editing grower group size ranged from 22 to 37 pigs which were confirmed to match farm group sizes. These data were merged with 35,582 pedigree records extending to January 1995.

Table 1. Characteristics of data including coefficients of variation (CV).

Trait	N	Means	Standard deviations	Min-Max	CV
Flight time (s)	9,460	2.1	1.1	0.3-9.4	54.3
Average daily gain (g/day)	9,606	672	68.6	463-881	10.2
Backfat (mm)	9,468	11.4	2.1	7-18	18.5
Test weight (kg)	9,622	103	9.4	75-132	9.1
Test age (days)	9,638	154	8.3	130-190	5.4

Records exceeding three standard deviations from the mean were deleted as were flight time records equalling exactly eight seconds, which were caused by an equipment failure. Fixed effect models were derived using the GLM procedure (SAS 1999) only retaining significant effects and their interactions. Breed and test month within year were fitted for all traits. Further fixed effects were encouragement score within staff (27 levels) for flight time as well as sex and diet for growth rate and backfat. The model for backfat also included weight as a linear covariate. Variance component estimates for all traits were obtained with univariate animal model analyses using the ASReml software (Gilmour *et al.* 2006). Random effects fitted were a combination of additive genetic, common litter and grower group. The significance of individual random effects was evaluated using a log likelihood ratio test. In addition, log likelihood values were generated for a range of estimates of two random effects by alternatively fixing either common litter or grower group effects at their estimated values. Genetic correlation estimates were obtained from one tri-variate analysis fitting all three random effects.

RESULTS AND DISCUSSION

The heritability of flight time was (0.15±0.02, Table 2) which was similar to previously reported heritability estimates (Hannson *et al.* 2005). Heritability estimates were 0.22±0.03 for growth rate and 0.40±0.04 for backfat. All three random effects were significant for the three traits analysed. Estimates of the ratios of common litter effect (c^2) and group effect (g^2) variances over phenotypic variance were low (0.02 to 0.05) for flight time and backfat. The low g^2 estimate of 0.02±0.005 for flight time (Table 2) implies that this behavioural trait is reasonably independent of the group that the animal was raised in. Conversely, for growth rate the estimates of c^2 and g^2 were 0.08±0.009 and 0.17±0.009, respectively. This relatively large g^2 estimate for ADG is indicative of the group expressing a considerable influence on this trait. Fitting pen as a fixed effect in the growth rate model did not alter the estimates. Using an equivalent model, Bergsma *et al.* (2008) found a larger g^2 effect of 0.27 for growth rate and a similar g^2 effect of 0.04 for backfat. However, in their study, fitting group as an additional random effect for growth rate reduced the heritability estimate from 0.36 to 0.25, whereas heritability estimates were not affected significantly in the current study. Various models were applied to simulated data with additive genetic effects, genetic social effects and group effects by Van Vleck and Cassady (2005). When genetic social and group effects were ignored, as models usually used in pig breeding do, additive genetic variances were

increased slightly but residual variances were substantially inflated, especially in data sets with large simulated group effects. Residual variances in this study were also higher for models that excluded group effects although additive genetic variances were minimally increased. Including grower group in models could reduce the likelihood of inflated residual variances and possibly inflated additive variances by improving variance partitioning.

In addition, Van Vleck and Cassidy (2005) did not expect the large overestimation of group variance when genetic social effects were ignored. The increase in group variance observed by Van Vleck and Cassidy (2005) corresponded well with the theoretical expectation of group variance of twice the covariance between the additive direct and additive social genetic effects plus (n-2) times the genetic social variance where n is the number of penmates (Bijma pers. comm.). Social genetic effects were not fitted explicitly in our study and group represents both additive and environmental social effects.

In models with multiple parameters, change in one parameter will often lead to a corresponding change in other parameters. Figure 1 illustrates the relative robustness of the parameter estimates providing some confidence that the g^2 effect is not a result of a confounding with c^2. These ranges illustrate the possible substitution between random effects until the log likelihood was changed significantly.

Table 2. Heritabilities (h^2), litter (c^2) and grower group (g^2) effects and log likelihoods (LogL) as well as residual and phenotypic variances for flight time (FT), average daily gain (ADG) and backfat (BF) for model 1 (M1: h^2+c^2), model 2 (M2: h^2+g^2) and model 3 (M3: $h^2+c^2+g^2$).

Trait	Model	LogL	h^2	c^2	g^2	Residual variances	Phenotypic variances
FT	1	-5890.23	0.14±0.02	0.03±0.008		1.09±0.02	1.33±0.02
FT	2	-5912.65	0.17±0.02		0.02±0.005	1.07±0.02	1.35±0.02
FT	3	-5884.97	0.15±0.02	0.02±0.007	0.02±0.005	1.07±0.02	1.34±0.02
ADG	1	-3699.63	0.26±0.03	0.12±0.010		2522±89	4062±89
ADG	2	-3391.92	0.31±0.03		0.17±0.014	2248±83	4335±109
ADG	3	-3356.09	0.22±0.03	0.08±0.009	0.17±0.009	2243±78	4207±103
BF	1	-9846.37	0.41±0.04	0.05±0.008		1.93±0.09	3.55±0.09
BF	2	-9803.40	0.45±0.03		0.04±0.006	1.89±0.09	3.63±0.09
BF	3	-9780.22	0.40±0.04	0.04±0.007	0.04±0.006	1.83±0.09	3.57±0.09

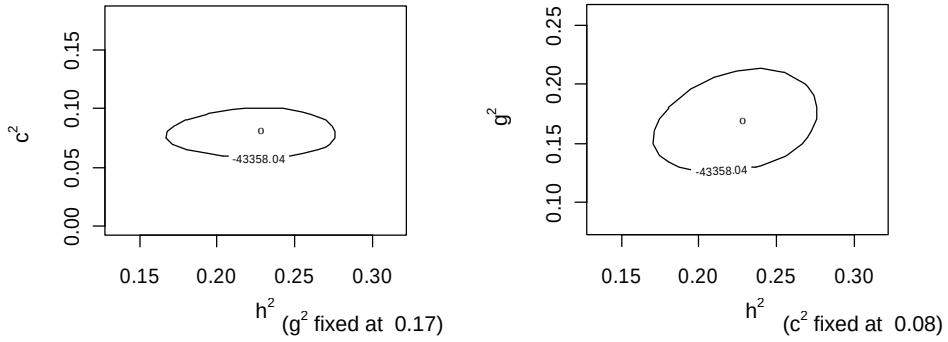


Figure 1. Log likelihood contour lines, illustrating where parameter estimates differ significantly from those obtained when the log likelihood is maximised (Model3) and either group (g^2) or litter (c^2) effects are fixed (trait=ADG).

Flight time was genetically correlated with backfat (0.21±0.09) but had no genetic relationship with growth rate (Table 3) confirming results by Hansson *et al*. (2005). Bunter (2005) also estimated positive genetic correlations between flight time and growth rate (0.34±0.16) or backfat (0.14±0.16). Adjusting flight time for weight did not significantly reduce the genetic correlation with growth rate (0.14±0.11 to 0.02±0.11). Overall, little correlated response is expected in flight time resulting from selection for higher growth and lower backfat.

In beef cattle, flight time is used as a selection criterion for tenderness. Bunter (2005) reported a significant genetic correlation between flight time and pH recorded 24 hour *post mortem* of -0.53±0.21. High final pH indicates dark, firm and dry pork and this genetic correlation suggests that selection for higher flight time would lead to a reduced incidence of dark, firm and dry pork.

Table 3. Genetic and phenotypic correlations between flight time, backfat and average daily gain.

Traits	Genetic	Phenotypic
Flight time – Average daily gain	0.14 ± 0.11	0.08 ± 0.01
Flight time – Backfat	0.21 ± 0.09	0.03 ± 0.01
Average daily gain - Backfat	0.13 ± 0.09	0.10 ± 0.02

CONCLUSIONS

The random additive genetic, litter (c^2) and group (g^2) effects were significant for flight time, growth rate and backfat, although estimates for c^2 and g^2 effects were low (0.02 to 0.05) for flight time and backfat. The g^2 estimate of 0.17 for growth rate indicates that this trait is affected by social effects and further analyses should fit genetic social effects explicitly. In addition, grower group should be recorded to enable analyses of genetic social effects. Flight time was heritable and had low positive genetic correlations with growth rate (not significant) and backfat. Little correlated response is expected in flight time resulting from selection for higher growth and lower backfat.

ACKNOWLEDGMENTS

The authors thank Brenden and Madelene McClelland of Eastern Genetic Resources for making the data available. Analyses of data were funded by Australian Pork Limited under Project 2133.

REFERENCES

Bergsma, R., Kanis, E., Knol, E.F. and Bijima, P. (2008) *Genetics* **178**:1559.

Bijma, P., Muir, W. and Van Arendonk, J. A. (2007) *Genetics* **175**:277.

Bunter, K. (2005) Project APL 1927, Final report for Australian Pork Limited.

Crump, R.E., Hansson, A.C., Graser, H.-U. and Sokolinski, R. (2005) *Proc. Assoc. Advmt. Anim. Breed. Genet.* **16**:91.

Gilmour, A.R., Gogle, B.J., Cullis, B.R. and Thompson, R. (2006) "ASReml User Guide Release 2.0" VSN International Ltd, Hemel Hempstead, HP1 1ES, UK.

Guardia, M., Estany, J., Balasch, S., Oliver, M., Gispert, M. and Diestre, A. (2005) *Meat Sc.* **70**:709.

Hansson, A.C., Crump, R.E., Graser, H.-U. and Sokolinski, R. (2005) *Proc. Assoc. Advmt. Anim. Breed. Genet.* **16**:141.

SAS (1999). Enterprise Miner, Release 9.1. SAS Institute, Cary, NC, USA.

Van Vleck, L.D. and Cassady, J.P. (2005) *J. Anim. Sci*. **83**:68.

Proc. Assoc. Advmt. Anim. Breed. Genet. 18:203-206

ASSOCIATIONS BETWEEN SOW BODY COMPOSITION, FEED INTAKE DURING LACTATION AND EARLY PIGLET GROWTH

K.L. Bunter[1], B.G. Luxford[2], R. Smits[2] and S. Hermesch[1]

[1]Animal Genetics and Breeding Unit[*], University of New England, Armidale, NSW 2351
[2]QAF Meat Industries Pty Ltd, Corowa, NSW 2646

SUMMARY

The genetic and phenotypic associations between sow body composition, early piglet growth and lactation feed intake (LFI) recorded during the first lactation were estimated using data collected from two maternal lines (N~2500). Heritability estimates for lactation feed intake, average piglet birth weight (ABW) and total born (TB) were 0.16±0.04, 0.27±0.03 and 0.10±0.04; genetic correlations between LFI and ABW or TB were positive but not significantly different to zero. Heritabilities for sow weight and fat depths prior to farrowing and at weaning ranged from 0.27 to 0.37 (±0.05) and within trait genetic correlations between these time points were less than one. Positive genetic (r_a) and phenotypic (r_p) correlations show that increased LFI is associated with higher sow weaning weight and fat depths (r_a: 0.52±0.13 and 0.21±0.16; r_p: 0.38±0.02 and 0.15±0.03) and higher litter gain (r_a: 0.10±0.24, r_p: 0.20±0.02). While correlations are not antagonistic between LFI and TB, ABW or litter gain, any correlated response in LFI to selection on these traits would be low.

INTRODUCTION

The group of desirable maternal traits includes large litter size, excellent mothering ability and adequate milk production to ensure high piglet survival and growth, followed by successful sow rebreeding after weaning. There are two major contradictory elements within this trait complex. Firstly, piglets from larger litters are generally lighter at birth, can suffer more hypoxia during farrowing, and have reduced access to colostrum and teats. Piglets from larger litters are therefore more vulnerable to environmental stressors that can result in piglet death (Knol 2001). Secondly, sows that successfully rear large, heavy litters have increased risk of longer weaning to conception intervals and reduced stayability in the herd (Tholen *et al.* 1996). These areas of antagonism are likely strongest for primiparous sows which must balance their own continuing growth and development against reproductive demands. This particular study focused on the genetic relationships between sow lactation intake and body composition, along with early litter growth, which is an area where information for modern sow genotypes is scarce.

MATERIALS AND METHODS

Approximately 2500 sows from two maternal lines (Large White and Landrace based, Primegro™ Genetics) were recorded for their first gestation and farrowing outcomes between January 2007 and June 2008 at QAF Meat Industries, Corowa, Australia. Records available for first parity sows in this study included aspects of sow body condition, described by sow weight and average fat depth at day 110 of gestation (W110, F110) and at weaning (SWW, SFD). Reproductive traits included total number of piglets born (TB), average piglet birth weight (ABW) of live born piglets, along with litter gain from day 1 (after cross fostering) until day 10 (LG10) and the average daily lactation feed intake of the sow (LFI). An estimate of the sow's own body weight (SW110) prior to farrowing was calculated as SW110=W110-(TB×ABW). During lactation feed was delivered 4 times per day to enable expression of appetite. Average lactation

[*] AGBU is a joint venture of the NSW Department of Primary Industries and the University of New England

intake was based on daily records averaged over a maximum lactation length of 35 days. Data were subsequently edited based on trait distributions. The UNIVARIATE procedure (SAS 2003) was used to identify outliers, whereby trait records that deviated by more than 3 times the interquartile range from the mean value were deleted. After editing, there were 2264 animals representing 206 sires and 1268 dams in a pedigree extended back to include all animals born since 2003 (N=53124).

Models for analyses were developed using ASReml software, which estimates variance components under a linear mixed model by residual maximum likelihood (Gilmour *et al*. 2006). Univariate analyses were used to develop models for systematic effects and to obtain initial estimates of genetic parameters under an animal model. Approximate F-tests were used to assess the significance of systematic effects and/or their interactions, only those effects significant at $P<0.05$ were retained. Systematic effects for all traits included year/month of farrowing (20 levels) and sow line (2 levels). Gestational treatment (4 levels) was fitted for W110, SWW and ABW. A factor categorising fostering events prior to day 10 (4 levels) was included in the model for LG10, while lactations of shortened or normal duration (2 levels) were modelled for LFI. Linear covariates included age at mating for all traits except LFI and LG10. The number of piglets on day 1, after cross-fostering, was a linear covariate for LG10 and SWW, while lactation length was fitted as a linear and quadratic covariate for LFI. Common litter effects, if present, were not estimated due to the low number of sows farrowing per source litter. Correlations between specific traits were subsequently estimated using the univariate model for each trait, fitted in a series of bivariate analyses.

RESULTS AND DISCUSSION

Characteristics of the data. The similarity of coefficients of variation (CV~9%) for weight or fat depths observed at mating (not presented), prior to farrowing and at weaning (Table 1) masks the much larger underlying variability between sows in how they transitioned between these time points. Gestational weight and fat gains calculated from this data had CV of 22% and more than 200%, respectively. Lactation feed intake averaged 4.99 kg/day (~2.4% of sow body weight), higher than was observed in a previous subset of this data predominantly recorded throughout summer (Bunter *et al*. 2007). The CV for TB (27%) was larger than the CV for ABW (17%), possibly indicating that piglet birth weight is under some form of physiological regulation to reduce variability generated by differences in litter size. In contrast, litter gain was highly variable (CV=61%), reflecting both variation in piglet losses and the weight gain of surviving piglets.

Table 1. Characteristics of the data after editing for outliers

Trait	Abbreviation	N	Mean (SD)	Min-Max	h^2	σ_p
Weight at D110 (kg)	W110	2244	224 (19.8)	150-289	0.28±0.05	16.9
Sow weight at D110 (kg)	SW110	2182	208 (18.8)	135-269	0.27±0.05	16.2
Fat depth at D110 (mm)	FAT110	2225	19.3 (3.98)	7.5-35.5	0.37±0.05	3.47
Sow weaning weight (kg)	SWW	1963	197 (18.0)	129-265	0.35±0.06	16.7
Fat depth at weaning (mm)	SFW	1867	17.4 (3.48)	6.5-32.0	0.34±0.05	3.35
Total born (N)	TB	2288	11.7 (3.18)	2-21	0.10±0.03	3.13
Average piglet birth weight (kg)	ABW	2223	1.41 (0.24)	0.63-2.42	0.27±0.03	0.23
Lactation feed intake (kg/day)	LFI	2034	4.99 (1.10)	0.50-9.00	0.16±0.04	0.78
Litter gain to 10 days (kg)	LG10	1970	10.0 (6.10)	-12.4 to 32.5	0.09±0.04	5.91

Genetic parameters. Heritability estimates from univariate analyses for sow weight and fat depths prior to farrowing and at weaning were moderate (range: 0.28 to 0.35, Table 1). The

estimate of heritability for LFI (h^2: 0.16±0.04) was similar to that previously reported by Bunter *et al*. (2007). For comparison, heritabilities for total lactation feed intake reported by Bergmsa *et al*. (2008) for sows recorded using different lactation feeding regimes over parities were 0.14±0.05 and 0.30±0.08, whereas the heritability for litter gain until 28 days was 0.18±0.05, not significantly higher than in this study. Genetic parameters for TB and ABW were consistent with averages from numerous studies reported by Rothschild and Bidanel (1998). Estimates of heritabilities from bivariate analyses (not presented) were similar to those estimated from univariate analyses.

Genetic correlations between sow body weights or fat depths before farrowing (W110 or SW110 and FAT110) and at weaning (SWW and SFW) were high but significantly less than one (Table 2). Phenotypic correlations for the same traits were much lower (~0.65) suggesting considerable variation amongst sows for changes in body weight and fatness as lactation progresses. Differences in weights pre- and post-farrowing include loss of conceptus products at farrowing, changes to sow body weight and composition during lactation, along with variation in mammary tissue development between these time points. Genetic and phenotypic correlations between weight and fat depths were moderate (range: 0.30 to 0.52) regardless of physiological state (pregnant or farrowed), and were similar to comparable estimates for weight and fat mass reported by Bergsma *et al*. (2008).

Table 2. Genetic (upper) and phenotypic (lower triangle) correlations (±se) (all values ×100)

	W110	SW110	FAT110	SWW	SFW	TB	ABW	LFI	LG10
W110	-	**98±0.6**	**29±11**	**79±5**	**30±12**	21±17	**32±12**	**34±16**	-1±18
SW110	**97±0.1**	-	**30±11**	**75±6**	**52±12**	3±18	16±13	29±17	-4±17
FAT110	**36±2**	**37±2**	-	**27±11**	**90±4**	-5±16	6±12	-12±15	-9±16
SWW	**65±1**	**64±1**	**24±2**	-	**43±10**	26±17	**-31±13**	**52±13**	**-39±15**
SFW	**30±2**	**41±2**	**66±1**	**46±2**	-	17±17	-12±13	21±16	-21±18
TB	**20±2**	-5±4	**-8±2**	3±2	0±2	-	-7±18	18±23	-24±24
ABW	**12±2**	**11±2**	2±2	**-8±2**	**-10±2**	**-46±2**	-	21±17	33±19
LFI	**-7±2**	**-10±2**	**-12±2**	**38±2**	**15±3**	**8±2**	**-4±2**	-	10±24
LG10	3±2	3±2	-1±2	**-18±2**	**-19±2**	-4±3	**14±2**	**20±2**	-

See Table 1 for trait abbreviations. Correlations significantly (P<0.05) different to zero are in bold.

Genetic correlations between TB and sow weight or fat traits were not significantly different from zero. Genes controlling ovulation rate and embryo survival, which determine TB, are largely independent of genes associated with body composition of the sow, as expected. The positive phenotypic correlation between W110 and TB (r_p: 0.20±0.02) arose from a part-whole relationship, since the phenotypic correlation was not different to zero between SW110 and TB (r_p: -0.05±0.04). Primiparous sows gestating larger litters were leaner (r_p: -0.08±0.02) prior to farrowing which suggests that sows have partially supported piglet development at the expense of accumulating their own body reserves during gestation. As has been observed from many studies, the phenotypic correlation between TB and ABW was strongly negative (r_p: -0.55±0.03).

Moderate negative genetic correlations between ABW or LG10 and SWW (r_a: -0.39±0.15 and -0.31±0.13) show that sows with heavier piglets at birth or with higher litter gains to day 10 were lighter and leaner at weaning, likely due to increased litter demands on sow resources. Additional estimates of correlations between ABW and sow weight or fat loss were strongly positive (r_a: 0.69±0.12 and 0.33±0.16; r_p: 0.20±0.02 and 0.15±0.02). Grandinson *et al.* (2005) and Bergsma *et al*. (2008) generally had similar results. Litter weight gains to day 10 of lactation were uncorrelated with sow weight or fatness prior to farrowing. Additional estimates of genetic and

phenotypic correlations between ABW and piglet weight at 10 days were 0.67±0.10 and 0.41±0.02. Piglet birth weight explains a large part of the genetic variability in piglet weight at day 10.

Genetic and phenotypic correlations between LFI and SWW, SFW or LG10 indicate that sows with higher lactation feed intake achieve higher body weight and condition at weaning (r_a: 0.52±0.13 and 0.21±0.16; r_p: 0.38±0.02 and 0.15±0.03) and higher litter gain, although only at the phenotypic level for the latter (r_a: 0.10±0.24, r_p: 0.20±0.02). While genetic correlations between LFI and ABW or TB were favourable, they were not significantly different from zero. Selection for larger surviving litters would therefore not be expected to generate a significant correlated response in LFI, with possibly detrimental effects for sow condition at weaning.

CONCLUSIONS

Genes controlling sow body weight and fatness at the end of lactation are either not identical or act differently to those controlling the same traits prior to farrowing. Sows with high genetic potential for farrowing and rearing heavier piglets are at risk of lower weight and fat depth at weaning. Gestating litter size had a negative phenotypic association with weight or fat gain of primiparous sows prior to farrowing. These effects can have negative consequences for sow longevity. Phenotypic correlations suggest that sows partially adjusted LFI according to their own body condition at farrowing and to the demands of the suckled litter. However, the absence of a substantial genetic correlation between TB and LFI indicates that a correlated response in LFI to selection on TB will not occur, potentially exacerbating the deficit between feed intake potential and requirements during lactation. Further research into achieving the best farrowing outcomes and treatment in the first parity is implicated for improving sow longevity.

ACKNOWLEDGMENTS

This research was funded by the Pork CRC under projects 2D-101-0506 and 2E-104-0506. The authors are grateful for diligent data collection by Matthew Tull and staff at QAF Meat Industries.

REFERENCES

Bergsma, R., Kanis, E., Verstegen, M.W.A. and Knol, E.F. (2008) J. Anim. Sci. **86**:067.

Bunter, K.L., Luxford, B.G. and Hermesch, S. (2007) *Proc. Assoc. Advmt. Anim. Breed. Genet.* **17**: 57.

Grandinson, K., Rydhmer, L., Strandberg, E. and Solanes, F.X. (2005). Anim. Sci. **80**:33.

Knol, E.F. (2001) PhD Thesis, Wageningen University.

Rothschild, M.F. and Bidanel, J.P. (1998) In "The Genetics of the Pig", p. 313, editors M.F. Rothschild and A. Ruvinsky, CAB International.

Gilmour, A.R. Cullis, B.R., Welham, S.J. and Thompson, R. (2006) "ASReml User Guide", Release 2.0 VSN International Ltd, Hemel Hempstead, HP1 1 ES, UK.

SAS (2003) Version 9. SAS Institute Inc. Cary, N.C..

Tholen, E. Bunter, K.L. Hermesch, S. and Graser, H.-U. (1996) Aust. J. Agric. Res. **47**:1261.

Proc. Assoc. Advmt. Anim. Breed. Genet. 18:207-211

BREEDING OBJECTIVES FOR SEASONAL PRODUCTION SYSTEMS: AN EXAMPLE FROM NEW ZEALAND VENISON SYSTEMS

J.A. Archer[1] and P.R. Amer[2]

[1]AgResearch Ltd, Invermay Agricultural Centre, Mosgiel, New Zealand
[2]AbacusBio, PO Box 5585, Dunedin, New Zealand

SUMMARY

Seasonal variation in prices received and input costs can have market effects on breeding objectives and farm system strategies. A breeding objective was developed for venison production systems in New Zealand, accounting for seasonal prices and feed costs. Two farm systems were compared, one with an early kill profile targeting premium prices in spring, and the other with a later kill profile which doesn't achieve spring price premiums. The impact of spring premiums on the value of additional growth was marked, but is somewhat negated by the additional feed costs incurred under an early kill system. The principles demonstrated in this example have application to other pastoral production systems.

INTRODUCTION

The value of genetic improvement in different traits can be strongly influenced by seasonal fluctuations in prices received and input costs such as feed under pastoral systems. In such systems, the commercial value of genetic improvement in different traits can vary depending on whether the farm management system targets seasonal premiums or not, and what the current success rate in achieving these premiums is.

The New Zealand venison industry has many similarities to other pastoral based meat production systems (e.g. lamb and beef), but a distinguishing feature of venison production systems is the very marked impact of seasonality. The venison industry operates under seasonal feed supply patterns (no different to many other pastoral systems), but also has strongly seasonal market demand and strong biological seasonal control over reproduction, feed intake and growth of deer. The biological and feed supply seasonality factors constrain supply of venison into periods of peak market demand.

The major markets for New Zealand venison are European, where a strong preference exists to consume venison during the northern hemisphere autumn, based on centuries of culinary tradition and hunting during "the rut". This means that chilled venison produced from July to November attracts premium prices. Venison produced during the remainder of the year is mainly exported as frozen product, generally attracting a lower price, and stored for consumption during the following European autumn. Despite attempts to market venison to European consumers "out of season", the fundamental consumer tradition and preference for venison consumption during autumn remains strong.

The schedule paid for venison carcasses to farmers reflects the seasonal demand for venison and the difficulty in producing venison of the desired carcase specification (55 to 70 kg carcase) from an animal less than 12 months old. Moreover, as the schedule declines during New Zealand's spring the value of a constant weight carcase declines at a rate which, depending on the absolute value of the schedule, often means that the additional carcase weight obtained by retaining growing animals for an additional week is not sufficient to compensate for the declining schedule leading to a loss of value. Thus venison production during this period is often driven by selling animals as soon as they reach a weight which will produce a carcase within specifications (around 55 kg carcase weight), and consequently the average weight of venison carcasses has not changed in many years, although the timing of the kill or the age of animals killed (rising one-year

olds vs animals retained for an extra year) may have changed significantly.

The consequences of seasonal schedules on the economic value assigned to traits can be marked, and these are dependent upon the current level of performance of the farm system being considered. In this paper we constructed a breeding objective for New Zealand venison systems operating under a seasonal schedule, and compare the economic value for different traits under scenarios where farm strategy and performance differs. While the example is specific to venison production, the principles can be applied to other livestock production systems where seasonal payment schedules are a feature.

DEVELOPMENT OF THE BREEDING OBJECTIVE

A bio-economic model for a self-replacing deer herd focussed on venison production only was created using an excel spreadsheet. The production system was based on slaughter of animals aged from 9 months to 21 months of age (ie. yearling production system), typical of most venison production in New Zealand. Two different scenarios were modelled, based on variation in average slaughter date. In an early kill system, median slaughter date for stags was set to 6th October, with median slaughter date for hinds being 6th December. This farm system targets the period in which spring premium prices are available, and while a proportion of animals are killed during the premium period, genetic gain in traits which lead to heavier animals at a fixed time of year will result in a larger proportion of animals achieving spring premiums. A later kill system was also modelled, with median slaughter dates being 6th December and 6th February for stags and hinds respectively. In this system most animals are killed after the spring premium period has ended, and incremental gain in traits leading to earlier slaughter is not sufficient to make a large difference to the proportion of animals achieving premium prices. However, as with the early kill system some savings in feed costs are potentially available from the reduction in maintenance requirements arising from earlier slaughter. Figure 1 shows the assumptions for seasonal variation in feed costs and schedule price.

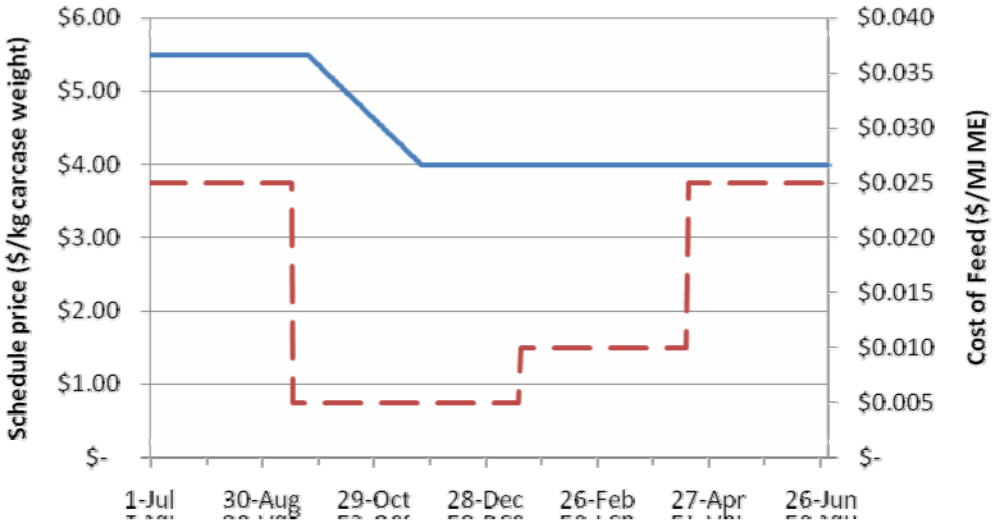


Figure 1. Assumptions used to describe seasonal changes in schedule price (solid line) and cost of feed (broken line).

The bio-economic model was constructed to reflect a system where animals are drafted weekly from 1st September, with all animals greater than 97 kg liveweight sold for slaughter. Feed costs for different classes of animals were calculated using assumptions for maintenance (0.7 for stags and 0.6 for hinds MJ ME per $kg^{0.75}$) and growth (37 MJ ME/kg liveweight gain). Feed costs were broken down into costs from birth to 3 months, 3 months to 6 months and 6 months to 12 months. Changes in weight at each age independent of the other ages were modelled to calculate an economic value, this approach leading to negative weights at earlier ages reflecting the slightly reduced maintenance requirements resulting from a growth path that is slow early on followed by rapid gains to reach slaughter weight. However, the economic weight for weight at 12 months is positive, reflecting the fact that a heavier weight at this age leads to earlier slaughter and consequently savings in maintenance requirements. Increases in hind mature weight resulted in increased annual maintenance requirements, increased feed costs to rear replacement hinds to a heavier weight, and increased cull value for hinds surviving to slaughter.

There are two potential pathways by which animals can be slaughtered earlier (but at a constant liveweight), namely improvement in growth rates or earlier calving. Reproduction in red deer is under strong seasonal control driven by photoperiod, but some genetic variation in conception date appears to exist both between breeds (Scott *et al.* 2006) and within breeds (Archer, unpublished data). The impact of seasonal changes in schedule price was calculated by deterministically simulating the drafting of animals at a fixed liveweight (97 kg), starting on 1st September. Economic values for carcase weight (at a constant age) were calculated based on the change in average price per kg resulting from earlier slaughter. During the finishing period, animals were assumed to grow at 0.3 and 0.25 kg/day for stags and hinds respectively. Standard deviations for liveweight at 12 months were 15 and 12 kg respectively. Economic values for calving date were calculated by shifting the mean calving date forward, and then calculating the change in average value per calf slaughtered and multiplying by the number of calves slaughtered per hind mated (0.59).

Carcass composition and reproductive traits were included in the index for future purposes, although little measurement of these traits currently occurs. Carcass composition was broken down into loin cuts and hind quarter cuts worth 4 times and 1.6 times the value of fore-quarter cuts respectively. These values were calculated at a constant slaughter weight, so that an increase in weight of one area displaces weight in other cuts. Economic values of reproductive success (pregnancy and calf survival) were calculated as the value of additional calves slaughtered minus feed costs.

The number of discounted genetic expressions for each trait type was calculated to account for differences in the frequency and timing of expressions (based on Amer 1999). A planning horizon of 25 years was taken with gene flows modelled through generations in the breeding herd. A discount rate of 0.07 was assumed.

RESULTS AND DISCUSSION

The economic values calculated for traits in the index under the early kill and late kill systems are given in Table 1. The early kill system differs from the late kill system by having higher feed costs per MJ of ME supplied, a greater opportunity to exploit early season premiums in a significant proportion of calves slaughtered, and slightly higher hind mature weights. The economic values for liveweight traits in the late kill system reflect the cheaper cost of feed, with a lower penalty on mature weight and a greater economic value on 12-month liveweight. The benefit from achieving greater spring premiums in the early kill system is seen in the economic weights for carcase weight and calving date. Under the scenario modelled higher BVs for carcase weight translate into earlier slaughter (due to killing at a fixed liveweight), while earlier calving (ie. negative BV for calving date) leads to heavier animals at a given time of year and hence earlier

slaughter at a fixed liveweight. Consequently, the index for the early kill system places considerably more emphasis on these two traits compared to the index for the late kill system. However, if carcase weight is considered to be directly related to 12 month weight by a factor of 0.55 (to account for dressing percent), then the total weighting on 12-month weight is 0.9015 for early kill and 0.8105 for late kill. Thus while additional growth is still more important in early kill systems, the advantage is somewhat diminished by the higher average cost of feed in this system. Economic weights for carcase composition traits are similar between the two systems, while traits describing reproductive success receive greater emphasis in the late kill system.

Table 1. Economic values for indices formulated for early and late kill systems.

Goal trait	Early Kill System			Late Kill System		
	EV[1]	DGE	EW	EV	DGE	EW
Growth rate						
Weight – 3 months kg	0.06	0.85	0.05	-0.33	0.85	-0.28
Weight – 6 months kg	-0.30	0.85	-0.25	-0.08	0.85	-0.06
Weight – 12 months kg	0.07	0.77	0.06	0.40	0.77	0.31
Hind mature weight kg			-0.45			-0.37
Replacement kg hind feed kg	-0.26	0.24		-0.22	0.24	
Annual hind feed kg	-0.69	0.95		-0.61	0.95	
Cull hind value	1.65	0.16		1.65	0.16	
Carcass Weight kg	2.81	0.54	1.53	1.67	0.54	0.91
Carcass yield (age constant BVs)						
Loin cuts kg	11.40	0.54	6.15	11.85	0.54	6.40
Hindquarter cuts kg	1.66	0.54	0.89	1.72	0.54	0.93
Forequarter cuts kg	-0.45	0.54	-0.24	-0.47	0.54	-0.25
Maternal						
Scanned pregnant – 2 yr old	113.54	0.24	27.14	129.04	0.24	30.84
Scanned pregnant – mixed age	113.54	0.76	87.69	129.04	0.76	98.65
Calf survival – 2 yr old	113.54	0.19	21.55	129.04	0.19	24.50
Calf survival – mixed age	113.54	0.76	87.69	129.04	0.76	98.65
Calving Date	-0.40	0.95	-0.39	-0.17	0.95	-0.16

[1]EV = Economic value ($) DGE = Discounted Genetic Expressions EW = Economic weight

These results highlight the impact of both seasonal premiums and seasonal feed costs on the composition of the index, and show that current farm system performance has a significant influence on the index and the farming strategy taken. Where a moderate change in genetic performance has a significant impact on the number of animals killed on premium schedules, it is worth putting significant emphasis on traits which will assist in targeting earlier kill. However, in situations where a large improvement in performance is required to meet premium schedules with earlier kill, a better strategy is to concentrate on reproductive performance and carcass yield, and to utilise the benefits of a lower cost of feed. Deer farmers make these decisions largely intuitively, but modelling approaches such as the one described here are useful to more objectively describe the trade-offs and determine the optimal direction to take in breeding programmes.

In practice the strategies used by many New Zealand deer farmers reflect this finding. With changing land use, the breeding herd has largely shifted to hill and high country where the cost of feed is lower, and the environment is conducive to better reproductive performance as calves tend to survive better in extensive areas. One of the industry issues currently being debated is the impact of high growth genetics on reproductive performance, with concerns that larger hinds will

lead to poorer reproductive performance. The evidence to support this view is anecdotal rather than experimental, but nevertheless it is consistent with experiences in other pastoral meat production systems.

To our knowledge there are few reports in the literature where seasonal premiums are explicitly accounted for when valuing genetic gain and developing breeding objectives. Jones et al (2004) used a model of fat and lean growth in lambs together with a seasonal change in schedule prices to model a drafting strategy and calculate economic values for component traits. Their study also valued differences in feed costs throughout the year, and they suggested that this was particularly important when concentrate supplements were used along with grazing to feed animals. Certainly the issues of seasonal feed prices are common across most livestock industries, and seasonal premiums often reflect the impact of seasonal feed prices on supply. The venison example described here has additional seasonal drivers as schedule prices are a result of seasonal consumer demand and biological seasonality of deer in addition to seasonal feed prices, but the principles have application in other livestock systems.

ACKNOWLEDGEMENTS

This work was funded by DEEResearch Ltd.

REFERENCES

Amer, P.P. (1999). *N. Z. J. Agric. Res.* **42**:325.

Jones, H.E., Amer, P.R., Lewis, R.M. and Emmans, G.C. (2004) *Livest. Prod. Sci.* **89:**1.

Scott, I.C., Asher, G.W., Lach, J.E. and Littlejohn, R.P. (2006) *Proc. New Zealand Soc. Anim. Prod.* **66:**270.

TRANS-TASMAN GENETIC EVALUATIONS OF SHEEP

M.J. Young[1], S-A. N. Newman[2], R. Apps[3], A.J. Ball[3] and D.J. Brown[4]

[1]SIL, Meat & Wool New Zealand, PO Box 39-085, Christchurch 8545, New Zealand
[2]AgResearch, Invermay Agricultural Centre, Private Bag 50034, Mosgiel 9053, New Zealand
[3]Meat & Livestock Australia, PO Box U254, Armidale, NSW 2351, Australia
[4]Animal Genetics and Breeding Unit*, University of New England, Armidale, NSW 2351

SUMMARY

Until recently it was not possible to compare the genetic merit of sheep between Australia and New Zealand in a technically valid manner. However some breeds now have sufficient sires with progeny in both countries to allow robust genetic evaluations based on performance and pedigree information from both countries. Corriedale and Coopworth breeders in the two countries have conducted across-flock, across-country evaluations using the Sheep Improvement Ltd (SIL) and Sheep Genetics (SG) genetic evaluation systems. Compared to within country evaluations, breeding values (BVs) from the across-flock evaluation were very similar (r=0.96-1.00) when the same system (SIL or SG) was used to estimate BVs. BVs generated from the combined data set by the two systems had lower but still strong correlations(r=0.69-0.89) for most traits. This variation was attributed to different analysis models, including genetic parameters, and performance being measured at somewhat different ages. Ultrasound scan traits were least well correlated (r=0.14-0.65), most likely due to SIL producing BVs adjusted to constant age while SG produces BVs adjusted to constant carcass weight. It was concluded that the best option for genetic evaluation of combined datasets from the two countries is to exchange and combine data but to conduct genetic evaluations and produce reports within-country.

INTRODUCTION

Genetic evaluation systems based on recorded performance measurements and pedigree, commonly use individual animal model BLUP (best linear unbiased prediction) methods for estimating genetic merit (*e.g.* Newman *et al*. 2000, Brown *et al*. 2007). Increasingly, breeders are using common sires to establish genetic connections on which to base across-flock genetic evaluations. These can extend to across-breed analyses *e.g.* LAMBPLAN Terminal Sire analysis (Brown *et al*. 2000; www.sheepgenetics.org.au/lambplan) or SIL-ACE (Young and Newman 2009; www.sil.co.nz) evaluations.

International genetic evaluations are well established for dairy and beef cattle (Schaeffer, 1994, Donoghue *et al*. 2007). To the author's knowledge, Dohne flocks in Australian and South African are the only sheep example of groups combining data for genetic evaluation from geographically separated countries. Many sheep breeders have sought improvements using overseas genetics through importation of "best bet" genetics from their current breed or "new" breeds to bring new or novel genetics into the industry. The true genetic merit of overseas animals relative to local animals is unknown prior to importation and takes time to estimate once used.

Local, specialized breeds of sheep occur in most countries but some breeds are common to a number of countries. Typically, such breeds have common selection objectives, albeit modified in response to local market signals. Modern methods used for across-flock genetic evaluation can lead to selection of highly related individuals and so to inbreeding in seedstock populations.

* AGBU is a joint venture of NSW Department of Primary Industries and University of New England

Populations of the same breed from other countries offer one option when seeking genotypes that may lift performance and manage inbreeding. However, fair comparison of animals requires a common genetic evaluation of similar traits and relies on adequate genetic connectedness between populations (also called "linkage" - Newman 2003; Huisman *et al.* 2006). Table 1 summarizes options for comparing sheep from different countries for genetic merit. Robustness of comparisons made increases as you move down the table.

Table 1. Options for objective comparison of genetic merit of sheep across-country

	Factors to consider				Best options		
Option	*Traits measured*	*Connectedness between countries*	*Genetic analyses*	*Report features*	*Dataset*	*Genetic analysis*	*Report formats*
A	Different	*Not critical*	*Not critical*	*Not critical*	Separate	Separate	Different
B	Similar	Poor	*Not critical*	*Not critical*	Separate	Separate	Customize for each country
C	Similar	Good	Differ	*Not critical*	Combined	Separate	Customize for each country
D	Similar	Good	Similar	Different indexes	Combined	Common	Customize for each country
E	Similar	Good	Similar	Similar indexes	Combined	Common	Same format

GENETIC EVALUATIONS

At the request of Corriedale and Coopworth breeders in New Zealand and Australia, SIL and SG exchanged pedigree and performance datasets in order to determine the best approach to assessing genetic merit for sheep of each breed across-country. Such Trans-Tasman (TT) genetic evaluations can be provided by SIL and SG at the request of breed groups in each country, where there is good genetic connectedness.

Working with representatives of the breeder groups in each country, SIL and SG set up protocols for data protection and customer service. It was agreed that results of "trans-Tasman evaluations" be made available to both groups at a similar time and that the data could not be used for any purpose other than these genetic evaluations without permission of the breeders.

Within breed, the dataset from each country and the combined dataset were analysed by both genetic evaluation systems (SIL and SG). Table 2 details size of the Australian and New Zealand datasets for key traits analysed.

RESULTS AND DISCUSSION

There were very high correlations between breeding values for the same traits generated from analysis of a within country dataset and those from analysis of the combined dataset (0.96-1.00) within each evaluation system (SIL and SG). So breeders within each country should feel confident that the addition of overseas data will have little impact on their evaluations.

Comparison of SIL BVs with SG BVs using the same NZ data (Table 3) show a high correlations for body weight traits (R^2=0.69 to 0.86) where data and models are reasonably consistent but weaker correlations (R^2=0.14 to 0.32) for carcass traits, where models fitted in the analyses and data collected differ more. Similar results were seen for correlations between BVs generated by the two systems for the combined dataset.

Four factors led to less than perfect correlations between BVs estimated by the two systems.

1. Genetic parameters used in evaluations by the two systems were similar but not identical,

2. Components of the genetic models fitted differ, including; fixed, maternal and permanent environment effects as well as adjustment for heterogeneous variance.
3. Weaker correlations seen for carcass traits occur because SIL adjusts these BVs to constant age whereas SG adjusts them to constant carcass weight,
4. Some measurements are collected at significantly different ages, or at a different site for carcass trait scanning, in each country (Young *et al.* 1992; Gilmour *et al.* 1994),

The greatest effects were seen for carcass traits. While we cannot determine the extent to which this was due to the adjustment to different bases (age or carcass weight) or to differences in scan site or age, we believe it is primarily due to different bases for adjustment of carcass traits.

Table 2. Number of animals and data counts for key traits in Corriedale and Coopworth datasets from Australia and New Zealand (NZ)

		Corriedale		Coopworth	
	Trait	Australia	NZ	Australia	NZ
Total animals		46,875	128,014	74,223	826,661
Growth	WWT – Weaning weight	23,421	90,084	56,952	637,468
Growth	LW8 – Live weight at 8 months	10,984	23,384	30,270	378,163
Growth	LW12 – Live weight at 12 months	14,078	14,857	29,683	90,314
Meat	EMD – Eye muscle depth	7,513	5,312	21,930	27,735
Wool	GFW - Greasy fleece weight	13,908	43,280	33,016	342,101

Table 3. Regression of BVs for NZ Corriedales between SIL and SG analyses.

	All animals with records			Sires		
Trait	Animals	R^2	Coefficient	Animals	R^2	Coefficient
WWT	132,439	0.86	0.56	5,749	0.81	0.53
LW8	130,960	0.69	0.90	1,851	0.64	0.75
LW12	131,060	0.81	1.00	2,269	0.79	0.99
FAT	66,507	0.14	0.42	2,628	0.23	0.36
EMD	83,747	0.32	0.73	144	0.51	0.87
GFW	130,989	0.82	0.03	37,633	0.84	0.03

Connectedness between countries was adequate to produce across country breeding values for each breed. Using current SG connectedness analyses, 27 of the 28 Corriedale flocks with recent data were sufficiently linked to report across flock breeding values. Twelve sires had recorded progeny in both countries and in excess of 100 sires had across country pedigrees. Approximately 10% of Australian born animals were from NZ sires where as only 1% of NZ born animals were from Australian sires. All active Australian flocks had direct sire connectedness with NZ flocks. Similar levels of connectedness were seen in the Coopworth dataset.

FUTURE INTERNATIONAL EVALUATIONS

Corriedale breeders involved in this evaluation have agreed to transfer updated datasets for trans-Tasman evaluations three times per year (January, May and October) to fit with data collection, selection and marketing decisions. In addition to routine reporting, top sires and young rams from both countries will also be reported to provide local breeders with the opportunity to identify new genetics to consider for importation. Coopworth breeders are considering how trans-Tasman evaluations can be best used for their breed.

Health concerns mean some countries will not accept live animals, semen, or embryos, from some other countries due to real and perceived disease risks. If such restrictions lie in only one direction, strong connections can be built by flocks in the less restrictive country using genetics from the other country. A downside is that genetics only flow in one direction so benefits to breeders in the country not allowing gene importation come only from sale of genes, as semen, embryos or live animals. If two countries do not allow importation of each others germplasm in any form, genetic connections cannot be created precluding across-country genetic evaluation.

Breeders in Australia and New Zealand can exchange sheep genetics to capture both genetic improvement and marketing opportunities. This should allow them greater flexibility when seeking high rates of genetic improvement while minimising inbreeding. Research and development will continue to enhance BVs effectiveness for across-country selection.

CONCLUSIONS

Across country genetic evaluation for sheep breeders in New Zealand and Australia is currently best achieved using a combined dataset with analysis and reporting from the local system breeders are familiar with. Differences in trait definition and analysis specification will cause minimal re-ranking of sires for non-carcass traits, between results of the two evaluation systems.

Breeding groups wishing to conduct across-country genetic evaluation should actively develop and maintain strong genetic links.

ACKNOWLEDGEMENTS

Representatives of the Corriedale Sheep Society of New Zealand and the Australian Corriedale Association for initiating the first trans-Tasman genetic evaluation. Meat & Wool New Zealand and Meat & Livestock Australia for their support of this work. Referees for valuable criticism.

REFERENCES

Brown, D.J., Huisman, A.E., Swan, A.A., Graser, H-U., Woolaston, R.R., Ball, A.J., Atkins, K.D. and Banks, R.B. (2007) *Proc. Assoc. Advmt. Anim. Breed. Genet.* **17:**187

Brown, D.J., Tier, B., Reverter, A., Banks, R. and Graser, H.U. (2000) *Wool Tech. & Sheep Breed.* **48**:285

Donoghue, K. A., Graser, H.-U., Johnston, D. J. and Tier, B. (2007) *Proc. Interbull Meeting, Aug 2007, Dublin, Ireland*. Bulletin 37 at http://www.icar.org/pages/Sub_Committees/sc_interbull.htm

Gilmour, A.R., Luff, A.F., Fogarty, N.M. and Banks, R. (1994) *Aust. J. Agric. Res.* **45:**1281

Huisman, A.E., Tier, B. and Brown D.J. (2006) *Liv.Sci.* **104**:254

Newman, S-A.N. (2003) *Proc. N.Z. Soc. Anim. Prod.* **63:**194

Newman, S.A., Dodds, K.G., Clarke, J.N., Garrick, D.J. and McEwan, J.C. (2000) *Proc. N.Z. Soc. Anim. Prod.* **60:**195

Schaeffer, L.R. (1994) *Jl. Dairy Sc.* **77:**2671

Young, M.J., Deaker, J.M. and Logan, C.M. (1992) *Proc. N.Z. Soc. Anim. Prod.* **52:**37

Young, M.J. and Newman, S-A.N. (2009) *Proc. N.Z. Soc. Anim. Prod.* **69:** 145

Breeding program design including MAS

COMBINING ESTIMATES OF SNP EFFECTS WHEN THEY ARE SUBPOPULATION SPECIFIC

P. R. Amer and G. M. Payne

AbacusBio Limited, Dunedin, New Zealand

SUMMARY

Three methods for combining subpopulation specific coefficients linking numerous genetic markers to phenotypic trait performance are compared using simulation. Resulting combined coefficients are used to predict the genetic merit of selection candidates. Sub populations varying in size, true QTL effects, and their degree of similarity in the extent of linkage disequilibrium between markers and quantitative trait loci to a test population of selection candidates are simulated. The methods of combining estimates that were considered differ in the way that weighting is placed on coefficients from the subpopulations. At one extreme, only coefficients from the sub population from which the target population is derived are used. At the other extreme, weighting is placed jointly on the standard errors of estimates of coefficients, as well as the similarity of coefficients between the subpopulation from which the test population was derived, and other subpopulations. An example is shown whereby the weighted by correlation method outperforms the other methods. The role of population specific coefficients in motivating investment in genotyping and trait recording at the subpopulation level is discussed in the context of these results.

INTRODUCTION

A major revolution in dairy cattle breeding programmes is currently underway as predictions of genetic merit of young bulls can now be made at moderate accuracy (e.g. Spelman *et al.* 2007), with huge scope for circumvention of progeny testing (Schaeffer 2006). There is considerable investment (at the level of industry and commercial service company level) currently under way with a view to this technology being applied in other industries, including sheep. Application in these industries may be more problematic for several reasons including: less genetically uniform populations with many breeds and subpopulation structures; less opportunity to shorten generation intervals; an absence of recording of profitable traits linked to genetic variation (e.g. maternal traits in adult breeding animals); and a lower commercial value of individual elite breeding animals (Dodds *et al.* 2007).

Because of current cost versus value trade-offs in sheep, it seems likely that a relatively modest subset of the total number of SNPs used in a discovery phase would be incorporated into a SNP key for whole genome selection (WGS). A likely application is that selection candidates that are genotyped for the SNP key can then have their genetic merit predicted directly from their genotype results, provided that robust and validated coefficients are available which translate genotype results into effects on traits. These SNP based predictions of genetic merit can then be integrated with phenotype based predictions of genetic merit (Estimated Breeding Values, EBV) generated from conventional genetic evaluation systems. Other potential applications can also be considered; for example, in theory both molecular and phenotypic based predictions of genetic merit could be computed in a simultaneous and co-ordinated way.

One key issue that must be addressed is the predictive performance of the SNP key in populations other than those in which the SNP/trait associations were discovered. Similarly, there may be considerable variability within subsets of the discovery population in the predictive performance of the SNPs. In these situations, the SNP key's ability to predict breeding values is related to the degree of linkage disequilibrium between SNPs and trait QTLs. Given the low SNP

density of the 60K SNP chip, the physical distance between a given SNP and QTL is relatively high. This translates to a high chance that phase associations between them can break down across populations or even among individuals within populations. This paper addresses the issue of how best to incorporate available SNP information from multiple subpopulations/breeds to increase accuracy of molecular breeding value prediction in target populations, where direct SNP effects are unknown.

MATERIALS AND METHODS

Context. We assume that a subset of SNP markers with reasonably robust links to quantitative trait performance in at least a subset of subpopulations has been identified following a research discovery phase of technology development. This so-called SNP key then becomes validated on a range of new animals in each subpopulation. The validation process involves genotyping sires with estimated breeding values derived from significant numbers of progeny records. An output of the validation process is a set of subpopulation specific coefficients that are capable of predicting the phenotype of an animal from that subpopulation based on SNP test results. The predictive ability of the coefficients is assumed to be independent of the predictive ability that arises when SNP frequencies are confounded with population substructure.

Simulation. Three simulations were run, one each for 50, 100 and 200 SNPs on the SNP key. Each SNP was associated with a single QTL, with the coefficient defining linkage disequilibrium (LD) between SNP/QTL pairs (expressed as deviation of observed haplotype frequency from expected) randomly drawn from a uniform distribution with 0.4 and 0.7 as the minimum and maximum values. The minor allele frequencies were randomly drawn from a uniform distribution from 0.3 to 0.5 and 0.1 to 0.5 for SNPs and QTLs respectively. When simulated levels of LD were incompatible with QTL and SNP frequencies, the LD was adjusted down to the maximum feasible value. QTL effects at each locus were simulated as a base effect with a random additional increment drawn from a beta distribution with parameters of 1.5 and 5. QTLs were assumed to be completely independent, and SNP's were linked only to their corresponding QTL. True breeding values were simulated for sires as the sum of simulated QTL effects plus a normally distributed polygenic component that is independent of any SNPs. It was assumed that 60% of total genetic variance was explained by the simulated QTLs linked to SNPs; irrespective of how many SNPs were being simulated. Estimated breeding values for sires were also simulated by adding a normally distributed error term onto the true breeding value and rescaling the result to account for the trait phenotypic variance and assuming that the accuracies of the estimated breeding values were 0.8.

Subpopulations. Four reference populations were derived from the founder population. Each reference population deviated from the founder population in regards to the level of LD between SNP and QTL pairings and QTL effects. Parameters used to create differences between each reference population and the founder population are described in Table 1. New QTL effects were simulated for each reference population so that the correlations between effects in the founder population with effects in the reference population were as specified in Table 1. In populations two and three, the level of LD simulated in the founder population for each SNP – QTL pairing was scaled by multiplication by a random deviate sampled from a triangular distribution with a mode of 1 and lower and upper bounds as specified in Table 1. SNP coefficients were generated for each SNP in each reference population by regressing SNP genotypes on de-regressed, sire estimated breeding values. Associated standard errors of SNP coefficients were also computed.

A test population consisting of 1000 animals was derived with identical parameters to reference population one. SNP based predictions of merit were then created using SNP genotypes and SNP coefficients derived by three different methods and their correlation with simulated true breeding values for genetic merit computed. In method one, coefficients from reference population one were used ignoring completely the coefficients from the other populations. In method two, a weighted average of coefficients from all populations was used, whereby the weightings were taken as the reciprocal of the squared standard error of the regression coefficient. In method three, the weightings used in method two were further updated to account for the similarity of coefficient estimates among the reference populations. To do this, correlations (r) among coefficient estimates from the 4 subpopulations were computed and incorporated into the method 2 weighting factors as follows:

$w_{i,j} = \frac{se_{i,j}^{-2}}{\sum_j se_{i,j}^{-2}} \cdot \frac{r_j}{\sum_j r_j}$ giving method 3 coefficients $\tilde{b}_i = \sum_j \frac{w_{i,j} b_{i,j}}{\sum_j w_{i,j}}$ where method 3 weighting factors (w) and standard errors (se) correspond to SNP regression coefficients (b) for locus i estimated from subpopulation j. Correlation estimates for coefficients between pairs of subpopulations incorporated their own weighting factors based on the inverse of squared coefficient standard errors. In order to determine the weighting applied to reference population one coefficients, relative to other populations (i.e. to estimate r_1), it was necessary to simulate a further validation set of animals from population one, and determine the correlation between coefficients from the two equally sized sets of animals from population one. Weightings were therefore higher for coefficients from subpopulations with a high correlation with population one's coefficients and higher for coefficients with a low standard error.

Table 1. Parameters used to define differences between 4 reference populations and a founder population.

Parameter	Reference population			
	1	2	3	4
Number of sires	500	1000	500	500
Lower bound for LD scaler	1.0	0	0	1.0
Upper bound for LD scaler	1.0	1.2	1.2	1.0
Average SNP-QTL LD	0.5	0.36	0.36	0.5
QTL effects correlation	1.0	0.2	0.5	1.0

RESULTS AND DISCUSSION

Average correlations (from 100 replicates) and their standard errors between true breeding values and SNP based estimates of true breeding values are presented in Table 2 for simulations with 50, 100 and 200 SNPs considered respectively. Because the proportion of genetic variance explained by QTLs linked to SNPs remained the same, irrespective of the number of SNPs simulated, individual QTLs had decreasing individual effects as the number of SNPs increased. With larger QTL effects, SNP coefficients can be estimated more accurately, and hence correlations are higher irrespective of methods used.

For the situation simulated, method 3 (incorporating standard errors and the general prediction of similarity between populations) performed better than the other methods. Method one performs poorly in many situations (not shown); an example is when the information from other

subpopulations is useful. Method two performs poorly when none of the sub populations have coefficients that have similar LD and similar QTL effects to the target population.

Further development of method three is required. In particular, its implementation for this simulation required an additional validation population for the reference subpopulation that matches the target population. In the real world, this would be a waste of resources, as it would be better to use the test population data to improve the coefficients for population one. A system of splitting data from subpopulation one into two sub populations of varying size and then extrapolating the correlations between the two subpopulations out to what they might be with the same number of animals as in population one is one approach that could be considered and tested. With future knowledge, it might also be possible to integrate other information about subpopulation similarity into the weighting factors for combining coefficients. Breed knowledge, and genetic distance information based on phylogeny type analysis are potential information sources that could be used. Such an approach would benefit from reformulating the problem in the context of Bayesian decision theory.

Table 2. Accuracy of prediction of true genetic merit using 3 methods of aggregating SNP coefficients from subpopulation averaged over 100 replicates (standard errors in brackets)

SNP coefficient method	50 SNPs		100 SNPs		200 SNPs	
1. Population 1 only	0.287	*(0.004)*	0.235	*(0.003)*	0.182	*(0.003)*
2. Weighted average	0.295	*(0.004)*	0.256	*(0.004)*	0.212	*(0.003)*
3. Weighted by correlation	0.325	*(0.004)*	0.276	*(0.004)*	0.220	*(0.003)*

The concept of customised coefficients might fit well in an industry context. It provides an incentive for breeders to invest in validation phenotyping and genotyping, while at the same time, creates a situation where there are modest mutual benefits from sharing data, without significant loss of intellectual property because coefficients are most relevant to the subpopulation they are estimated in.

CONCLUSIONS

Development of population specific SNP coefficients may provide superior predictors of genetic merit for most animals and could lead to mechanisms for incentivising private investment by breeding companies and farming organizations to undertake SNP genotyping in animals recorded for novel phenotypes and for historically recorded animals. This work has identified a method which, after further refinement, could provide an opportunity for custom derivation and application of SNP prediction coefficients that use information from multiple subpopulations in a robust way.

ACKNOWLEDGMENTS

We gratefully acknowledge funding for this work by Ovita Limited.

REFERENCES

Dodds, K.G., Amer, P.R., Spelman, R.J., Archer, J.A. and Auvray B. (2007) *Proceedings of the New Zealand Society of Animal Production* **67**:162.

Schaeffer, L. R. (2006) *Journal of Animal Breeding and Genetics* **123**:218

Spelman, R. J., Arias, J., Keehan, M., Obolonkin, V., Winkelman, A., Johnson, D. and Harris, B. (2007) *Proc. Assoc. Advmt. Anim. Breed. Genet.* **17**:471.

SNP PREDICTORS TO ACCELERATE THE RATE OF GENETIC PROGRESS IN SHEEP

J.A. Sise and P.R. Amer

AbacusBio Limited, Public Trust Building, 442 Moray Place, Dunedin, New Zealand

SUMMARY

A selection index model was used to calculate the benefit of using SNP panels in a ram breeding program for 5 different target traits including growth, carcass yield, reproduction, survival and parasite resistance. Breeding programs using either a mixture of young, emerging and mature rams (based on industry average ram ages) or the use of predominantly young untested rams were considered. Using an index with modest emphasis on litter size, response to selection without markers was higher (111 index cents per annum) in the young ram program than that observed with an industry average program (106 index cents per annum). Whilst more information on mature rams leads to better selection decisions, the reduced generation interval achieved through use of young rams leads to a higher overall response. When SNP markers were included as selection criteria with a prediction accuracy of 0.5, the increase in overall breeding program response ranged from 2-10 cents of index per animal per year, with the largest gains observed in markers for growth and carcass yield. Proportionately greater increments of response are expected as the accuracy of the genomic scores increases beyond 0.5.

INTRODUCTION

In breeding programs, an index of economic merit is commonly used to assist in selection decisions. Recorded trait and pedigree information is used to estimate breeding values, which are combined into an index according to their relative economic weightings. Over the last 5 years, much research has gone into the development of high density DNA marker chips, containing hundreds or even thousands of single nucleotide polymorphisms (SNPs), which are linked to traits of economic importance. Associations of large numbers of markers spread across the genome with phenotypes linked to economic traits, has allowed development of marker assisted selection methods, used to predict breeding values for individuals using their combined marker genotypes (Meuwissen *et.al* .2001).

This paper uses a selection index model to predict the response to genomic selection in dual purpose sheep breeding flocks. Parameters and assumptions are based around selection indexes widely used within the New Zealand sheep industry (http://www.sil.co.nz/). Ages of candidates at selection were based on Amer (2009). Selection index methods can be used to determine the potential benefits of recorded trait and DNA marker information in selection decisions for young animals (Dekkers 2007). Results are presented as the breeding program response with and without genomic markers.

MATERIALS and METHODS

A selection index model was used to predict the response to conventional and marker assisted selection in dual purpose sheep breeding flocks. The total merit index was comprised of prolificacy, growth, carcase, wool, lamb survival and internal parasite resistance traits of direct economic performance. Selection criteria included, lamb survival observations, 2 live weights, recorded at weaning (WWT) and as the estimated weight at slaughter age (CW), a hogget fleece weight (HFW), ultrasonic fat depth (FD), eye muscle area (EMA), and faecal egg count measurements (FEC2).

Within a breeding program, animals available for selection include young rams and ewes with no progeny at the time of selection, and older rams with recorded progeny. The amount of data available for use in the selection process varies according to the age and sex of animals. Rams are split into 3 different ram types according to their age.

1. Young untested "new" rams have no progeny information available, but do have their own trait records along with WWT and SUR records from up to 120 paternal half sibs. Other trait records available include CW, EMA and FD on 100 paternal half sibs, 80 FEC2 and 60 HFW paternal half sib records. Information is also derived from their mother and 15 paternal aunts, who each have 2 performance records for NLB.
2. Emerging rams have been used once either as a hogget or a two tooth ram and have progeny information, including 100 records for WWT and SUR, 80 for CW, 60 for FEC2 and 50 for HFW, EMA and FD but have no daughter lambing records, so NLB information is still derived from their mothers and paternal aunts.
3. Mature rams have a complete set of progeny data, including 120 progeny records for WWT, SUR, CWT, and HFW. They also have 100 progeny records for both EMA and FD, and 80 FEC2 progeny records. NLB information is available from performance records of 40 female progeny and an additional 40 paternal half sibs.

Ewes are assumed to have the same records as those of young rams, with the exception that they have no HFW records for themselves or on their half sibs. Genetic and phenotypic correlation matrices are incorporated into the selection index calculation to account for known positive and negative associations between correlated traits.

Five genomic scores have been modeled, where each genomic score has a heritability of 0.95 (allowing for genotype error) and a correlation of 0.25, 0.5, or 0.75 with the corresponding breeding objective traits; litter size, carcase weight, meat yield, lamb survival and internal parasite resistance.

Selection index calculations are made specifically for each candidate type, where the superiority of the top 5% of rams selected for mating and the top 60% of ewe lambs going into the flock as replacement ewes each year is estimated (based on index). Results are reported for each specific target trait individually and as an overall economic index response from all target traits. Results have then been aggregated into an overall annual breeding program response (cents) taking into account the generation interval associated with each candidate type.

The generation interval of the ewes was fixed based on an average age of 3.3 years when their lambs are born. Rams in the average industry breeding program design had an average age of 2.28 years, comprising 44% young untested rams (aged 1 and 2) , 40% emerging rams (aged 2 and 3) and 16% mature rams (aged 3 and 4). In comparison the predominantly young ram program had an average ram age of 1.5 years, with 75% being untested rams (aged 1 and 2) and 25% emerging rams (aged 2 and 3).

RESULTS AND DISCUSSION

Response with conventional selection criteria. Responses for individual traits by candidate type and for the two industry breeding programs were estimated. The overall superiority of the top 5% of untested young rams over the total group of young rams available for selection is 381 cents. This increases to 574 cents for the top 5% of mature rams which have progeny information available for selection decisions. The selection intensity on ewe hogget replacements is much weaker with the top 60% of ewes valued at 109 cents above the average for all candidates. When the breeding program response is calculated taking into account the effect of generation interval, the overall response for the industry average breeding program was 106 cents, compared to 111

cents for the predominantly young ram breeding program. These modest differences reflect the trade-off between the increased information available when making decisions on mature rams versus the reduced generation interval that can be obtained by selection of young rams. These rates of progress are slightly above the average rates of progress for dual purpose flocks reported by Amer (2009).

Response with genetic markers. Single or multiple genomic scores may be used alongside the production trait information to assist in the selection of young animals. If multiple genomic scores are used, then the additional gains made by use of genomic scores for a single objective trait are diluted by the competing objective traits which also have additional genomic information. Table 1 shows the breeding program response for the objective traits, when no genomic scores are used, a single genomic score is used for each objective trait, and all 5 genomic scores are used simultaneously.

Table 1. Annual breeding program response using production traits and genomic scores (cents)

	Industry average ram program			Young ram program		
	Traits only scores	SingleGS score	MultipleGS	Traits only scores	SingleGS score	MultipleGS
NLB	3.0	6.2	5.3	2.2	6.8	5.5
Growth	58.8	67.3	60.1	63.1	76.8	67.0
Disease	17.2	24.1	20.8	17.5	28.0	23.3
Survival	4.1	10.9	9.4	3.7	12.8	10.6
Yield	11.7	22.4	20.4	12.4	25.8	22.5
Fleece	10.6	10.2[1]	8.8	11.8	11.2[1]	9.3
Total	105.5	109.7[2]	124.8	110.7	116.8[2]	138.2

[1] Average response to fleece weight when a genomic score for other target traits was applied
[2]Average increase in overall response due to selection using a single genomic score

Results are shown for industry average and young ram breeding programs. When any single genomic score is used, the breeding program response to that target trait increases, resulting in selection being diverted away from other traits. When all five genomic scores are used, selection pressure is applied to all traits affected by the genomic scores resulting in an increased response for all affected traits, which is only marginally smaller than that seen when a single genomic score is applied. No genomic score was modeled for fleece weight resulting in a reduction in its response, for all situations when one or more genomic scores were included as selection criteria.

Effect of genomic score accuracy on response. Figure 1 shows the overall breeding program response to use of individual genomic scores, as the accuracy of prediction for the genomic scores is increased. As the correlation between the genomic score and the objective trait increases, non linear gains were observed in both individual trait and overall responses. Largest gains were observed using markers for growth and carcass yield with lesser gains observed in disease, survival and NLB.

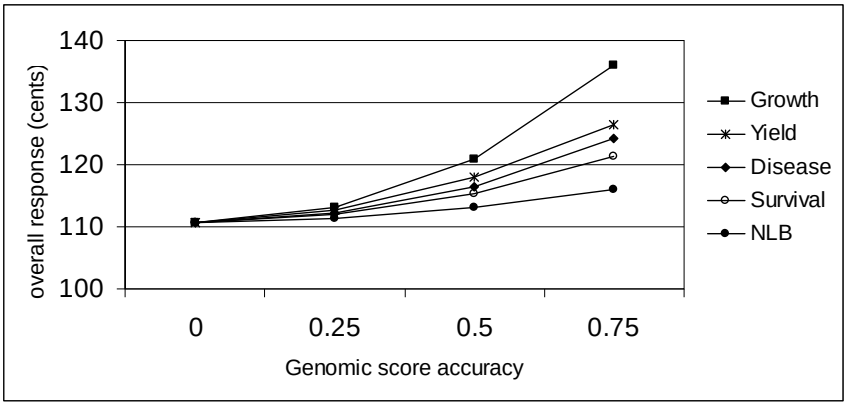


Figure 1. Annual response to selection at varying accuracies for genomic scores

CONCLUSIONS

Response to selection predicted using selection index theory was found to be relatively robust to the balance of ages of male selection candidates. However, additional response to selection when genomic scores were added as additional selection criteria were higher in the breeding program with a greater use of young rams. With genomic score accuracies of 0.5 for key breeding objective traits, rates of genetic progress are expected to increase by approximately 5% per annum. Proportionately greater increments of response are expected as the accuracy of the genomic scores increases beyond 0.5.

REFERENCES

Amer, P.R. (2009) *Proc NZ Society of Animal Production* **69:** (In Press)
Dekkers J.C.M (2007). *J Anim Breeding & Genetics* **124**:331
Meuwissen, T.H.E., Hayes, B. and Goddard, M.E. (2001) *Genetics* **157**:1819

ASSOCIATION OF POLYMORPHISMS IN CANDIDATE GENES WITH CARCASS AND TASTE PANEL ASSESSED MEAT QUALITY TRAITS IN A COMMERCIAL POPULATION OF ANGUS-SIRED BEEF CATTLE

J.L. Gill[1], S.C. Bishop[1], C. McCorquodale[1], J.L. Williams[2] and P. Wiener[1]

[1]The Roslin Institute and R(D)SVS, University of Edinburgh, Edinburgh, UK,
[2]Parco Tecnologico Padano, Lodi, Italy

SUMMARY

Associations between polymorphisms in candidate genes and economically important meat quality traits were assessed in a commercial population of Aberdeen Angus-cross animals. A number of traits were measured including 20 carcass and sirloin measurements recorded shortly after slaughter and also following maturation, 1 mechanical measure of tenderness and 7 taste panel assessed sensory traits. Polymorphisms tested included those in the calpain, leptin and myostatin genes. A number of significant and potentially important associations were found. An association was observed between a SNP in the calpain gene and meat tenderness, measured by both the tenderometer and the taste panel (P=0.01 for both), where the allele associated with tenderness was also associated with heavier hindquarters. Additionally we found significant associations between a leptin SNP and overall liking (P=0.02) and a DGAT1 SNP and sirloin fat depth (P<0.05).

INTRODUCTION

Meat quality is of great importance to the beef industry, but quality traits such as tenderness and juiciness often have low heritabilities and can only be measured post-slaughter, reducing the effectiveness of traditional phenotype-based breeding strategies. Marker-assisted selection has the potential to solve these problems and several markers associated with quality traits have recently been discovered in genes such as calpain (*CAPN1*), calpastatin (*CAST*), bovine growth hormone receptor (*BGHR*) and acylCoA:diacylglycerol acyltransferase 1 (*DGAT1*). Before such information can be used in breeding programmes, unbiased and independent validation studies should be carried out to establish whether the effects are found in the breed and population of interest. The aim of this study was to test for associations using polymorphisms in various candidate genes in a sample of over 400 Scottish Aberdeen Angus-cross animals collected through a commercial abattoir.

MATERIALS AND METHODS

Data. The sample set used to test the majority of polymorphisms consisted of 443 commercial crossbred beef cattle with purebred Aberdeen Angus sires sourced through the Scotbeef abattoir (Bridge of Allan, Scotland). Cattle originated from 14 breeder finisher farms (i.e. farms where animals are bred and finished on the same farm) and were selected to be representative of British commercial cattle slaughtered for beef production, i.e. a mix of heifers and bullocks (castrates). The data set used to test the myostatin polymorphism included a further 93 animals from the same source. Twenty carcass and sirloin traits, such as hot carcass weight, hindquarter weight, sirloin weight and eye muscle area were measured shortly after slaughter and also following maturation. In addition, 7 taste panel assessed sensory traits were measured, as well as a mechanical measure of tenderness, recorded using a tenderometer machine.

Genotyping. Samples were genotyped for 9 polymorphisms in 6 different genes. The genes (and polymorphisms) were *CAPN1* (*CAPN316*, *CAPN4751*), *CAST* (*UoGCAST*), Leptin (*UASMS1*, *UASMS2*, *Exon2FB*), *BGHR* (1 single nucleotide polymorphism (SNP)), *DGAT1* (1 dinucleotide substitution) and myostatin (*GDF-8*) (an 11 base-pair deletion referred to as *del*11). All animals with phenotype information were genotyped, as were all available sires.

Data Analysis. The associations between the different genotypes of each SNP and the various traits recorded were evaluated using a single marker mixed model association analysis. Data were analyzed by fitting a linear mixed model using the restricted maximum likelihood method (REML) provided in Genstat, release 10 (Payne *et al.* 2007). The statistical model included fixed effects of farm, genotype, sex and the genotype-sex interaction, and random effects of sire, slaughter date (panel date for the taste panel traits), interaction of sire and slaughter date (panel date for the taste panel traits) and interactions of sire and slaughter date (panel date for the taste panel traits) with the genotype/sex groups. These latter interactions took into account the possibility of genotype/sex effects varying with sire or slaughter date (panel date for the taste panel traits) or both. Additionally, random effects were constrained to be non-negative, i.e. effects that were estimated to be negative were set to zero. Statistical significance for the fixed effects was determined using approximate F-statistics with denominator degrees of freedom estimated in the Genstat REML procedure (Kenward and Roger 1997).

Additive effects and dominance deviations were calculated from the predicted genotype means. The additive effect was estimated as half the difference between the mean of the 2 homozygotes, and dominance was estimated as the deviation of the heterozygote from the mean of the 2 homozygotes (Falconer and Mackay 1997).

RESULTS AND DISCUSSION

Using single-marker, mixed-model association analysis 5 of the polymorphisms were found to be significantly associated with one or more of the traits tested. These include the *CAPN316* SNP in the *CAPN1* gene, the *UASMS2* SNP in the Leptin gene, the *DGAT1* SNP, the *BGHR* SNP and the *del*11 polymorphism in the *GDF-8* gene. Table 1 shows the genotype means and P-values for these significant associations.

Significant associations were found between *CAPN316* genotype and meat tenderness, measured using both the tenderometer machine and the taste panel, which is in agreement with previous findings (Page *et al.* 2005). Animals with the CC genotype at this SNP had more tender meat (P=0.01) as expected from the earlier study, however, they also had heavier hindquarters. No associations were found between the *CAPN4751* SNP, and tenderness, or between the *CAST* SNP and tenderness as previously documented (Casas *et al.* 2006). No associations were found between the SNPs in the leptin gene and carcass quality traits as had been previously shown (Nkrumah *et al.* 2005; Schenkel *et al.* 2005). However, the leptin SNP, *UASMS2*, was found to be significantly associated with overall liking (P=0.02) measured by the taste panel, where panellists gave animals with the TT genotype significantly higher overall liking scores than animals with CC or CT genotypes. This SNP has previously been associated with both backfat thickness and marbling score with TT animals having higher values for both traits (Nkrumah 2005). Additionally, the present study showed a significant, although non-additive, association between this SNP and sirloin fat depth where animals with the CC genotype had the lowest fat depth. The *DGAT1* polymorphism has been shown to be significantly associated with milk fat yield and fat percentage, with AA animals having increased levels of both (Grisart *et al.* 2005). The present study found that this polymorphism was associated with sirloin weight after maturation (P=0.04) and sirloin fat depth (P<0.05). In both cases the A allele was associated with the higher value indicating that the increase in sirloin weight was probably due to the increase in the depth of fat

surrounding the muscle. Regarding the myostatin gene, there were no *GDF-8* homozygous *del*11 animals in the population studied and only 39 *wt*/*del*11 animals. However, as with previous studies that have shown that a single copy of the *del*11 allele has effects on carcass characteristics (Wiener *et al.* 2002; Casas *et al.* 2004), the data reported here show that animals inheriting the mutant *del*11 allele were heavier at slaughter ($P<0.05$), with higher Conformation Class scores ($P<0.05$) and heavier hindquarters ($P<0.001$); they had heavier sirloins, both before ($P<0.001$) and after ($P<0.01$) maturation and larger eye muscle areas ($P<0.05$), suggesting a general increase in muscle mass for those animals with a single copy of the *del*11 allele. Bonferroni corrections for multiple hypothesis testing reduced the number of apparently significant associations; however with a moderately stringent correction (6 genes as SNPs in the same genes were found to be in strong linkage disequilibrium) those associations with P-values less than ~0.01 remained significant.

Table 1. Genotype means, standard errors and P-values for polymorphisms with significant trait associations

Trait	SNP	Genotype means ± se			Nominal P-value
	CAPN316	CC	CG	GG	
Tenderometer score (kPa)		22.25 ± 1.16	24.24 ± 0.70	25.18 ± 0.67	0.01
Hindquarter weight (kg)		75.67 ± 1.57	72.05 ± 0.84	71.84 ± 0.77	0.04
Tenderness		6.00 ± 0.16	5.79 ± 0.08	5.63 ± 0.07	0.01
	UASMS2	CC	CT	TT	
Overall liking		5.59 ± 0.08	5.55 ± 0.08	5.80 ± 0.10	0.02
	DGAT1	AA	AG	GG	
Sirloin weight after maturation (kg)		8.31 ± 0.46	7.17 ± 0.12	7.14 ± 0.11	0.04
Sirloin fat depth, mm		11.11 ± 1.65	6.62 ± 0.4	6.53 ± 0.33	0.05
	BGHR	AA	AT	TT	
Odour		5.64 ± 0.19	5.24 ± 0.08	5.16 ± 0.06	0.02
	GDF-8	*wt*/*wt*	*wt*/*del*11	*del*11/*del*11	
Hot carcass weight (kg)		314.7 ± 3.2	332.1 ± 6.6	-	0.01
Sirloin weight before maturation (kg)		7.25 ± 0.1	7.88 ± 0.2	-	0.001
Sirloin weight after maturation (kg)		7.17 ± 0.1	7.70 ± 0.22	-	0.01
Confirmation class		7.19 ± 0.15	7.92 ± 0.32	-	0.02
Eye muscle area (mm^2)		11077 ± 152.2	11890 ± 362.8	-	0.02
Hindquarter weight (kg)		72.2 ± 0.64	77.0 ± 1.44	-	0.001

There were significant sex-genotype interactions for 6 of the significant trait-polymorphism combinations tested. Analyses of the mean trait value for each genotype in each sex for those trait-SNP pairs indicated that, for the majority of associations, the effect was primarily in female animals. The exception was the effect of *CAPN316* on tenderometer values which was observed in the male animals. Differences between male and female genotype effects were seen in 5 traits: taste panel assessed tenderness, weight of hindquarter, odour, sirloin weight after maturation and sirloin fat depth. Differences could be partly due to the limited number of females (135) in the analysis when compared to the males (308) although allele frequencies for both sexes are similar.

Alternatively, trait expression could be strongly correlated with fatness. Means for each sex showed that females tended to have higher fat class scores than males (data not shown). Therefore, it is possible that the female animals are more likely to express genetic differences in traits correlated with fatness.

CONCLUSIONS

Polymorphisms in 6 genes were tested for their associations with various economically important traits in a commercial population of Aberdeen Angus-cross animals. The results presented here confirm some of the previously-documented associations, for example, the association between *CAPN316* genotype and tenderness, the most important quality trait for consumers. Furthermore, novel associations were identified which, following validation in other populations, could be incorporated into breeding programmes to improve meat quality. Finally, whilst some previously reported associations were not replicated in the current study, it is important to note that validation is dependent on the specific nature of the population screened and that genetic background may influence the size of the effect of a polymorphism. Validation failure may be due to a lack of true associations between the trait and marker but could also be caused by differences in SNP frequencies, different SNP marker-causative mutation linkage phases, genotype-by-environment interactions or epistasis (Dekkers 2004), or the way the trait is measured. Nevertheless, for those associations confirmed here, the additional validation instils confidence in using these markers in selection programmes for improved meat quality.

ACKNOWLEDGEMENTS

This work was funded by the Biotechnology and Biological Sciences Research Council, Scotbeef (Stirling, UK) and Genesis Faraday.

REFERENCES

Casas, E., Bennett, G.L., Smith, T.P.L. & Cundiff, L.V. (2004) *J. Anim. Sci* **82**:2913.

Casas, E., White, S.N., Wheeler, T.L., Shackelford, S.D., Koohmaraie, M., Riley, D.G., Chase, C.C. Jr., Johnson, D.D., Smith, T.P.L. (2006) *J. Anim. Sci.* **84**:520.

Dekkers, J.C.M. (2004) *J. Anim. Sci*. **82**:E313.

Falconer, D.S. and Mackay, T.F.C. (1997) "Introduction to Quantitative Genetics", 4th ed., Longman Ltd., England.

Grisart, B., Coppieters, W., Farnir, F., Karim, L., Ford, C., Berzi, P., Cambisano, N., Mni, M., Reid, S., Simon, P., Spelman, R., Georges, M., Snell, R., (2002) *Genome Res.* **12**:222.

Kenward, M.G. and Roger, J.H. (1997) *Biometrics* **53**:983.

Nkrumah, J.D., Li, C., Yu, J., Hansen, C., Keisler, D.H., Moore, SS. (2005) *J. Anim. Sci.* **83**:20.

Page, B.T., Casas, E., Quaas, R.L., Thallman, R.M., Wheeler, T.L., Shackelford, S.D., Koohmaraie, M., White, S.N., Bennett, G.L., Keele, J.W., Dikeman, M.E., Smith, T.P.L. (2004) *J. Anim. Sci.* **82**:3474.

Payne, R.W., Harding, S.A., Murray, D.A., Soutar, D.M., Baird, D.B., Welham, S.J., Kane, A.F., Gilmour, A.R., Thompson, R., Webster, R., Tunnicliffe, Wilson, G. (2007) GenStat Release 10 Reference Manual, VSN International, Hemel Hempstead.

Schenkel, F.S., Miller, S.P., Ye, X., Moore, S.S., J.D. Nkrumah, J.D., Li, C., Yu, J., Mandell, I.B., Wilton, J.W., Williams, J.L. (2005) *J. Anim. Sci.* **83:**2009.

Wiener, P., Smith, J.A., Lewis, A.M., Woolliams, J.A. & Williams, J.L. (2002) *Genet. Sel. Evol.* **34**:221.

Breeding program design including MAS

THE EFFECT OF THE INCLUSION OF PEDIGREE DATA ON ESTIMATES OF CARRIER STATUS AT THE AGOUTI LOCUS IN SHEEP

J. M.Henshall[1], J. McNally[1] and B.J. Norris[2]

[1]CSIRO Livestock Industries, FD McMaster Laboratory, Armidale, NSW 2350
[2]CSIRO Livestock Industries, St Lucia, Queensland 4067

SUMMARY

A simulation study was conducted to examine the effect of including pedigree data when estimating genotype probability at the agouti locus. The effect of errors in the pedigree was also examined. The proportion of non-carrier progeny identified increased by as much as 20% when pedigree information was included, particularly if genotype estimates for maternal grandsires were available. Pedigree errors had to be at very high rates before they adversely affected the accuracy of genotype estimates.

INTRODUCTION

Approximately 0.5% of Merino lambs born each year are recessive black, and therefore around 13% of the Australian Merino flock are estimated to be carriers of the black pigmentation allele. Variation in the Agouti region is known to be responsible for the recessive black condition (Parsons, Fleet and Cooper 1999), with the causative allele identified as having 1 copy of agouti signalling protein coding sequence, while dominant alleles responsible for no pigmentation in the fleece are composed of between 2 and 5 tandem copies of the gene (Norris and Whan 2008). An estimate of the number of copies can be obtained using a DNA based test where the 'junction point' between each copy of the gene is amplified (Norris and Whan 2008). The assay (referred to here as the "agouti assay") is an asymmetrical competitive PCR where a fluorescently-labelled common reverse primer is added at 100 fold less then the specific forward primers allowing quantitative measurement of the total gene copy number. In a diploid animal the total gene copy number does not distinguish how many copies of the gene are on each chromosome. Thus, the distinction between white animals with multiple copies of the gene at both loci and carrier animals that are heterozygous for the single copy allele (recessive black allele) and a multiple copy allele, cannot be determined. Further, the quantitative nature of the assay produces probabilities for each gene copy number rather than calls made with certainty. The assay returns a real number with expected value $\frac{n-2}{n}$, where n is the total gene copy number. This series converges as copy number increases, so high copy counts in particular are not known with certainty.

The utility of the assay can be enhanced by taking account the constraints imposed by Mendelian transmission of copy number alleles in pedigreed populations. Assay values can be used to specify penetrance values relating to the unobserved genotype, and segregation analysis on pedigreed populations can then provide estimates of the probability of each underlying genotype configuration. For commercial sheep populations a half-sib data structure will be usual, but it is unlikely that the pedigree will be without error as, unless paternal and maternal DNA parentage is carried out, pedigree errors will occur on the dam side even in flocks that lamb in sire paddocks. In this paper we examine the benefit that can be gained by including pedigree in addition to sire in analysing assay results. As this increases the risk that incorrect information is included in the analyses we also consider the impact of pedigree errors on the estimated genotype probabilities.

MATERIALS AND METHODS

Simulation. The population was modelled on 10 small stud flocks each joining 5 rams to 200 ewes each year over 20 years. Of the 5 rams, 4 were home bred and 1 came from a linked flock. Selection was on a trait uncorrelated to fleece colour except that no black animals were selected in the final generation. Allele frequencies estimated from a Merino sheep population were simulated in the base population. In the final generation, assay values were simulated for 16 progeny from each sire and for the sires themselves. Variance around the expected assay value was simulated with a coefficient of variation (CV) of 3%. The simulation was repeated 20 times producing a total of 1,000 halfsib families for analysis. For each simulated dataset, pedigree errors were superimposed as mismothering events where pairs of lambs were assigned to each other's mother at rates of 0%, 1%, 2%, 4%, 8% or 16%. The mismothering events were within halfsib groups, such as might occur when lambing takes place in sire paddocks or when paternity is determined through a DNA test. No paternity errors were simulated as when performing the agouti assay it is only a small amount of additional work to do sire pedigree verification using markers, thus the sire can be assumed to be known without error.

Analysis. Initial analyses took place using the software package Mendel (Lange *et al*. 2001), which computes the full likelihood for complex pedigrees. This restricts the size of pedigrees that can be analysed. Other software packages are available that lift the restriction by applying approximations to the full likelihood, however these would have required modification to allow for the penetrance function derived from assay data. To explore the effect of including pedigree information (with errors) in analyses for larger pedigrees, we used halfsib family analyses with modified penetrance function values for the un-assayed dams to simulate information flow from relatives of the dams. Our motivation for this approach was as follows: when analysing halfsib families, with the sire and progeny assayed, sire genotype probabilities are estimated with high accuracy. Therefore, when the analysis includes pedigree information above the halfsib family the sire estimates will not change much, but information will flow through the dam to the progeny. Most of this information will come from the maternal grandsire, whose own genotype may be known with high accuracy. This can be modelled in a halfsib analysis by modifying the dam penetrance values to account for knowledge of her sire's genotype. We considered levels of certainty of the maternal grandsire's genotype of 100%, 50% and 0%. In all of the analyses the penetrance function was calculated assuming a CV of 3% for the assay values.

RESULTS AND DISCUSSION

Analyses using the complete likelihood. Using Mendel, we had difficulty in estimating genotype probabilities in data with pedigree errors. In many cases the software execution aborted as the problem was too complex. This may be due to very small likelihoods, exacerbated by the pedigree errors. Even without pedigree errors we were unable to reliably analyse the 5 halfsib families from a flock with more than 4 progeny per family and with more than grandparents linking the families. Results for non-carrier progeny analysed using 4 progeny per family, the most we could achieve using Mendel, are presented in Table 1. There was a small benefit to including the pedigree information, but it was less than the benefit gained by assaying 8 progeny per family and analysing the resultant data as a halfsib family.

Halfsib analyses using modified dam penetrance values. Probabilities for non-carrier progeny are presented in Table 2. A maternal grandsire genotype certainty of 0.0 is equivalent to performing a half-sib analysis, and mismothering is irrelevant as only sire data is used in the

analysis. The effect of including information about the genotype of the maternal grandsires is shown by the difference between the results with a maternal grandsire genotype certainty of zero and a non-zero maternal grandsire genotype certainty. Including maternal grandsire genotype data increased the precision of the assay, lifting the proportion of lambs declared non-carriers with 99.9% certainty from 61% to as much as 73%. For a flock that has applied the assay for more than one generation the genotype of the maternal grandsires may well be known with high accuracy, and it is clear that there is great benefit from including their data in the analysis. Mismothering has a negligible effect on progeny genotype probabilities. Even at very high rates (16%) the effect is barely noticeable. Estimated carrier probabilities for non-carrier sires and dams were similarly improved by the inclusion of information about the paternal grandsire's genotype, and the effect of pedigree errors was minor (results not shown).

Table 1. Proportions of non-carrier progeny achieving <0.1%, <1%, <5%, >95%, >99% and >99.9% estimated probability of being a carrier for the undesirable allele. No pedigree errors were simulated

Analysis method	Family size	Estimated probability of being a carrier					
		P<0.001	P<0.01	P<0.05	P>0.95	P>0.99	P>0.999
Pedigree	4	0.39	0.65	0.78	0.00	0.00	0.00
Halfsib	4	0.38	0.62	0.76	0.00	0.00	0.00
Halfsib	8	0.50	0.71	0.81	0.00	0.00	0.00

Table 2. Proportions of non-carrier progeny achieving <0.1%, <1%, <5%, >95%, >99% and >99.9% estimated probability of being a carrier for the undesirable allele. There were 16 progeny in each halfsib family, with progeny and sires assayed

Grandsire genotype certainty	Mismothering (%)	Estimated probability of being a carrier					
		P<0.001	P<0.01	P<0.05	P>0.95	P>0.99	P>0.999
0.0	NA	0.61	0.74	0.81	0.00	0.00	0.00
0.5	0	0.67	0.77	0.82	0.00	0.00	0.00
0.5	1	0.67	0.77	0.82	0.00	0.00	0.00
0.5	2	0.67	0.77	0.82	0.00	0.00	0.00
0.5	4	0.67	0.77	0.82	0.00	0.00	0.00
0.5	8	0.67	0.77	0.82	0.00	0.00	0.00
0.5	16	0.66	0.77	0.82	0.00	0.00	0.00
1.0	0	0.73	0.79	0.85	0.00	0.00	0.00
1.0	1	0.73	0.79	0.85	0.00	0.00	0.00
1.0	2	0.73	0.79	0.85	0.00	0.00	0.00
1.0	4	0.73	0.79	0.85	0.00	0.00	0.00
1.0	8	0.73	0.79	0.85	0.00	0.00	0.00
1.0	16	0.72	0.79	0.85	0.00	0.00	0.00

For carrier progeny, estimates of probabilities of being a carrier were improved by the inclusion of information about the maternal grandsire's genotype, and again the effect of pedigree errors was minor (Table 3). Most importantly, the probability of being declared a non-carrier was

not increased by the inclusion of maternal grandsire data, even in pedigrees with high rates of mismothering. For carrier sires and dams, inclusion of maternal grandsire genotype information had a similar effect on estimated probabilities of being a carrier (results not shown).

Table 3. Proportions of carrier progeny achieving <0.1%, <1%, <5%, >95%, >99% and >99.9% estimated probability of being a carrier for the undesirable allele. There were 16 progeny in each halfsib family, with progeny and sires assayed

Grandsire genotype certainty	Mismothering (%)	Estimated probability of being a carrier					
		P<0.001	P<0.01	P<0.05	P>0.95	P>0.99	P>0.999
0.0	NA	0.00	0.00	0.01	0.65	0.61	0.57
0.5	0	0.00	0.00	0.01	0.68	0.65	0.59
0.5	1	0.00	0.00	0.01	0.68	0.65	0.59
0.5	2	0.00	0.00	0.01	0.68	0.64	0.59
0.5	4	0.00	0.00	0.01	0.67	0.65	0.59
0.5	8	0.00	0.00	0.01	0.67	0.65	0.59
0.5	16	0.00	0.00	0.01	0.67	0.64	0.59
1.0	0	0.00	0.00	0.01	0.71	0.70	0.66
1.0	1	0.00	0.00	0.01	0.71	0.70	0.66
1.0	2	0.00	0.00	0.01	0.71	0.69	0.66
1.0	4	0.00	0.00	0.01	0.71	0.69	0.65
1.0	8	0.00	0.00	0.01	0.71	0.69	0.65
1.0	16	0.00	0.00	0.01	0.71	0.68	0.65

CONCLUSIONS

When using Mendel to analyse small families of 4 halfsibs, including pedigree information in the analysis produced a small increase in the precision of estimates of genotype probabilities. Analyses on larger families of 16 halfsibs using our simplified model, where pedigree information was simulated by including knowledge of the maternal grandsire's genotype, suggests that greater improvements are likely to be achieved where information on genotype probability is available for maternal grandsires. Although the estimation can be carried out without software for segregation analysis in deep pedigrees, ideally all of the available data would be analysed together. Pedigree errors on the dam side had negligible effect on accuracy of estimates.

REFERENCES

Lange K., Cantor R., Horvath S., Perola M., Sabatti C., Sinsheimer J. and Sobel E. (2001) *Amer. J. Hum. Genetics.* **69(Supplement)**: A1886.

Norris, B.J. and Whan, V.A. (2008) *Genome Research*, **18**:1282.

Parsons, Y.M., Fleet, M.R. and Cooper, D.W. (1999) *Aust. J. Agric. Res*. **50**:1099.

I HAVE A DRAFT GENOME FOR MY SPECIES…WHAT NOW?

Claire M. Wade

Faculty of Veterinary Science, University of Sydney, Sydney NSW 2006

SUMMARY

As researchers, we need to question our objectives in gene mapping. Are we seeking merely a genetic test, or do we want to find the causative mutation so that we can understand the underlying biology. For a number of practical applications, a DNA test may be sufficient to assist the research sponsors, but as scientists, we have the opportunity to learn more about biology from our research. Additionally, it is becoming increasingly difficult to publish work that does not report a functional relationship with the association. This might be established through the discovery of an exonic mutation or through means such as expression analysis or other functional analysis if the mutations are regulatory.

Given that we have decided to proceed to mutation discovery, we require an effective approach. Our experiences with trait mapping in dog and horse have shown us that many mutations are regulatory, and that these can be elusive. Our use of target enrichment and sequencing with Massively Parallel techniques have shown us that there might be many mutations that are in LD with our trait of interest and that prioritizing these is the first step in establishing a functional basis to the phenotype.

INDRODUCTION

As we enthusiastically rush to gather the tools to enable genetic trait mapping in our species of interest that we often fail to stop and imagine what might be needed once we are successful in actually mapping something. In the past few years, more than 24 mammalian genomes have been taken to full draft coverage. The projects that generate the assemblies of these genomes frequently provide other essential resources such as SNP maps and the computational annotation of genes and features onto browsers that are easily accessible to the general public. These SNP resources have been used to produce genotyping arrays in most domestic species. This talk will focus on some lessons learned from array design and gene mapping thus far in the mouse, dog and horse. Some preliminary experiences with the use of Massively Parallel Sequencing (MPS) for targeted sequencing in the horse will also be discussed.

EFFECTS OF POPULATION HISTORY ON LINKAGE DISEQUILIBRIUM AND MAPPING SUCCESS

The architecture of linkage disequilibrium (LD) in mammalian genomes is a product of their particular population histories. In particular, factors leading to population bottlenecks have two large effects. First, population bottlenecks create founder effects that reduce the number of alleles that may occur at any given locus. Second, the bottlenecks have an effect of re-setting the "recombination clock". That is, whole haploid chromosomes are forced into the new population and from that point onward, the alleles or haplotypes on those chromosomes are inherited as a unit until separated by recombination. The effect of this is to drastically lengthen LD in the short to medium term.

In the genome of the mouse, inbreeding and the creation of laboratory mouse strains from relatively few founders in the past 200 years has led to an extreme of this process. In essence, the inbred laboratory mouse strains are recombinant inbred lines of the different mouse sub-species, with the main contributors being *Mus musculus domesticus* and *Mus musculus musculus* but other contributors are *Mus musculus castaneus and Mus musculus molossinus*. This suggests that it

should be relatively easy to map genetic traits in mice, but that the resolution of the mapping will be coarse. The LD across inbred laboratory mice measures over an average of two megabases (Mb). Even when a trait is successfully mapped, there is still much territory to sort through to try and find the mutation(s) responsible. Mutation detection through mapping in the mouse has not been extremely successful to date because of this problem of coarse resolution.

The population history of the dog provides advantages for gene mapping relative to the mouse. The history of dog domestication and breed creation is such that this species has been through two different major bottlenecks. The first bottleneck occurred at the point of human domestication of dogs. It is well established now that domestic dogs have been derived from Grey Wolves from Europe and Asia. Because wolves can be dangerous, relatively few individuals were taken from the wild to be tamed by humans. Those few that were taken forced a situation where relatively few chromosomes entered the dog population, but the long time period since domestication commenced (10-40K years ago) has recombined these few chromosomes very well so that if the domestic dog population is examined as a whole, the LD is very short. In humans the mean LD is regarded as short at 15-25 kilobases (Kb), while across all dog breeds it is even shorter and is of the order of 8-10Kb.

This would at first suggest that to undertake successful gene mapping in the dog, we will need more than twice as many genetic markers (most commonly used now are SNPs) than for mapping similar traits in humans. Fortunately, that is not the case. A second feature of dog genetic history is that during the Victorian era, humans became fascinated with the concept of dog breeding and showing. As a result, kennel clubs were formed and groups of dogs "breeds" that were genetically isolated from other domestic dogs were created. The effect of this second period of recent population bottlenecking is to create a situation in which the LD within a breed is of mouse-like proportions. In fact, the mean within-breed LD in the dog is of the order of one megabase. Indeed, this particular population history creates an ideal situation. If we have a mendelian trait that is segregating within a breed, then we can map the genes influencing the trait at coarse resolution with relatively few markers. Then if we wish to narrow the interval of association we can simply use other breeds that segregate the trait and make use of the short across-breed LD.

The genome of the horse has been recently completed. Horses are estimated to have been domesticated between 4,000 and 6,000 years ago and no populations of undomesticated horses exist in the wild other than perhaps the Przewalskii (Mongolian-wild) horse. This also has been shown to be introgressed with domestic stock. It seems that, unlike the dog, the domestication process of the horse has resulted in the capture of all the horses from the wild and so there is no true "domestication bottleneck". The second bottleneck akin to the breed creation bottleneck in dog does however exist. But the genetic isolation of horse breeds is substantially reduced compared with dog breeds. An exception is the thoroughbred horse. The thoroughbred has been developed from few founders and has not allowed introgression from other breeds for a long period of time. The LD in this breed is much like that of a dog breed. The average within-breed LD in horses is approximately 150Kb and for the Thoroughbred is closer to 500Kb. Across breeds, there is much haplotype sharing and the LD remains relatively long at 50-75Kb. A SNP map of 1.2 million SNPs was generated as part of the horse genome assembly project. The SNPs were derived from light whole genome shotgun sequencing of 9 horse breeds sampled from the horse populations of Europe, the Americas and Asia.

THE RIGHT TOOL FOR THE JOB: GENOTYPING ARRAYS

All of the aforementioned species have genotyping arrays now available. For the dog, arrays were designed on both Affymetrix and Illumina iSelect platforms. The long LD within dog breeds suggested that between 15 and 20,000 SNP would be needed for trait mapping. Affymetrix arrays

holding initially 63K and later 128K SNPs were designed and an Illumina iSelect array of 23K SNPs that had been validated on the Affymetrix arrays.

The Affymetrix platform uses a restriction digest of the genomic DNA with either one or two restriction enzymes (usually only one for non-human species). The fragments from the restriction are amplified and sized and a particular size range of fragments is used to take part in the genotyping. This strategy has two disadvantages. First, in the design phase of the array, the number of SNPs occurring in parts of the genome residing in these expected fragments is much reduced. Only about 10% of the discovery SNPs can be used for the assay design. On this system, the most efficient approach is to pre-digest the genome with the chosen enzyme and then to sequence only the fragments in the desired size range for SNP discovery at the beginning. Second, the restriction process requires that the DNA quality should be very good and this typically results in an unsuitability of samples derived from buccal swabs or hair samples for whole genome genotyping purposes. This is because DNA degradation pre-cuts the genome, so that when the restriction digest is carried out the DNA in the desired fragment size range may not be the DNA that you are expecting and array performance is extremely poor. The high genotyping success rates reported for arrays such as the Human arrays are never observed on non-human species arrays because the SNPs used for non-human mammals are not usually pre-validated on the Affymetrix system. For the mouse arrays, two 256K SNP arrays (512K SNP total) yield 148K usable, polymorphic SNP. For the dog arrays, the 63K SNP array yields 27K usable, and the 128K array design yields 50K.

The Illumina iSelect platform makes use of whole genomic DNA and so is relatively unaffected by sample degradation. Also, because it does not require pre-digestion of the genome in the array design, all known SNPs are available for array design. The yield from the arrays is very good – typically of the order of 85-90%. For the dog Illumina array, 24K SNP were designed for a yield of 22K SNP. For the horse array, exactly 60K SNPs were designed and the yield is 54K. This array is considerably more expensive than the Affymetrix platform and the array processing facilities are less available but the data quality is exceptionally good.

In 2007 the horse research community formed a consortium to produce a horse genotyping array. Power calculations suggested that an array of 150,000 SNP was desirable. While the Affymetrix platform offered 1 million SNP designs in an affordable package, the horse community only had 1.2 million polymorphisms available for design of which only 10% were expected to co-occur with the fragment sizes produced by restriction digest. Thus the only Affymetrix option available was the 128K design with an expected yield of 50K SNPs. The Illumina 60K option had a similar expected yield. The Illumina iSelect genotyping system was chosen for its high data fidelity and for the capacity to use the greatest proportion of the SNPs available. The expense of this array meant that the community was unable to afford more than 60,000 bead-types but the samples used by the community came from predominantly swab or hair DNA and so were subject to degradation.

SUCCESS MAPPING WITH GENOTYPING ARRAYS IN THE DOG

Successful coarse-resolution mendelian trait mapping in the dog has been carried out with as few as 7 cases and 10 control animals. Because of the structure of the genome in dog populations, there is the advantage that if the same trait segregates in multiple breeds, the other breeds can be brought in for fine mapping to reduce the interval of association. Of course, the density of markers required for fine mapping is very much greater than that of the genome-wide mapping runs. There are several examples of the successful use of this approach in the dog (Karlsson *et al*. 2007; Salmon *et al.* 2007; Drogemuller *et al.* 2008; Awano *et al.* 2009).

A different breed segregating the trait of interest is, however, not always available and this can create a substantive impediment to mutation detection. Even so, SNP and other polymorphisms

within the broader interval can be used for robust genetic testing if mutation discovery can be delayed as is the case for the test for Hyperparathyroid tumors in Keeshonden (http://www.akcchf.org/news/index.cfm?article_id=145). Otherwise, alternate bio-informatic approaches can be used to identify candidate genes in the interval of association as has been used for mutations associated with Degenerative Myelopathy and Rod-Cone dystrophy in dogs (Wiik *et al.* 2008; Awano *et al.* 2009). In both cases, literature searches were used to identify candidate genes within the interval of interest and exonic mutations were identified. In the case of the Keeshonden, the causative mutation remains elusive although the genetic test is 100% effective in this breed. Of the traits described, the mutations of three are exonic (Rod-Cone Dystrophy and Degenerative Myelopathy, Chinnese Crested), two are regulatory (White boxer, ridge) and one is unknown (Keeshonden). We expect that many non-lethal mutations will be regulatory in nature.

PROOF OF PRINCIPLE MUTATION DETECTION IN THE HORSE INTEGRATING MASSIVELY PARALLEL SEQUENCING

As part of the horse genome analysis (Wade *et al.* 2009) we chose to integrate positional mapping with new sequencing technologies including MPS to gain an idea of the likely success of these technologies for mutation detection. To do this we chose to study four mendelian coat-colour traits, but only three were successful because two horses with the fourth phenotype failed in sequencing. In each case, exons within the mapped interval had been already been assayed by limited sequencing with PCR and no exonic mutations had been identified. For this reason the mutations were expected to be regulatory.

While MPS is a very cheap way of generating sequence, it is indiscriminant with respect to target. Eight lanes of sequence with the Illumina Genome Analyzer generates about 2Gb of random sequence and costs about $15K US. If you give the sequencer genomic DNA, then you will get light cover of the entire genome and perhaps considerable cover of mitochondrial DNA depending on the tissue from which DNA was extracted. The sequences produced by MPS have considerable error rates with error types that vary depending on the platform used. This means that you must have at lease five fold cover of the target to reliably call SNP or Insertion-deletion events and possibly even more cover to call copy number variants. A more efficient method of carrying out mutation discovery using MPS is to enrich the sequencing for target DNA so that the cover can be effectively increased and more individuals (including replicated affected and unaffected individuals) can be sequenced.

Achieving effective enrichment of sequence in the target region presents one of the major challenges of the new sequencing technologies. A number of methods of target enrichment are available. Some involve long-range PCR and others involve target capture using tiled genomic sequence from a draft genome. Long-range PCR is suitable if it is expected that there will be considerable divergence from the draft genome, or when no draft genome is available. For the horse analysis we used a hybrid capture technique. For this, we sent draft-genome sequence from the target region to Nimblegen and had high resolution comparative genome hybridization slides constructed for the four regions. All four regions were on the same hybridizing array. DNA from affected horses for the four traits was hybridized to the slides and the slides were washed to remove the extraneous DNA. Next, the hybridized DNA was eluted from the slide and amplified using PCR. The amplified products were sent for Illumina Genome Analyzer II (Solexa) sequencing.

In horse, the strategy employed was to take regions that had been previously identified to harbour genes influencing the traits by genetic mapping with microsatellite markers. The regions to be assayed varied considerably in size, with the smallest being 300Kb and the largest 10Mb. We used Sequenom Mass Spectrometry genotyping with one or two pools of SNPs (each pool may contain up to 35 SNP) and many horses from different breeds to reduce the intervals to

manageable size for MPS. To do this effectively we needed sufficient density of SNP to see haplotype breaks in individual horses. In the horse, this ideally requires at least 5 SNP per 100kb. For the largest region, to save funds we gambled and focused the fine-mapping in the vicinity of a gene that had an expression difference and so we used sparser than desired genotyping over 2MB. If this had not revealed an association we would have broadened the search but this was not required. In each case, by fine mapping we were able to reduce the interval of association to 200-300Kb. These intervals were tiled on hybridizing arrays and became our sequencing targets.

The success of this approach was varied. The best performing samples achieved coverage of 150× of the target region with 70% of all sequence tags falling within the enriched region. This was from 1 lane of sequencing (tag size 35 base pairs). The worst sample had 0.2× cover of the target (mainly in repetitive sequence). Alignment to draft genomic sequence was carried out by three methods. MAQ (Li *et al.* 2008), Spines-aligns (Maucelli pers. comm.) and Smatch (Kirby pers. comm.). The latter two methods are under development at the Broad Institute of Harvard and MIT. All of these alignment methods worked well over the limited regions assessed. MAQ is freely available on the web. The samples with low cover were determined to be primarily affected by hybridization failure.

One of the traits assessed (Grey) was known to be associated with sequence duplication and this was readily detected in the normalized sequence tags by assessing mean draft genome coverage base by base. As expected, many mutations (predominantly SNP and some insertion-deletion events) were observed to be in LD with the expected associations. To prioritize these mutations for functional study, we made use of transcriptome profiling data(Coleman 2009) and also conserved element analysis (Garber *et al.* 2009). Briefly, the transcriptome profiling resulted from mRNA-seq using Illumina Genome Analyzer II on eight horse tissues (none were skin). The conserved elements analysis is the result of multiple alignments of sequences from 24 mammals with at least draft genome (7×) coverage. The conserved elements are detected with word sizes of 8 bases or more. These are considered to be regulatory elements of evolutionary significance. At the time of writing, two of these mutations are maintaining association over a larger set of horses from the affected breeds.

I HAVE A DRAFT GENOME FOR MY SPECIES…WHAT NOW?

For many genomes so far, there has been sufficient community fundraising to enable the production of large scale commercial genotyping arrays. Typically the funding required to create an array is around $1million USD and so this may be beyond the reach of species without commercial or public importance. Once it is decided to create an array, the choice of array technology must be influenced first by the population genetics of the species. For species without recent population bottlenecks, it is likely that a large number of SNP will be required to map even mendelian traits effectively, and should sufficient SNP exist, care must be taken to identify those that can be successfully used on the genotyping platform chosen and the tissue types that will be commonly used for DNA extraction. For communities unable to generate sufficient funding for commercial array production other methodologies are required, such as large scale hybridization arrays combined with MPS technologies. At this time the success of such approaches can only be speculated but with improvements in assembly of paired-end MPS data e.g. Velvet (Zerbino and Birney 2008) a draft genome may not even be required.

Given sufficient SNP density, mapping projects using commercial large scale genotyping arrays will inevitably lead to the successful mapping of mendelian traits at the very least. Depending on the distance to the last strong population bottleneck, the resolution of this mapping might be quite coarse and the researcher may be left with a considerable region to analyse to discover the causative mutation. This region should ideally cover unassociated flanking sequences to ensure that the mutation lies within the assayed interval. Affected and unaffected individuals

should be sequenced as replicates with sufficient coverage to enable confident mutation detection. Fine mapping can be used in multi-breed data to effectively reduce the size of assayed region so that coverage may be increased.

As MPS becomes cheaper it may be feasible to sequence an entire region of association from within a single breed affordably or possibly the entire genome. Given that in our limited regions we typically identified more than 100 associated mutations, the number of potential mutations that would be identified by larger scale sequencing would be truly staggering and require significant bioinformatic analysis for everything from assembly to alignment and the detection of high quality disparities between affected and normal individuals. Once disparities are identified, rational approaches must be employed to reduce the search space for the causative mutation. Bioinformatic approaches based on literature review, transcriptomics and conservation have the ability to reduce the search space by more than 90%. Enormous amounts of data result from even limited use of MPS technologies. Tera-/Peta-byte levels of space for processing and storage of data are required along with the computational expertise and equipment to analyse the data. These represent significant impediments to the successful application of MPS without target resolution in the short to medium term.

REFERENCES

Awano, T., Johnson T.G., Wade, C.M., Katz, M.L., Johnson, G.C., Taylor, J.F., Perloski, M., Biagi, T., Baranowska, I., Long, S., March, P.A., Olby, N.J., Shelton, G.D., Khan, S., O'Brien, D.P., Lindblad-Toh K. and. Coates J.R. (2009) Proc Natl Acad Sci U S A **106**: 2794

Coleman, S. J., Zeng S., Wang K., Khrebtukova I., Luo S., Mienaltowski M.J., Schroth G., Liu J., and MacLeod J.N. (2009) I*n preparation.*

Drogemuller, C., Karlsson, E. K., Hytonen, M. K., Perloski, M., Dolf, G., Sainio, K., Lohi, H., Lindblad-Toh, K., and Leeb, T. (2008). **321**: 1462.

Garber, M., Guttman, M., Clamp, M., Zody, M. C., Friedman N. and X. Xie (2009) *Bioinformatics* **25**: i54

Karlsson, E. K., Baranowska, I., Wade, C. M., Salmon Hillbertz, N. H., Zody, M. C., Anderson, N., Biagi, T. M., Patterson, N., Pielberg, G. R., Kulbokas, E. J., Comstock, K. E., Keller, E. T., Mesirov, J. P., von Euler, H., Kampe O., Hedhammar, A., Lander, E. S., Andersson, G., Andersson, L. and Lindblad-Toh, K. (2007). *Nat Genet* **39**: 1321

Li, H., Ruan J. and Durbin, R. (2008) *Genome Res* **18**:1851

Salmon Hillbertz, N. H., Isaksson, M., Karlsson, E. K., Hellmen, E., Pielberg, G. R., Savolainen, P., Wade, C. M., von Euler, H., Gustafson, U., Hedhammar, A., Nilsson, M., Lindblad-Toh, K., Andersson L. and Andersson, G. (2007) *Nat Genet* **39**:1318

Wade, C. M., Giulotto, E., Sigurdsson, S., Zoli, M., Gnerre, S., Imsland, F., Lear, T. L., Adelson, D. L., Bailey, E., Bellone, R. R., Blöcker, H., Distl, O., Edgar, R. C., Garber, M., Leeb, T., Mauceli, E., MacLeod, J. N., Penedo, M. C. T., Raison, J. M., Sharpe, T., Vogel, J., Andersson, L., Antczak, D. F., Biagi, T., Binns, M. M., Chowdhary, B. P., Coleman, S. J., Della Valle, G., Fryc, S., Guérin, G., Hasegawa, T., Hill, E. W., Jurka, J., Kiialainen, A., Lindgren, G., Liu, J., Magnani, E., Mickelson, J. R., Murray, J., Nergadze, S. G., Onofrio, R., Pedroni, S., Piras, M. F., Raudsepp, T., Rocchi, M., Røed, K. H., Ryder, O. A., Searle, S., Skow, L., Swinburne, J. E., Syvänen, A. C., Tozaki, T., Valberg, S. J., Vaudin, M., White, J. R., Zody, M. C.,Broad Institute Genome Sequencing Platform,

Broad Institute Whole Genome Assembly Team, Lander, E. S., and Lindblad-Toh, K. (2009) *(submitted)*

Wiik, A. C., Wade, C., Biagi, T., Ropstad, E. O., Bjerkas, E., Lindblad-Toh, K. and Lingaas, F. (2008) *Genome Res* **18**:1415

Zerbino, D. R. and Birney, E. (2008) *Genome Res* **18**:821

Proc. Assoc. Advmt. Anim. Breed. Genet. 18:239-242

SNP ORIGIN BIAS ON POPULATION STRUCTURE ANALYSIS: AN AUSTRALIAN BEEF CATTLE CASE STUDY

L. R. Porto Neto[1,2] and W. Barendse[1]

Cooperative Research Centre for Beef Genetic Technologies
[1] CSIRO Livestock Industries, Queensland Bioscience Precinct, 306 Carmody Road, St. Lucia 4067, Australia.
[2] The University of Queensland, School of Animal Studies, St. Lucia 4072, Australia.

SUMMARY

The use of single nucleotide polymorphism (SNP) in cattle molecular genetics studies has increased in the last few years by several factors including the identification of new markers and the development of new genotyping technologies. Most of the cattle SNP markers were developed by comparison of a Hereford genome sequence to a sequence of an animal of a different breed, leading to different breed of origin of the SNP markers. In this Australian case study we analysed 302 SNP markers of two different origins (Brahman and Holstein) in a population study including eight cattle breeds. We demonstrate that the breed of origin of the marker can potentially bias this analysis, showing that it is important to find a balance between the origin of the markers and the composition of the population being studied.

INTRODUCTION

Single nucleotide polymorphisms (SNP) are the most common molecular markers in the genome of an organism. Their use has increased through the development of high throughput genotyping platforms that have a low cost per genotype. In cattle the vast majority of SNP markers were identified by comparison of a Hereford genome sequence with sequences of another taurine or Brahman animal (ftp://ftp.hgsc.bcm.tmc.edu/pub/data/Btaurus/). The breed of the DNA sequences, which was compared to the Hereford genome sequence to identify the marker, is called the breed of origin of the SNP marker in this study. A recent study comparing the minor allele frequencies (MAF) of different breeds showed that common SNP (MAF > 0.2) identified originally in taurine breeds often have MAF < 0.10 when evaluated in zebu animals (The Bovine HapMap Consortium, 2009).

Considering that differences in allelic frequency between populations forms the basis of population structure analysis and the knowledge that the SNP origin can bias the allelic frequency, we studied the influence of SNP origin on population structure analysis in a sample of Australian beef cattle.

MATERIAL AND METHODS

The animals used in this study and the genotyping has been reported previously (Barendse *et al.* 2009). There were 179 animals of eight cattle breeds used in this analysis. They were the taurine dairy Holstein (HOL, n=25), taurine meet Hereford (HFD, n=24), Murray Grey (MGY, n=16), Shorthorn (SHN, n=24) and Angus (ANG, n=25); the composite Belmont Red (BEL, n=21) and Santa Gertrudis (SGT, n=25); and the zebu Brahman (BRM, n=19). These animals were genotyped and quality control measures implemented as reported previously. Briefly, 9260 SNP, distributed in all chromosomes, were genotyped using the MegAllele™ Genotyping Bovine 10k SNP Panel (Hardenbol *et al.* 2005) by ParAllele Inc. on an Affymetrics GeneChip. Animals with more than 10% of missing data, and then loci with more than 10% of missing data were excluded from the analysis.

SNP markers. SNP originating in Brahman and Holstein were compared. Holstein were the most common taurine breed that was compared to the Hereford and Brahman was the only zebu breed used in SNP discovery. Among the approximately 300 Brahman markers genotyped only 151 were polymorphic in our population. One hundred and fifty one markers from Holstein were then selected by numeric order, distributed in most chromosomes (Table 1). Only SNP polymorphic in at least one breed were used.

Table 1. Distribution of Brahman (BRM) and Holstein (HOL) markers per cattle chromosome.

Chr	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	X
BRM	9	4	2	2	9	8	3	10	3	7	11	1	4	6	5	9	8	0	4	1	0	7	2	14	0	0	0	1	0	3
HOL	4	5	8	11	4	4	6	5	2	7	7	5	6	5	6	3	1	6	6	2	4	4	4	1	2	4	2	5	1	0

*two Holstein markers and one Brahman marker were not assign to a chromosome.

Population analysis. The minor allelic frequency (MAF) of the markers were observed and grouped into Brahman, composite and taurine subpopulations for comparison. The population stratification was evaluated using the STRUCTURE Software 2.2 (Pritchard *et al.* 2000) and visualized using *Distruct* (Rosenberg 2004). Three runs were performed for each of K=2 to 8 with burn in of 20,000 and 100,000 MC iterations without previous knowledge of breed assignment. The data shown is the analysis of one representative run. A major cluster was considered a cluster that contains more than 50% of the individuals of a breed. We observed the ability of the markers to determine subpopulations (clusters) that were compatible with the breed designations. To determine if a cluster was a good representation of a breed, the number of individuals of pure taurine or Brahman with more than 0.85 genetic composition assigned to the main cluster of their respective breed was counted and the results between the sets of markers were compared.

RESULTS AND DISCUSSION

Allelic frequency analysis. Seventy-one Brahman derived markers had lower (<0.10) MAF in the taurine group than in the Brahmans (23) and composites (21) (Figure 1). A similar trend in the opposite direction was also observed in the Holstein markers, Brahmans having a higher percent of MAF < 0.10 (61) but the difference to the other groups was not as pronounced as for the Brahman markers (Taurines 38 and Composites 44).

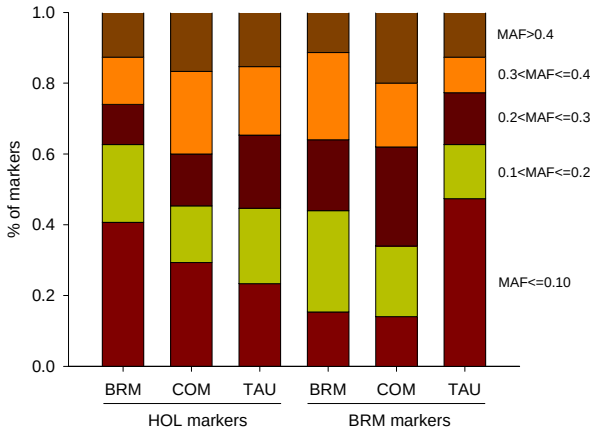


Figure 1. Deciles of minor allele frequency by genetic group and SNP origin

Population stratification analysis. To determine the number of underlying groups in the data, we determine K, the number of subpopulations. The biggest initial increase in likelihood was to K=2 and then a progressively smaller increases until an optimal K=7 for the combined set of markers (data not shown). There was not much difference in likelihood between K of a small number in the range 2-8. We found that with K=2, instead of a taurine vs zebu split, the two groups were meat taurine vs dairy taurine plus zebu, with the Belmont Red and Santa Gertrudis showing near equal contributions for the two groups regardless of SNP origin. For larger K (Fig 2) the groups split progressively into more breeds until with K=8, the Santa Gertrudis and Belmont Red did not only appear as composite taurine and zebu, but as separated groups. Throughout the process, Angus and Murray Grey were indistinguishable, as would be expected given their known history as well as from the principal components analysis for these animals (Barendse *et al.* 2009).

To determine what effect SNP origin had on the STRUCTURE plots we analysed each marker set separately. Most of the discrimination in the combined panel came from the Brahman markers, the increase in likelihood was 5 times greater with the Brahman SNP than the Holstein SNP (data not shown). At K=3, the Brahman panel clearly separated the Holstein from the other beef breeds and at K=8 was able to put most animals correctly into their breeds despite or perhaps because of having more loci with lower MAF (Table 2). The Holstein panel was less discriminatory than the Brahman panel and its analysis led to more than one set of assignments with approximately equal likelihood. The combined set of loci performed as well as the Brahman set, as would be expected given the greater information content in the Brahman SNP. It is worth noting that the Brahman SNP performed worse on Brahman and the Holstein SNP performed worse in classifying Holstein as 100% of that breed group.

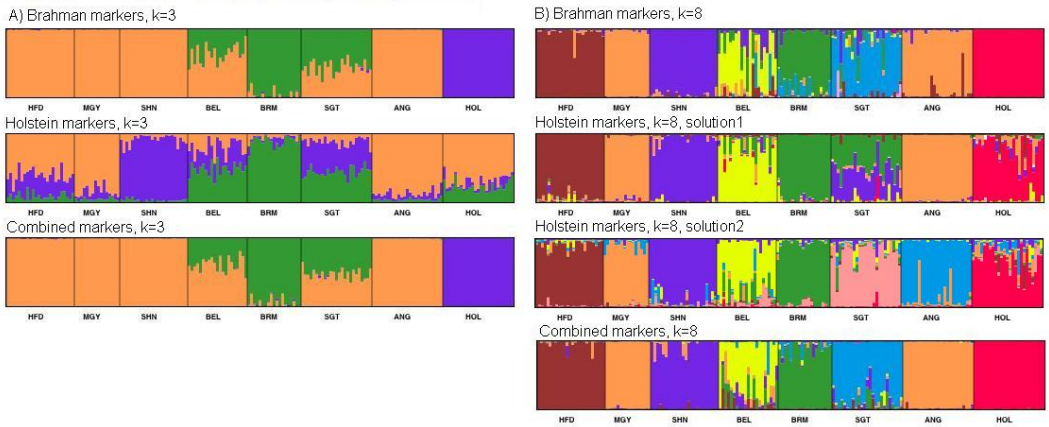


Figure 2. Population structure determined by Brahman, Holstein and Combined set of SNP markers from K=3 (A) and K=8 (B)

The poor performance of Brahman SNP in classifying Brahman and Holstein SNP in classifying Holstein is clearly of interest and suggests why the Holstein SNP have performed worse than the Brahman SNP in the STRUCTURE analysis (Table 2). The Brahman SNP have the highest MAF in the Brahman, as would be expected, and the Holstein SNP have the highest MAF in the Holstein. The Holstein SNP have higher MAF in the taurine breeds, the taurine breeds tend to have similar allele frequencies and so are more difficult to separate from each other using the Holstein panel. On the contrary, the Brahman SNP have greater differences between breeds, they have a more U-shaped distribution, and so they have greater power to discriminate between

breeds. Combining the two set of markers did not materially influence the discriminatory power of the panel of SNP and it is doubtful whether a many more SNP would be able to bring greater clarity in discrimination.

Table 2. Number of individuals of a particular breed that have at least 85% of its genetic composition on its respective breed main cluster at K=8

	N of animals	Brahman markers	Holstein markers[1]		Combined set
		Number of animals in the main cluster (%)	Number of animals in the main cluster (%)		Number of animals in the main cluster (%)
HFD[2]	24	23 (0.96)	19 (0.79)	21 (0.87)	22 (0.92)
MGY	16	16 (1.00)*	9 (0.56)	16 (1.00)*	16 (1.00)*
SHN	24	23 (0.96)	16 (0.67)	21 (0.87)	19 (0.79)
ANG	25	21 (0.84)*	20 (0.80)	25 (1.00)*	25 (1.00)*
HOL	25	25 (1.00)	6 (0.24)	11 (0.44)	25 (1.00)
BRM	19	8 (0.42)	14 (0.74)	18 (0.95)	14 (0.74)

[1] The Holstein markers lead to two different clustering, which are analysed separately.
[2] HFD Hereford, MGY Murray Grey, SHN Shorthorn, ANG Angus, HOL Holstein, BRM Brahman.
* The MGY and the ANG shared the same main cluster.

CONCLUSIONS

A relatively small number of SNP can be used to reconstruct the genetic divisions of breeds within cattle. However, an ascertainment bias in the origin of the SNP would generate spurious conclusions in the degree of similarity between breeds that is not consistent with other molecular work. As the ancestries of cattle are well known, these biases can be seen for what they are. In less well known species, a bias in ascertainment could generate groups that do not reflect the true links between groups.

ACKNOWLEDGMENTS

We thank J.W. Kijas for discussion on using STRUCTURE and *Distruct* analyses. LRPN is supported by an Endeavour International Postgraduate Research Scholarship, a University of Queensland International Student Living Allowance and a Beef CRC scholarship; WB is supported by CSIRO and a Beef CRC Research Grant.

REFERENCES

Barendse W., Harrison B.E., Bunch R.J., Thomas M.B., Turner L.B. (2009) *Bmc Genom* **10:**178.
The Bovine HapMap Consortium (2009) *Science* **324:**528.
Hardenbol P., Yu F.L., Belmont J., MacKenzie J., Bruckner C., Brundage T., Boudreau A., Chow S., Eberle J., Erbilgin A., Falkowski M., Fitzgerald R., Ghose S., Iartchouk O., Jain M., Karlin-Neumann G., Lu X.H., Miao X., Moore B., Moorhead M., Namsaraev E., Pasternak S., Prakash E., Tran K., Wang Z.Y., Jones H.B., Davis R.W., Willis T.D., Gibbs R.A. (2005) *Gen. Res.* **15**:69.
Pritchard J.K., Stephens M. & Donnelly P. (2000) *Genetics* **155:**945.
Rosenberg N.A. (2004) *Mol. Ecol. Notes* **4**:137.

Proc. Assoc. Advmt. Anim. Breed. Genet. 18:243-246

POPULATION STRATIFICATION, NOT GENOTYPE ERROR, CAUSES SOME SNPS TO DEPART FROM HARDY-WEINBERG EQUILIBRIUM

Y.D. Zhang and B. Tier

Cooperative Research Centre for Beef Genetic Technologies
Animal Genetics and Breeding Unit*, University of New England, Armidale NSW 2351

SUMMARY

Large scale whole genome scans generate massive amounts of genotype data. It is essential to check genotype integrity and identify genotype errors prior to association analysis. Departure from Hardy-Weinberg Equilibrium has been adopted as one of the main methods to identify genotype errors. However population stratification also causes departure from Hardy-Weinberg Equilibrium, which is a disadvantage of this approach. This study used 2 sets of SNP genotypes to show that after basic editing using Call Rate and minor allele frequency, up to 13% of SNPs departed from Hardy-Weinberg Equilibrium (HWD) and about one third of these HWD SNPs could be falsely identified as genotype errors, were attributable to population subdivision (eg herd of origin, cohort) for one dataset and corresponding numbers for the second dataset are 21% and 16%, respectively. This approach can avoid improper culling of a considerable proportion of SNPs.

INTRODUCTION

Genotype errors exist in mass generated SNP genotype data. Genotype error may have significantly deleterious effects on genetic tests, such as affecting LD measures and linkage distance in family-design. Such errors may also lead to a high false positive rate, in particular in case-control design (Mitchell *et al.* 2003); for example in the case-control design, difference in allele frequencies at a locus between case and control groups could be interpreted as linkage between this locus and the causal gene.

There are many sources of genotype errors, eg assay failure. Prior to association analysis, it is necessary to identify genotype errors and control them at a certain level. Though some study argued that various combinations of quality control did not reduce much the proportion of false association (Chan *et al.* 2009). Methods proposed to detect genotype errors can be classified into 4 groups: testing Mendelian Inheritance using family based or/and trio data; checking closely linked loci for double recombination events; checking experimental error using duplicates (not useful for systematic genotyping errors); checking errors in population-based or controls of case-control design using Hardy-Weinberg Equilibrium (HWE) test. Deviation from Hardy-Weinberg Equilibrium (HWD) has been widely used for detecting genotype errors (eg Hosking *et al.* 2004). However, besides genotype error, there are a number of other factors causing HWD, such as small population variation and population structure (inbreeding, assortative mating, stratification/admixture). This is particularly relevant to genotypes from livestock, because livestock populations have been subject to decades of selective breeding and breed formation (inbreeding, assortative mating, selection). As illustrated by Hartl and Clark (2007) using Wright's example, that estimated frequencies of the recessive allele for blue flower colour in a population of *Linanthus parryae* in approximately 900 square miles of the Mohave desert exhibited in remarkable geographical subdivision. The average allele frequencies in the East, West and Central regions were 0.515, 0.013 and 0.189, respectively, and the corresponding heterozygosity were 0.50, 0.03 and 0.31. Geographic isolation in the East, West and Central regions, implying

*AGBU is a joint venture of NSW Department of Primary Industries and University of New England.

population subdivision, causes a reduction in heterozygous genotypes, relative to that expected under random mating.

This demonstrates that low levels of heterozygosity and population sub-division can lead to marked deviation from HWE. This implies that some genotype errors are falsely identified by HWE test, because departure from HWE may be actually due to population subdivision. The aim of this study is to describe a method for identifying HWD caused by population structure using SNP genotypes, and demonstrate the method in beef data..

MATERIAL AND METHODS

Reduction in heterozygosity is calculated as the difference between the expected heterozygosity under random mating and that observed in the whole population or subpopulation. The fixation index or Wright's F-statistic is defined as the reduction in heterozygosity expected with random mating at any one level of a population hierarchy relative to another, more inclusive level of the hierarchy (Hartl and Clark 2007). In this assessment, the hierarchical F-statistics is expressed as $F_{ST} = (H_T - H_S) / H_T$, where H_T is heterozygosity of the total population and H_S is the average heterozygosity of subpopulations, for instance the sire groups. Wright (1978) suggested the following interpretations for F_{ST}:

- The range of 0 to 0.05 indicates little genetic differentiation.
- The range of 0.05 to 0.15 indicates moderate genetic differentiation.
- The range of 0.15 to 0.25 indicates great genetic differentiation.
- The value of F above 0.25 indicates very great genetic differentiation.

Modified χ^2 test. In general, Hardy-Weinberg Equilibrium tests were performed using a χ^2 test. This study employed a modified χ^2 test to carry out HWE test. The expected frequencies for three genotypes were adjusted using F_{ST} as the weight factor:

- Aa: $2pq - 2pq\ F_{ST}$
- AA: $p^2 + pq\ F_{ST}$
- aa: $q^2 + pq\ F_{ST}$

Data. SNP genotypes in this assessment were mainly derived from two beef cattle whole genome scan projects (designated as P1 and P2), generated using Affymetrix 10K platform. The SNP genotype data from P1 were the main data which were generated for 579 heifers on 9065 SNPs. In P1, animals can be further classified into subclasses by herd of origin, cohort and sire group (half-sib family). In P2, 9421 SNPs were genotyped for 191 animals derived from 7 breeds. The P2 SNP genotypes were used in this assessment for comparison purposes. Three parameters were available to assess the integrity of SNP data: SNP Call Rate (CR, an indication of genotype completeness, at a scale of 0 to 100%), minor allele frequency (MAF) and deviation from Hardy-Weinberg Equilibrium (HWE). The SNP genotypes were initially edited against MAF and CR. The empirical culling thresholds for CR and MAF were suggested as >93% and >0.05. In 9065 SNPs of the P1 animals, 1043 SNPs showed significant departure from HWE ($p<0.05$) and 1740 out of 9241 SNPs in P2 dataset. In this process, SNPs that showed departure from HWE were used in this subdivision test. The subdivision tests were applied to sire groups, herd of origin and cohort for the P1 dataset and to breed for the P2 dataset.

RESULTS AND DISCUSSIONS

Individual SNP Call Rate is the quality control indicator for an experimental assay. For the P1 dataset (Table 1), after applying CR>93%, there were 8716 SNPs remaining, and 831 of 8716 SNPs departed from HWE. When SNPs were culled with CR>93% and MAF>0.05, 5678 SNPs

remained and 751 (13.5%) of them departed from HWE ($p<0.05$). The majority of the HWD SNPs (91%) showed high reduction of heterozygosity (0.15).

Two examples illustrated in Table 2. Although having high Call Rate (99.8 and 96.0) and moderate MAF (0.14 and 0.33), SNPs A and B showed significant departure from HWE and high reduction in heterozygosity. As indicated by their F-statistics, their HWD were clearly due to subdivision caused by sire. Both SNPs showed reduction in heterozygosity (0.34 and 0.47).

Table 1. Distribution of SNPs in P1 dataset after culling against Call Rate (CR), Minor Allele Frequency (MAF) and Hardy-Weinberg Equilibrium test (χ^2 $p<0.05$). The hierarchical F-statistics were assessed against sire group, cohort and herd of origin. The modified Chi Square test showed that the departure from HWE of about 30% of SNPs was due to subdivision of sire group or 10% due to cohort or herd of origin (χ^2_F $p>0.05$)

CR	MAF	Total	HWD χ^2 ($p<0.05$)	Modified χ^2_F ($p>0.05$)
				Sire Group
0.0	0.0	9065	1043	355
93	0.0	8716	831	300
0.0	0.05	5908	930	296
93	0.05	5678	751	262
				Cohort
0.0	0.0	9065	1043	119
93	0.0	8716	831	106
0.0	0.05	5908	930	85
93	0.05	5678	751	79
				Herd of Origin
0.0	0.0	9065	1043	111
93	0.0	8716	831	100
0.0	0.05	5908	930	80
93	0.05	5678	751	76

Table 2. Examples of departure from Hardy-Weinberg Equilibration due to subdivision, illustrated using 2 SNPs with high Call Rate and moderate MAF

	Genotype 0	Genotype 1	Genotype 2	χ^2	P
SNP A					
Observed	37	97	461		
Expected	12.3	146.4	436.3	23.3	0.0001
Expected, F_{ST}(sire)=0.34	37.3	96.3	461.3	0.004	0.998
SNP B					
Observed	323	133	177		
Expected	264.8	249.5	58.8	60.53	0.000001
Expected, F_{ST}(sire)=0.34	311.2	156.7	105.2	2.78	0.24

The hierarchical F-statistics values for each SNP were assessed against sire groups, cohort or herd of origin (as shown in Table 1). Using the hierarchical F-statistics derived against sire groups in the modified χ^2 test on the 831 HWD SNPs, 531 showed significant departure from HWE ($p<0.05$), *ie.* 300 SNPs were not eliminated due to HWD ($p<0.05$). The modified Chi Square test (χ^2_F) showed that about one-third of SNPs that departed from HWE were attributable to subdivision caused by sire groups. When using the F-statistics derived from cohort or herd of

origin for P1 SNPs, an additional 106 or 100 SNPs remained, respectively. Collectively, 357 of 831 SNPs were retained. As a result, after applying HWE test using the hierarchical F-statistics 374 of 8716 SNPs were culled due to departure from HWE ($p<0.05$). This result suggested that about 10% of SNPs departed from HWE due to subdivision by cohort or herd of origin.

On examination of the P2 SNPs, the F-statistics were estimated within breed. After culling on CR and MAF 7461 SNPs remained (Table 3), 1536 of the remained SNPs (21%) departed from HWE ($p<0.05$). Similarly, the majority of these SNPs showed reduction in heterozygosity. After applied with the modified χ^2 test, 1233 of 1536 SNPs (80%) were in Hardy Weinberg Equilibrium (χ^2_F $p>0.05$).

Table 3. Distribution of SNP genotypes for the P2 animals after culling against Call Rate (CR), Minor Allele Frequency (MAF) and Hardy-Weinberg Equilibrium test (χ^2 $p<0.05$). The hierarchical F-statistics was assessed against breed. The modified χ^2 (χ^2_F) showed about 80% of SNPs departed from HWE were due to subdivision of breed (χ^2_F $p>0.05$)

CR	MAF	Total	HWD χ^2 ($p<0.05$)	F-statistics χ^2_F ($p> 0.05$)
0.0	0.0	9241	1740	1353
93	0.0	8634	1562	1242
0.0	0.05	7944	1709	1344
93	0.05	7461	1536	1233

CONCLUSIONS

The HWE test can be used to detect genotype errors. However, in populations with some sub-structure, steps should be taken to identify possible sources underlying the HWD other than genotype errors. Possible sources of subdivision could be natural grouping, management process etc. We have demonstrated that application of the modified χ^2 test using the hierarchical F-statistics can identify some SNPs with HWD due to subpopulation. In this assessment for the P1 dataset, subdivision was assessed against sire group, cohort and herd of origin. Sire group is the main source causing subdivision. Collectively, about one-third of SNPs showing HWD can be corrected by accounting for sub classification and should be attempted when data are analysed for trait-genotype associations. In the P2 dataset, when genotypes from 7 breeds were pooled as population data, significant subdivision due to breed was apparent. When HWE test is applied to this dataset, about 16% of SNPs (1233 out of 7461) could be wrongly culled because of their departure from HWE.

ACKNOWLEDGMENTS

The work is part of the Cooperative Research Centre for Beef Genetic Technologies and data was made available from the Cooperative Research Centre for Beef Genetic Technologies.

REFERENCES

Chan, E. K. F., Hawken, R. and Reverter, A. (2009) *Animal Genetics* **40**:149.

Hartl, D. L. and Clark, A. G. (2007) "Principles of Population Genetics" 5th ed. Sinauer Associates, Sunderland, MA, USA.

Hosking, L., Lumsden, S., Lewis, K., Yeo, A., McCarthy, L., Bansal, A., Riley, J., Purvis, I. and Xu, C. F. (2004) *Eur. J. Hum. Genet.* **12**:395.

Mitchell, A. A., Cutler, D. J. and Chakravarti, A. (2003) *Am. J. Hum. Genet.* **72**(3):598.

Wright, S. (1978) "Evolution and the Genetics of Populations", Vol. 4: Variability within and among Natural Populations. University of Chicago Press, Chicago.

Proc. Assoc. Advmt. Anim. Breed. Genet. 18:247-250

GENOME STRUCTURE IN AUSTRALIAN HOLSTEIN FRIESIAN CATTLE REVEALED BY COMBINED ANALYSIS OF THREE HIGH DENSITY SNP PANELS

M.S. Khatkar[1,2], B. Tier[1,4], M. Hobbs[1,2], D. Khatkar[1,2], J.A.L. Cavanagh[2], R. Crump[1,4], G. Moser[1,3] and H. W. Raadsma[1,2]

[1]The CRC for Innovative Dairy Products, [2]ReproGen, Faculty of Veterinary Science, The University of Sydney, Camden, Australia, [3]Bellbowrie, QLD 4070, Australia, [4] Animal Breeding and Genetics Unit, UNE, Armidale, NSW 2351, Australia.

SUMMARY

We genotyped overlapping samples of Australian dairy bulls using three different SNP chips (15k, 25k and 54k). These chips have different but complementary coverage hence increasing the number of animals and the density and coverage of SNPs to 74k in a combined dataset. A combined analysis of the data from these three SNP chips showed a four fold increase in the coverage of the genome by haplotype blocks over bovine hapmap reported previously (Khatkar *et al.* 2007). An analysis of contiguous runs of homozygosity revealed long stretches (up to 49.39 Mb) of homozygosity on chromosome 1 in many bulls. Distribution of these segments of homozygosity in a sample of bulls is presented. The results for one chromosome are described in detail.

INTRODUCTION

The success of genomic selection (GS) and genome wide association studies (GWAS) largely depends on the number of animals in the design and density of markers in the genome screen. The latter depends on the genome structure in terms of the extent and variation in linkage disequilibrium in the population. With the availability of high density and low cost SNP chips, it is now possible to genotype large numbers of animals with high density SNP markers. Such initial datasets generated in cattle have enabled the elucidation of haplotype block structure (Khatkar *et al.* 2007), selection signatures (Hayes *et al.* 2008), extent of homozygosity (MacLeod *et al.* 2007) and the extent of linkage disequilibrium (McKay *et al.* 2007; Khatkar *et al.* 2008; Sargolzaei *et al.* 2008), in the bovine genome. These studies recommended the requirement of higher density SNP markers for association mapping and genome structure analysis. Due to sparse and unequal SNP coverage it was not possible to explore fine scale LD in the bovine genome.

Recently, there has been a significant and rapid increase in density of SNPs available on diverse sources of bovine SNP genotyping arrays. We genotyped three different SNP chips in overlapping samples of Australian dairy bulls. This combined dataset provides higher density SNP coverage compared to any individual SNP chip presently available. Here we present the results from the analysis of this combined dataset with detailed results focused on chromosome 1.

MATERIALS AND METHODS

Genotypic data. Data from three SNP chips viz. 15k, 25k and 54k genotyped for 1536, 441 and 377 Australian Holstein-Friesian (HF) bulls, respectively, were combined into a single dataset. This combined dataset has 73,569 unique SNPs and 1,943 bulls. After excluding the SNPs with low (<0.05) minor allelic frequency (MAF) and deviation (P<0.0001) from Hardy Weinberg Equilibrium (HWE), 3,102 SNPs on chromosome 1 (BTA1) were used for the present analysis.

Construction of Haplotype block map. Haplotype blocks were identified using Haploview software (Barrett *et al.* 2005) and are defined as detailed in our previous bovine HapMap study (Khatkar *et al.* 2007).

Detection of runs of homozygosity (ROH). A run of homozygosity was defined as a contiguous segment of homozygosity The ROH was detected across the chromosome by inspecting the genotype of an animal starting from the beginning of the chromosome. The segment ended when a heterozygous SNP was encountered and new segment was started at the next homozygous SNP genotype. The only segments consisting of a minimum 6 SNPs and a minimum length of 100 kb were counted as runs of homozygosity. These thresholds were applied to filter out numerous, small segments resulting from local LD.

RESULTS AND DISCUSSION

Haplotype blocks. A total of 3,102 SNPs located on BTA1 were used to construct the haplotype block map of chromosome 1. The individual SNPs are well distributed across the range of MAF (0.05-0.5) (Figure 1a). The mean spacing between SNPs is 51.9 ± 9.7 kb. Figure 1b shows the distribution of spacing between adjacent SNPs which is fairly even. There are only 3 gaps of more than 500 kb between consecutive SNPs. Figure 2 presents the haplotype block map of chromosome 1. The map contains 298 haplotype blocks consisting of more than 2 SNPs. The mean length of these haplotype blocks is 113.4± 9.18 kb. Figure 1c shows that most blocks are relatively short in length (< 100 kb). These blocks cover 11.3 % of BTA1. This is a four fold increase in the coverage as compared to the earlier bovine HapMap of BTA1 constructed based on 528 SNPs (Khatkar *et al.* 2007)). Pair-wise tag analysis selected 81 % of the 3,102 SNPs as tag SNPs, suggesting limited redundancy, at this density of SNP coverage.

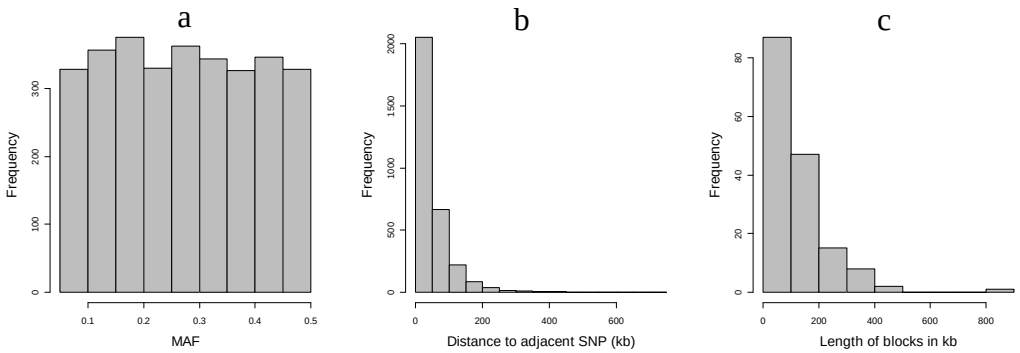


Figure 1. Distribution of a) MAF of 3102 SNPs b) SNP spacing c) length of haplotype blocks on bovine chromosome 1.

Runs of homozygosity. On an average there were 73.2 ± 0.58 (range from 32 to 119) runs of homozygosity on chromosome 1, when a threshold of a minimum length of 100 kb and at least 6 SNPs in a segment was applied. These ROH, on average, cover 28.3 % of the chromosome. The number and coverage of ROH, applying different thresholds for minimum segment length, are presented in Table 1. Increasing the minimum length of segment from 500kb to 5Mb showed a rapid decline in mean number of ROH across the chromosome suggesting that longer ROH occur with low frequency. The longest ROH on chromosome 1 is 49.4 Mb spanning across 861 SNPs. In summary, there is high level of homozygosity in Australian HF bulls. Figure 3 shows the variation

in the proportion of bulls, carrying ROH longer than 500kb, along the chromosome 1 with highest peak in the region ranging from 84Mb to 87 Mb. The figure clearly shows a non-uniform distribution of regions of homozygosity and regions showing strong signs of loss of heterozygosity. To what extend this is a function of population structure and/or influence of selection, remains unclear.

Figure 2. Heatmap of haplotype structure of BTA1 showing confidence bounds of D'. Dark grey colour indicates strong evidence of LD, light grey uninformative and white suggests strong evidence of recombination.

Table 1. Regions of Homozygosity on chromosome 1 in 377 HF bulls.

Threshold for minimum length	Number of contiguous homozygous segments				% of chromosome			
	Mean	SE	min	Max	mean	SE	min	max
100 Kb	73.2	0.58	32	119	28.3	0.42	11	66
250 Kb	55.9	0.46	26	99	26.3	0.42	10.3	64.2
500 Kb	25.9	0.3	11	56	19.5	0.45	8.5	60.3
1 Mb	7	0.14	2	19	11.6	0.49	1.7	56.4
2 Mb	2	0.08	0	10	7.6	0.5	0	51.3
5 Mb	0.6	0.05	0	5	5	0.48	0	51.3

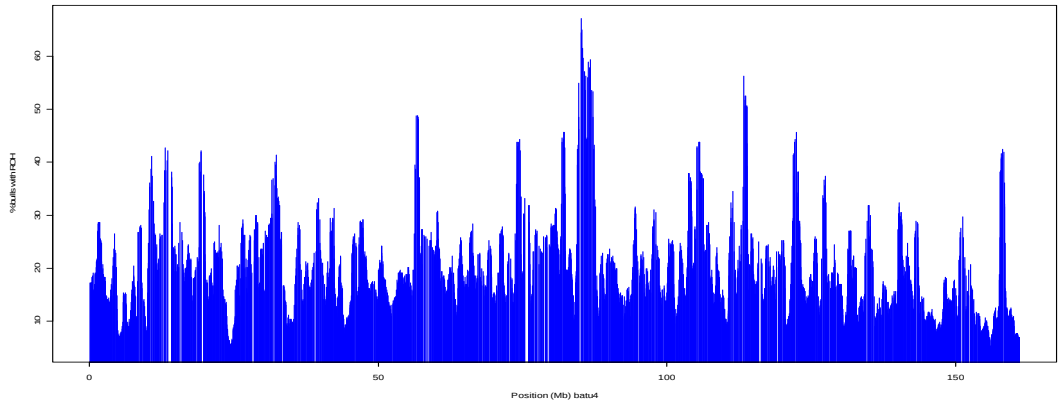


Figure 3. Percentage of bulls carrying runs of homozygosity larger than 500 Kb counted at different points along the chromosome 1.

Animal genomes

Recently, large numbers of structural variations (SVs) including insertion-deletions, , translocations/inversions and copy number variations have been discovered in normal human individuals (Redon *et al.* 2006; Frazer *et al.* 2009). Many of these variations were found associated with phenotypic differences and disease status (Smith 2009). The SNP panels were unable to tag all of the known SVs. It is expected that these types of structural variations will be present in the bovine genome as well. There were 5 % of SNPs with MAF above 0.01, which deviated from HWE in the present study. Some of these deviations from HWE may arise from difficulties in calling the genotype due to the structural variations mentioned above. It is possible to use SNP genotypic data to discover some of structural variations (Simon-Sanchez *et al.* 2007), but it requires very high density SNP panels. However, with the availability of cheaper high throughput whole genome sequencing technology (Kidd *et al.* 2008), it will be possible to discover these structural variations even in livestock species. Studying these new types of DNA variants along with high density SNPs will provide deeper insight into the complex relationship of genotype with phenotype.

ACKNOWLEDGMENTS

The genotype data was provided by bovine genome resources from the Co-operative Research Centre for Innovative Dairy Products.

REFERENCES

Barrett, J.C., Fry, B., Maller, J. and Daly, M.J. (2005) *Bioinformatics* **21:**263

Frazer, K.A., Murray, S.S., Schork, N.J. and Topol, E.J. (2009) *Nat Rev Genet* **10**:241.

Hayes, B.J., Lien, S., Nilsen, H., Olsen, H.G., Berg, P., Maceachern, S., Potter, S. and Meuwissen, T.H. (2008) *Anim Genet* **39:**105.

Khatkar, M.S., Nicholas, F.W., Collins, A.R., Zenger, K.R., Cavanagh, J.A., Barris, W., Schnabel, R.D., Taylor, J.F. and Raadsma, H.W. (2008) *BMC Genomics* **9:**187.

Khatkar, M.S., Zenger, K.R., Hobbs, M., Hawken, R.J., Cavanagh, J.A., Barris, W., McClintock, A.E., McClintock, S., Thomson, P.C., Tier, B., Nicholas, F.W. and Raadsma, H.W. (2007) *Genetics* **176:**763.

Kidd, J.M., Cooper, G.M., Donahue, W.F., Hayden, H.S., Sampas, N., Graves, T., Hansen, N., Teague, B., Alkan, C., Antonacci, F., Haugen, E., Zerr, T., Yamada, N.A., Tsang, P., Newman, T.L., Tuzun, E., Cheng, Z., Ebling, H.M., Tusneem, N., David, R., et al. (2008) *Nature* **453**:56

McKay, S.D., Schnabel, R.D., Murdoch, B.M., Matukumalli, L.K., Aerts, J., Coppieters, W., Crews, D., Dias Neto, E., Gill, C.A., Gao, C., Mannen, H., Stothard, P., Wang, Z., Van Tassell, C.P., Williams, J.L., Taylor, J.F. and Moore, S.S. (2007) *BMC Genet* **8**:74.

MacLeod, I.M., Hayes, B.J., Chamberlain, A.J., Savin, K.W., McPartlan,H.,Haile-Mariam, M., Bowman, P. and Goddard, M.E.(2007) *Proc. Assoc. Advmt. Anim. Breed. Genet.* **17:** 549.

Redon, R., Ishikawa, S., Fitch, K.R., Feuk, L., Perry, G.H., Andrews, T.D., Fiegler, H., Shapero, M.H., Carson, A.R., Chen, W., Cho, E.K., Dallaire, S., Freeman, J.L., Gonzalez, J.R., Gratacos, M., Huang, J., Kalaitzopoulos, D., Komura, D., MacDonald, J.R., Marshall, C.R., et al. (2006) *Nature* **444**:444

Sargolzaei, M., Schenkel, F.S., Jansen, G.B. and Schaeffer, L.R. (2008). J Dairy Sci 91. 2106-17.

Simon-Sanchez, J., Scholz, S., Fung, H.C., Matarin, M., Hernandez, D., Gibbs, J.R., Britton, A., de Vrieze, F.W., Peckham, E., Gwinn-Hardy, K., Crawley, A., Keen, J.C., Nash, J., Borgaonkar, D., Hardy, J. and Singleton, A. (2007) Hum Mol Genet **16:**1

Smith, M. (2009). *Ann N Y Acad Sci* **1151:**1

Proc. Assoc. Advmt. Anim. Breed. Genet. 18:251-254

A PECULIARITY OF GENE FREQUENCY ESTIMATION

J.W. James[1], Vicki A. Whan[2] and Belinda J. Norris[2]

[1]Faculty of Veterinary Science, University of Sydney, NSW 2006
[2] CSIRO Livestock Industries, St Lucia, Queensland 4067

SUMMARY

Black wool is due to a recessive gene, the dominant white trait being caused by a duplication, with one or more repeats. The total number of repeats in the diploid genotype can be detected, but a 1/1 and a 2/0 repeat cannot be distinguished. This complicates gene frequency estimation, and in the absence of black sheep, it is shown that very similar data on repeat numbers can give drastically different gene frequency estimates. The (possible) presence of black sheep can resolve the problem.

INTRODUCTION

The presence of black sheep in a flock is highly undesirable, and if such animals appear they are usually culled. Black wool has long been known to be due to a recessive gene, with heterozygotes for the recessive allele indistinguishable from the dominant white homozygotes. This has made it very difficult to eliminate the recessive allele, as only the production of black progeny indicates that a ram is a carrier. However, the molecular basis of the trait has recently been shown by Norris and Whan (2008) to be a duplication at the ovine agouti locus, and a test has been developed. Unfortunately, there is more than one dominant allele, with some having more than one repeat. The test counts the number of repeats in the diploid genotype and thus cannot distinguish between a homozygous dominant 1/1 and a heterozygous 2/0. The first genotype would produce all white progeny, but the second could produce black offspring, and half of all progeny would be carriers of the recessive allele. The practical value of the test clearly depends on the frequencies of the various alleles in the population to which it is applied. Samples of white sheep were therefore taken and tested. In the course of analysis of these data a peculiarity of gene frequency estimation was found, and this is reported in this paper.

DATA and ANALYSIS

For the present purpose only a part of the data is relevant, so other data collected are not mentioned. From the Falkiner Field Station research flock, 60 White Suffolk and 76 Poll Dorset ewes were sampled and tested for number of repeats. The results are shown in Table 1.

Table 1. Distributions of number of repeats in White Suffolk and Poll Dorset ewes.

Breed	Number of ewes	Number of repeats					
		1	2	3	4	5	6
WS	60	2	23	18	11	5	1
PD	76	2	31	20	11	9	3

The two distributions look quite similar, and a chi-squared test gave a value of 1.56 on 5 degrees of freedom with a probability of about 0.9, confirming the apparent similarity.

If we assume that there are four alleles present with 0, 1, 2 and 3 repeats, with gene frequencies p_0, p_1, p_2 and p_3, and that the parents mated at random, the expected proportions of the six phenotypes are:

$f_1 = 2p_0p_1/(1 - p_0^2)$, $f_2 = (p_1^2 + 2p_0p_2)/(1 - p_0^2)$, $f_3 = (2p_1p_2 + 2p_0p_3)/(1 - p_0^2)$, $f_4 = (p_2^2 + 2p_1p_3)/(1 - p_0^2)$, $f_5 = 2p_2p_3/(1 - p_0^2)$, $f_6 = p_3^2/(1 - p_0^2)$. (1)

These frequencies have been calculated assuming a Hardy-Weinberg distribution at birth, with any black lambs being culled.

Letting n_j denote the number of ewes with j repeats, the logarithm of the likelihood function is

$L = n_1 \log f_1 + n_2 \log f_2 + n_3 \log f_3 + n_4 \log f_4 + n_5 \log f_5 + n_6 \log f_6$. (2)

Maximising the likelihood for general values of f_j gives estimated frequencies as n_j/N, where N is the total number of animals. Assuming the values of f_j satisfy the relations given in (1), finding the maximum likelihood estimates and comparing the resulting maximum of the likelihood with the general maximum gives a test for the adequacy of the assumed model.

An analytical approach to finding the maximum with the assumed model is not rewarding, so a Monte Carlo method was used. The method was to sample 4 values from a uniform distribution on (0, 1), then to divide each by their sum to give 4 gene frequencies adding to unity. These were then used in the equations (1) to find the f_j which were inserted in (2) to find the corresponding likelihood. This was replicated 10^8 times, and the largest value of L and its associated values of f_j and gene frequencies were taken as the maximum likelihood and the ML estimates. This is not necessarily computationally efficient, but is quick to program, and each run takes a couple of minutes. A chi-squared was computed from the difference in likelihoods for the general and assumed models, with 2 degrees of freedom as 3 independent parameters were estimated for the assumed model. In addition, the expected numbers were calculated from the estimated f_j and tested against the observed numbers by chi-squared. The two methods of computing chi-squared agreed very well. As an additional test of the model, an additional allele with 4 repeats was added, and the analysis repeated, but the outcome was that the estimated frequency of the extra allele was negligible, and the likelihood was essentially unchanged, so the results are not given here. To compute standard errors, a simplex was constructed starting from the ML estimates, and the method of Nelder and Mead (1965) was used.

RESULTS

The maximum likelihood estimates of gene frequencies in the two samples are shown in Table 2.

Table 2. Maximun likelihood estimates (standard errors) of gene frequencies in White Suffolk and Poll Dorset ewes, and chi-squared values for testing the assumed model.

Breed	p_0	p_1	p_2	p_3	Chi-squared
WS	.0276 (.0336)	.5970 (.0798)	.2487 (.0815)	.1267 (.0501)	0.67
PD	.5125 (.1007)	.0201 (.0271)	.2998 (.0713)	.1676 (.0437)	0.37

Clearly there is no need to reject the assumed model. However, the great discrepancy in estimates of p_0 and p_1 is very surprising, since the distributions do not differ significantly. The Poll Dorset result seems to be discrepant, since it implies that about one quarter of Poll Dorset lambs would be born black, which is not the case. It appears that in the absence of black lambs, the frequencies of the recessive and single repeat alleles can compensate for each other. This can be illustrated by the fact that if the estimates for White Suffolk are used to compute expected numbers for Poll Dorset and vice versa, the chi-squared values for comparing the observed and expected values are 8.77 and 4.05 respectively on 5 degrees of freedom, with probabilities of

approximately 0.15 and 0.5. Thus, although the ML estimates fit the data most closely, the ML estimates from the other breed do not fit significantly worse. In this case it was not hard to see that the Poll Dorset ML estimate was misleading, but this will not always be the case.

DISCUSSION

The data used here were obtained from flocks of white ewes, but if the true frequency of the recessive allele is about 0.02 it is very likely that these sheep came from a population in which no black lambs were actually culled. If we therefore assume that in fact there was an observed value of zero for ewes with no repeats, we have an expected proportion of $f_0 = p_0^2$, while the f_j values in (1) are all multiplied by $(1 - p_0^2)$. With these modified data we have estimated gene frequencies using the method described above with the obvious changes. The results are shown in Table 3.

Table 3. Maximum likelihood estimates of gene frequencies (standard errors) for White Suffolk and Poll Dorset ewes with data augmented with zero homozygous recessives.

Breed	p_0	p_1	p_2	p_3
White Suffolk	.0275 (.0324)	.5974 (.0790)	.2476 (.0810)	.1275 (.0503)
Poll Dorset	.0221 (.0269)	.5954 (.0660)	.2230 (.0602)	.1594 (.0446)

The chi-squared tests gave values of 6.30 and 0.74 respectively on 3 degrees of freedom, so that the augmented data are satisfactorily fitted. It is striking that by forcing the estimates in the Poll Dorset to give low expected numbers of homozygous recessives the analysis has resulted in a switch between p_0 and p_1. On the other hand, the White Suffolk estimates from the augmented data are nearly identical to those from the actual data.

As an approach to obtaining reasonable estimates for the Poll Dorset without augmenting the data we have modified the estimation program to find the maximum of the likelihood over the other gene frequencies for fixed values of p_0 and the resulting values of chi-squared comparing this likelihood with the unrestricted one are shown in figure 1. The likelihood curve plotted against fixed values of p_0 has two maxima for both the Poll Dorset and the White Suffolk, one near 0.02 and one near 0.5. As can be seen, the two cases show similar curves, but the lower of the two minimum chi-squareds is different in the two breeds. Obviously, the variation in the samples has had the effect of shifting the curve for the Poll Dorset so that the "wrong" likelihood is globally maximum.

Clearly in this case the information in the data does not allow the gene frequencies to be found with great precision, nor does it allow the two models, with or without an "observed" zero for the number of recessives, to be definitively distinguished. However, the fact that both breeds show two maxima for the likelihood, and that the global maximum may be incorrect in some cases, even when the sample distributions do not differ significantly, is a warning to be careful when analysing data of this kind.

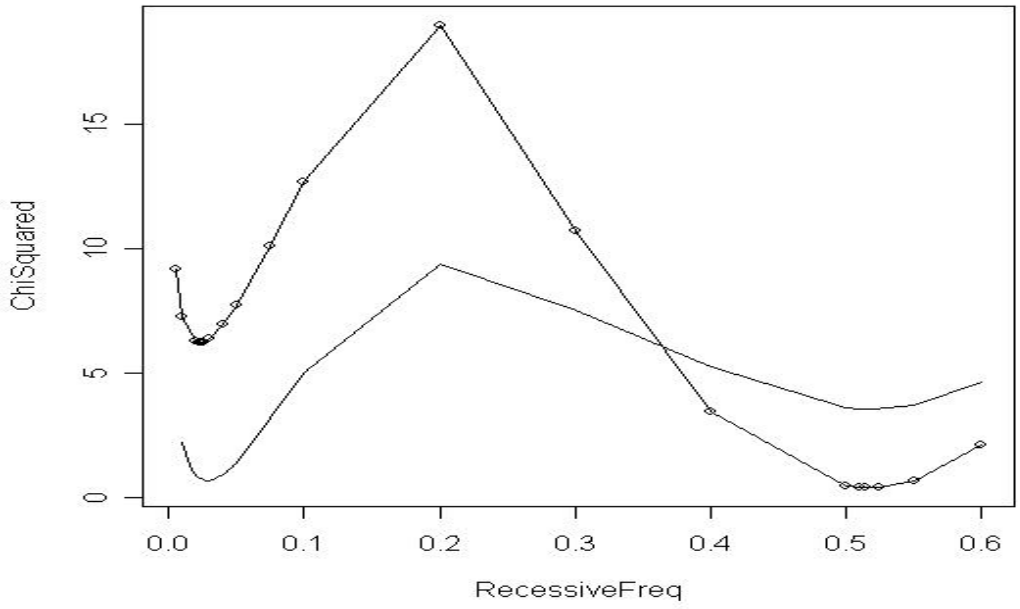


Figure 1. Chi-squared values from unrestricted maximum likelihood versus maximum likelihood for different set values of p_0 in Poll Dorset (with dots) and White Suffolk (no dots).

The main lesson to be learnt from this is that even estimates obtained by highly regarded methods such as maximum likelihood can be seriously wrong, and need to be regarded with scepticism if they do not make sense. Another lesson is that the absence of a class from the data can have important consequences for the analysis. An obvious possible situation is with a recessive lethal. A third lesson is that gene frequency estimation is always harder when the genes present in the sample cannot be counted.

ACKNOWLEDGMENTS

Professor Frank Nicholas was responsible for JWJ's involvement in this work. We thank a referee for suggesting the calculation of standard errors.

REFERENCES

Nelder, J.A. and Mead, R. (1965) A simplex method for function minimization. *The Computer Journal* **7**:308.

Norris, B.J. and Whan, V.A. (2008) A gene duplication affecting expression of the ovine ASIP gene is responsible for white and black sheep. *Genome Research* **18**:1282.

Proc. Assoc. Advmt. Anim. Breed. Genet. 18:255-262

INTERPRETATION AND PREDICTION IN DISEASE GENETICS

S.C. Bishop

Roslin Institute and R(D)SVS, University of Edinburgh, Roslin, Midlothian EH25 9PS, UK

SUMMARY

This paper summarises general issues relating to the interpretation of field disease data and the prediction of responses to selection for disease resistance. Prioritisation of diseases to study is a non-trivial task as there are many criteria by which disease importance can be assessed; a structured approach is described herein. The interpretation of field data and the prediction of responses to selection require an understanding of disease epidemiology, as infection transmission dynamics will affect both the interpretation of genetic parameters and the likely observable outcomes. For many bacterial and viral diseases, field data will likely contain noise due to incomplete exposure to infection and imperfect diagnosis of infection. These factors will result in heritabilities or SNP associations being underestimated; hence a weak genetic signal from such data may mask stronger underlying genetic effects. Interpretation of field data for parasite resistance is more straightforward, provided animals have faced sufficient challenge. Mathematical models that predict responses to selection reveal non-linear relations between mean host genotype and outcomes such as observed infection levels, animal performance, and the likelihood and severity of epidemics. In many cases, total benefits are predicted to be larger than suggested by quantitative genetic theory alone, justifying inclusion of disease resistance in selection goals. These concepts are illustrated for nematode infections and footrot in sheep, and ruminant mastitis.

INTRODUCTION

Genetic variation in host resistance to infectious disease is ubiquitous, and disease resistance is now a major focus for animal geneticists. Further, the availability of dense single nucleotide polymorphism arrays (i.e. SNP chips) has given rise to hitherto unforeseen opportunities to dissect this between-host variation and identify genes contributing to this variation, by means of genome wide association studies. This, coupled with more traditional quantitative genetic variance-partitioning approaches, enables detailed descriptions of genetic aspects of disease resistance and the identification of individuals with extreme (high or low) risk of infection or disease.

It is usually necessary to use field data to get sufficient animals to reliably estimate genetic parameters for disease resistance or detect robust SNP associations. Data may be captured from a population undergoing an epidemic such as bovine tuberculosis, from an endemic disease such as mastitis, where herd-level disease occurrence is predictable, or from diseases such as nematode parasite infections where (depending on weather conditions) all animals will be challenged. However, such field data is very 'noisy': diagnosis of infection or disease may be imprecise; it can be difficult to determine when infection of an animal occurred; and it is often unclear whether or not apparently healthy animals have been exposed to the infection. This paper explores the nature of field disease data, assessing the opportunities and describing factors affecting its interpretation and analysis. Further, it explores predictions of the benefits of selecting animals for increased disease resistance, accounting for disease epidemiology and transmission of infection.

PRIORITISING DISEASES

There are many examples of host genetic variation in resistance to infection or in disease resistance, i.e. the side effects of infection, with even cursory literature reviews revealing more than 50 diseases for which genetic variation has been published (Bishop, 2005). It is a reasonable assumption that genetic variation will exist for resistance to most diseases and, therefore, most

diseases are potentially targets for selective breeding of the host. In reality, only a small subset of diseases can ever be considered and prioritisation must be made. Davies et al. (2009) discuss and demonstrate criteria by which diseases can be ranked. These include economic impact, industrial concern, public concern, human health implications, animal welfare and international trade restrictions caused by the disease. Assigning scores to each category and summary across categories gives an empirical ranking of diseases. Although the ranking will differ between countries or production systems, it will tend to throw obvious suspects such as foot and mouth disease (FMD) to the top of the list. But from an animal genetics perspective, other factors such as feasibility of data collection, prior evidence for the extent of host genetic variation in resistance and the available genomic tools also influence the rankings. The combined rankings of Davies et al. (2009) suggest that for host genetic studies, the highest ranking diseases are salmonella infections, Marek's disease and coccidiosis in chickens, mastitis in cattle, *E. coli* infections and porcine reproductive and respiratory syndrome in pigs, and mastitis and gastrointestinal (GI) parasites in sheep. Whilst these are the most amenable diseases for genetic research, from a breeding perspective it is also important to consider the logistics of performing selection, the epidemiological benefits of selection and the compatibility with other control measures. If a disease meets these criteria, then is it a suitable candidate for breeding for host resistance.

FRAMEWORK FOR INTERPRETING FIELD DATA

Transmission of infection. The key to interpreting disease data and predicting responses to selection lies in understanding the infection transmission pathways. A schema for typical infections is shown in Fig. 1. Infection is transmitted from animals that are infected and infectious (i.e. capable of transmitting infection) to animals that are immunologically susceptible, i.e. capable of being infected. For some diseases infection may be transmitted via a reservoir, e.g. pasture or insect vectors. Animals advance through disease states according to the disease and the force of infection. For macroparasitic diseases, such as GI parasite infections in grazing ruminants, the infection pressure is essentially continuous and, depending on climate and treatment strategies, most animals are challenged almost continually. Thus, a population will mainly comprise animals in the *Infectious* state, and the *Recovered* state generally is not relevant. For microparasitic infections (bacteria or viruses), animals progress through some or most of the infection states in Fig. 1, and it may be difficult to quantify the extent to which animals are infected. A population will comprise animals in every state, with the proportion in each state depending on the time since the commencement of the epidemic and the resistance of the population to the disease in question.

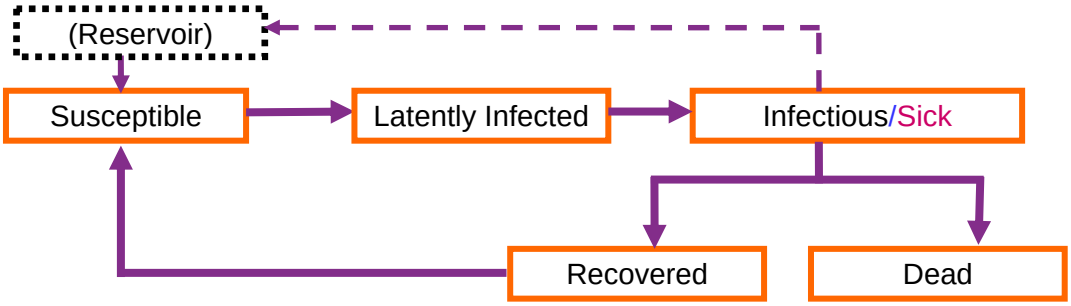


Figure 1. Potential Pathways of Transmission of Infection and Animal Infection States

Interpreting field data: GI nematode parasite infections. Continuous between-animal variation will be seen in indicator traits used to describe nematode resistance. These traits fall into three broad categories: (i) indicators of how heavily infected an animal is, such as faecal egg count (FEC), (ii) indicators of animals' immune responses, such as antibody levels or eosinophilia and (iii) indicators of the impact of infection on the animal, such as fructosamine or pepsinogen concentrations, packed cell volume (PCV) for *Haemonchus contortus* infections, or even growth rate. Provided it can be assumed that animals grazing the same pasture are equally challenged, then data analysis is straightforward: the data are transformed to render them approximately normally distributed, and analysed in the same way as any other trait. Animals with extreme EBVs are selected, i.e. for decreased FEC, or for increased PCV or growth rate.

Interpreting field data: bacterial or viral infections. Data collected will often be categorical, describing the state that an animal is in, e.g. alive vs. dead or diseased vs. apparently healthy. The severity of infection or disease is often not known or may be unobtainable. The apparently healthy category may comprise animals that have yet to be infected, either because they haven't had sufficient exposure or have a high degree of resistance to infection, are latently infected, or have been infected but have subsequently recovered.

The timing and extent of exposure to infection is generally unknown. Consider a population where a proportion e has faced sufficient exposure to pathogen challenge to have a chance of becoming infected. Let the virtual prevalence (p^*) be the prevalence if all animals were exposed to the pathogen. If exposure is random and independent of animal genotype, then observed prevalence is ep^*. Of the $1-ep^*$ proportion of animals that are healthy, $e(1-p^*)$ are exposed and apparently resistant, and $(1-e)$ have not yet been exposed. It can be shown (Bishop and Woolliams, in prep.) that incomplete exposure to infection has a near-linear impact on the heritability of liability to infection, as illustrated in Fig. 2. However, assuming exposure is random across families and genotypes, the impact on EBV estimation and hence genetic progress is somewhat less (Nieuwhof, 2009), because the incomplete exposure effects average out at the family level. Effects of non-random exposure, e.g. mother-offspring transmission, have yet to be explored.

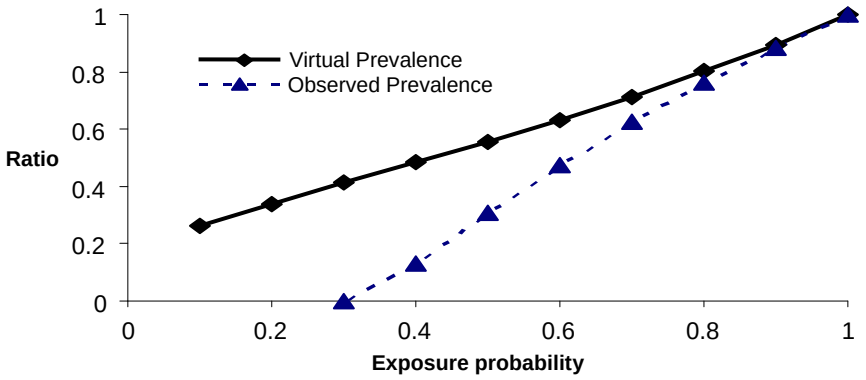


Figure 2. Ratio of estimated heritability to true heritability on the liability scale for varying levels of incomplete exposure, assuming virtual or observed prevalences of 0.3.

Imperfect diagnosis also impacts upon inferences from field data and heritability estimation. Fundamental to all diagnostic tests are the concepts of specificity and sensitivity. Specificity (F) is the probability that truly healthy animals (i.e. not infected by the pathogen of interest) are classified as such, and sensitivity (T) is the probability that truly infected or diseased animals are

correctly classified. True prevalence (p) may be estimated from estimated prevalence (p') as: $p = (p'+F-1)/(F+T-1)$. Imperfect specificity and sensitivity impact on the heritability of liability to infection (Bishop and Woolliams, in prep.). In general, for true disease prevalence less than 0.5, imperfect specificity will markedly reduce the estimated heritability and imperfect sensitivity will result in a minor reduction in the estimated heritability. For example, for a true prevalence of 0.3, sensitivity or specificity of 0.8 will result in heritability of liability being underestimated by 20% and 45%, respectively. The effects of imperfect sensitivity and specificity are reversed for prevalences greater than 0.5. The implication of this underestimation is that even low heritabilities under field conditions may be indicative of substantial true heritable variation in resistance.

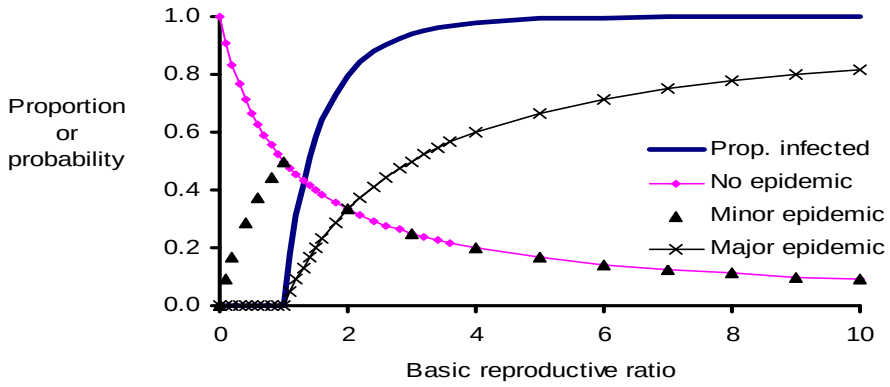


Figure 3. Epidemic probabilities, given the presence of an infected animal, and severity of major epidemics, as a function of the basic reproductive ratio, for hypothetical viral or bacterial diseases.

Predicting responses to selection. The flip side of the interpretation of field data is the prediction of responses to selection, using field data or genetic markers. The key point is that genetic theory has to be extended to include the epidemiological impact of infection, i.e. to account for the altered transmission of infection along the pathways shown in Fig. 1. This will result in non-linear relationships between average host genotype for resistance and disease outcomes at the population level. For macroparasitic infections, such as ruminant nematode infections, the outcome traits are simply indicators of the severity of infection or disease, measured at the individual animal level, and these impacts are described below. In general the feedback loops will lead to additional genetic progress (Bishop and Stear, 1997). The situation for microparasitic infections is more complex, requiring the development of genetic-epidemiological models to quantify the outcomes (MacKenzie and Bishop, 1999). Here, the outcomes may be assessed in terms of the probability and severity of epidemics, and the benefits depend on the infectiousness of the disease which is described by R_0, the basic reproductive ratio (i.e. the expected number of secondary cases after the introduction of a single infected animal). Increasing host resistance to infection has a largely linear impact in terms of reducing R_0, however the relationship between R_0 and epidemic outcomes is non-linear, as illustrated by Mackenzie and Bishop (2001). For highly infectious diseases (e.g. FMD) a large decrease in R_0 would be required before any noticeable improvement in epidemic outcomes would be observed, whereas for diseases whose R_0 is only slightly above 1.0 (e.g. scrapie) modest improvements in resistance may lead to large reductions in epidemic likelihood or severity. These impacts are illustrated in Fig. 3 for the probabilities of no, minor or major epidemics, given the introduction of a single infected animal into the population, and the

proportion of animals infected during the course of a major epidemic (Bishop and MacKenzie, 2003). A minor epidemic is one that dies out through stochastic events without intervention.

APPLICATIONS

Nematode infections in sheep. Genetic selection has often been used to help to control nematode infections in sheep, usually based on FEC (see summary by Bishop and Morris, 2007). In almost all cases FEC, once appropriately transformed, is a moderately heritable trait and one which responds to selection. Genome scans to detect QTL are now well advanced in many countries. With the exception of a QTL near the interferon gamma locus on chromosome 3, a feature of these studies is the difficulty in detecting QTL that are consistent between studies. Selection based on either phenotypic data or whole genome results obtained using a dense SNP chip would appear to be the most promising ways of achieving genetic progress.

An important feature of selection for nematode resistance is the interaction between host genotype and disease epidemiology, as altering host genotype can also change the force of infection faced by the population as a whole. In this case, by creating a population of animals that has lower mean FEC, the return of eggs to the pasture will decrease and the larval contamination on pasture will tend to decrease. This, in turn, will lead to reduced parasite challenge to all animals, furthering the benefits of selection. Therefore, the total benefits from selection are larger than those arising directly from genetic change in the host. This phenomenon was quantified *in silico* by Bishop and Stear (1997 and 1999). The benefits are manifested to some extent by decreased FEC, but more strongly by improved performance due to decreased larval challenge. Experimental verification of this phenomenon has been provided by Gruner et al. (2002) and Leathwick et al. (2002). This experimental demonstration is important, as the extra benefits are largely invisible if all animals are grazing the same pasture.

The model of Bishop and Stear (1997) has recently been extended to include a more mechanistic description of the development of immunity and the interactions between host and parasite (Vagenas et al., 2007a and b). In particular, this model incorporates host nutrition and, thus, when parameterised at the population level (Vagenas et al., 2007c) it can be used to explore interactions between host genotype and nutrition, and their impact on genetic parameters. This is important as it may go some way to resolving apparent contradictions seen across datasets and countries for genetic correlations between FEC and performance; for example, is the genetic correlation between FEC and growth rate favourable (negative) or unfavourable (positive). Assuming no linkage or pleiotropy between the genes underlying resistance to nematodes, Vagenas et al. (2007c) showed a marked predicted impact of level of nutrition; for poor nutrition moderate favourable correlations were predicted between FEC and lamb growth rate, however these correlations became essentially zero on high protein diets. These results were extended by Doeschl-Wilson et al. (2008) who demonstrated likely impacts of genetic correlations between underlying growth and immunological traits on predicted genetic parameters for production and resistance traits. Extreme genetic correlations observed from field studies could only be reproduced by assuming genetic relationships between the underlying input resistance traits. Altering preferences in the resource allocation between growth and immune response functions had less pronounced effects on the genetic parameters for the same traits. Effects were stronger when the allocation priority shifted towards growth, in which case worm burden and faecal egg counts increased and genetic correlations between these resistance traits and body weight became stronger. The results suggest that moderate pleiotropy and linkage may have large impacts on observed genetic parameters, and hence on outcomes of selection for nematode resistance.

Footrot. Footrot is a common cause of lameness in both lambs and mature sheep and it is a major welfare problem in sheep. Footrot is a highly contagious bacterial disease caused by *Dichelobacter* (*Bacteroides*) *nodosus*. In addition to the welfare concerns, it is also a major cause of economic loss. In the UK context it is estimated to have economic costs to the UK sheep industry of £24 million per annum (Nieuwhof and Bishop, 2005).

Substantial genetic variation in resistance to footrot has been demonstrated by Raadsma et al. (1994) using deliberate challenge data, and by Nieuwhof et al. (2008) using field data. Nieuwhof et al. (2008) found that data describing 'affected or not' was at least as heritable as data giving more detailed descriptions of the severity of infection, possibly because these data only describe the ~10% of animals that have clinical signs of disease and not the ~90% of animals that do not. Essentially this heritability describes the probability of animals being in the 'infectious/sick' category in Fig. 1. The heritabilities would rise if the non-affected grouping of animals could be distinguished into the other non-affected categories in Fig. 1, and if those that have not been exposed to infection could be identified. A further finding from Nieuwhof et al. (2008) was that heritability of liability to footrot appeared to increase with flock disease prevalence, even when corrections were made for prevalence effects. This suggests that the greater force of infection in high prevalence flocks has allowed genetic variation in resistance to be more strongly expressed.

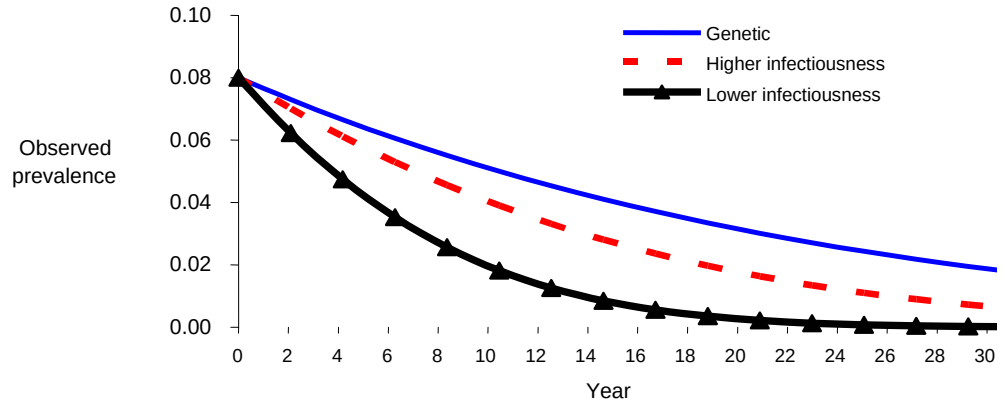


Figure 4. Responses to selection for footrot resistance, as assessed by observed prevalence of disease, predicted either by standard genetic theory or by a genetic epidemiological model, for situations where the bacteria have high or low infectivity.

Potential responses to selection for footrot have been quantified using a genetic-epidemiological model which accounts for transmission of infection via pasture as well as host genotype (Nieuwhof et al., 2009). Because footrot affected sheep can recover, show short-lived immunity to infection and ultimately become susceptible to reinfection (as shown in Fig. 1), epidemic properties are complex and an endemic equilibrium state can be reached in which the proportion of animals affected at any time point is largely constant, given constant weather conditions, stocking densities, management strategies, etc. Nieuwhof et al. (2009) derived mathematical terms describing the expected endemic equilibrium infection levels, and explored the properties of these equilibria as genetic progress was made for footrot resistance, assuming that resistance is expressed through the acquired immune response (hence the speed with which animals recover from infection. The results are shown in Fig. 4 for cases where genetic progress is estimated ignoring epidemiological impacts, and cases where the disease is assumed to have relatively high or low levels of infectiousness. For both assumptions about infectiousness, genetic

progress was somewhat faster than predicted by quantitative genetic theory alone, making selection a more attractive proposition. The more infectious the pathogen, hence the more quickly the disease spreads through the flock, the more closely the selection dynamics resemble those for normal production traits.

The footrot example illustrates several key lessons. Firstly, simple scoring criteria, such as presence or absence of hoof lesions, are heritable traits which should respond to selection. Secondly, responses to selection for disease resistance may be greater in higher prevalence herds. Lastly, total gains may exceed those predicted by quantitative genetic theory alone.

Mastitis. Mastitis, inflammation of the mammary gland, is usually caused by bacterial organisms such as *Staphylococcus spp.*, *Streptococcus spp.*, *Pseudomonas spp.*, *Mycoplasma spp.* and various coliforms such as *E. coli*. Mastitis incidence in the European dairy industries has been estimated at 30% of cows per year, and each case has been estimated to cost between 150 to 300 euros per diseased cow. Selection for increased milk yield will generally worsen the incidence of mastitis, due to the unfavourable genetic correlation between milk yield and mastitis susceptibility. Therefore, efforts to reduce mastitis, or prevent its incidence from rising, are a part of most dairy cattle evaluations. Currently, selection to reduce the incidence of mastitis is based on udder conformation, somatic cell count (SCC) and mastitis infection history. SCC and clinical mastitis generally have low heritabilities, usually in the range 0.05 to 0.15 (Rupp and Boichard, 2003). Mastitis resistance is probably due to structural attributes of the udder or teat, as well as immune responses. QTL associated with mastitis resistance traits have been reported on almost all of the 29 bovine chromosomes, in a variety of populations and breeds (see Khatkar et al., 2004). This large number of QTL suggests that gene-assisted selection, using causative mutations underlying these QTL, may be inefficient if each of these mutations explains only a small proportion of the observed variation in SCC or clinical mastitis. A genome-wide selection approach using a dense SNP chip would be advantageous.

However, there is also likely to be considerable benefit in redefining traits describing mastitis resistance, and this may also address concerns as to whether continued selection for reduced SCC is a long-term solution to mastitis. SCC is used as an indicator of mastitis, with high SCC values indicating that an animal is likely to be infected. However, SCC measurements on a group of animals comprise a mixture distribution trait describing baseline SCC in unaffected animals as well as elevated SCC in infected animals. The concern is that reducing SCC too far may reduce baseline SCC levels, hence animals' ability to respond to infection (Rupp and Boichard, 2003). In actual fact, the trait that is of interest to the breeder is liability to mastitis. Therefore, a rational approach when considering SCC data is to decompose it into baseline SCC values for uninfected animals, response SCC values for animals that are infected, along with the probability that a particular animal falls into one distribution and not the other. This concept was introduced by Odegard et al. (2005), and the quantitative genetic properties of such mixture distribution traits were formalised by Gianola et al. (2006). The primary selection criterion arising from this data decomposition is the liability of an animal to be affected by mastitis. The secondary question is whether to increase or decrease SCC; however, this question must be asked separately for baseline and response SCC. To answer this, it is necessary to calculate genetic correlations between SCC and mastitis liability, separately for baseline and response SCC. It is suggested that selection for resistance should not be for SCC *per se*, but for liability to infection, either conditional upon SCC or assessed independently through diagnoses of infection or clinical signs of mastitis.

The fact that mastitis is caused by different species of bacteria raises further issues of potential importance. Can these separate categories of infection be teased apart and is it beneficial to do so? Further, for the infectious (as opposed to 'environmental') sources of mastitis, further insight may be gained by assessing cow liability to mastitis in relation to the force of infection. For example,

are there genetic influences on the order in which animals become infected, and do genetic effects alter as disease prevalence changes? If these effects exist, they may point to additional epidemiological benefits from selection for increased resistance. However, it may require considerable quantities of detailed data to assess these effects.

CONCLUSIONS

This key take-home message from this paper is that the interpretation of data and estimation of benefits of selection differ for disease data compared to standard performance data, and these interpretations also differ between categories of disease. It is important to understand the disease in question, rather than simply treat data as if it were performance trait data. From a breeding perspective, the impacts of accounting for transmission of infection and disease epidemiology are generally positive. For example, weak genetic signals in field data may often be masking stronger underlying genetic effects. Further, true benefits of selection may well be greater than those predicted using naïve quantitative genetic theory that ignores disease epidemiology.

In summary, considerable opportunities exist to breed animals for enhanced disease resistance, often simply using noisy field data either directly for selection or to calibrate SNP chip data.

ACKNOWLEDGEMENTS

This paper has summarised work funded by the BBSRC, Defra, MLC, REERAD, GENUS and the EU. I wish to thank many colleagues for inputs, most of whom have been cited as coauthors.

REFERENCES

Bishop, S. C. (2005) In: "Encyclopedia of Animal Science", p. 288. editors W. G. Pond and A.W. Bell, Marcel Dekker Inc., New York.

Bishop, S.C. and MacKenzie, K.M. (2003) *Genet. Sel. Evol*. **35**:S3.

Bishop, S.C. and Morris, C.A. (2007) *Small Rum. Res.* **70**:48.

Bishop, S.C. and Stear, M.J. (1997) *Anim. Sci*. **64**:469.

Bishop, S.C. and Stear, M.J. (1999) *Anim. Sci.* **69**:515.

Davies, G., Genini, S., Bishop S.C. and Giuffra, E. (2009) *Anim*. **3**:415.

Doeschl-Wilson, A.B., Vagenas, D., Kyriazakis, I., Bishop, S.C. (2008) *Genet. Sel. Evol*. **40**:241.

Gianola, D., Heringstad, B. and Odegaard, J. (2006) *Genet*. **173**:2247.

Gruner, L., Cortet, J., Sauve, C., Limouzin, C. and Brunel, J.C. (2002) *Vet. Parasit.* **109**:277.

Khatkar, M.S., Thomson, P.C., Tammen, I. and Raadsma, H.W. (2004) *Gen. Sel. Evol*. **36**:163.

Leathwick, D.M., Atkinson, D.S., Miller, C.M., Brown, A.E. and Sutherland, I.A. (2002). Proc. of Novel Approaches III, a workshop on helminth control in livestock in the new millennium. Moredun Research Institute, Scotland. July 2-5, 2002.

MacKenzie, K. and Bishop, S.C. (1999) *Anim. Sci.* **69**:543.

MacKenzie, K. and Bishop, S.C. (2001) *J. Anim. Sci*. **79**:2057.

Nieuwhof, G.J. (2009) PhD Thesis, University of Edinburgh.

Nieuwhof, G.J. and Bishop, S.C. (2005) *Anim. Sci.* **81**:23.

Nieuwhof, G.J., Conington J. and Bishop, S.C. (2009) *Genet., Select., Evol.* **41**:19.

Nieuwhof, G.J., Conington, J., Bűnger, L. Haresign, W. and Bishop, S.C. (2008) *Anim*. **2**:1289.

Odegard, J., Madsen, P., Gianola, D., Klemetsdal, G., Jensen, J., Heringstad, B. and Korsgaard, I.R. (2005) *J. Dairy Sci.* **88**: 652.

Raadsma, H.W., Egerton, J.R., Wood, D., Kristo, C. and Nicholas, F.W. (1994) *J. Anim. Breed. Genet*. **111**:367.

Rupp, R. and Boichard, D. (2003) *Vet. Res.* **34:**671.

Vagenas, D., Bishop, S.C. and Kyriazakis, I. (2007a) *Parasit*. **134**:1263.

Vagenas, D., Bishop, S.C. and Kyriazakis, I. (2007b) *Parasit*. **134**:1279.

Vagenas, D., Doeschl-Wilson, A., Bishop, S.C. and Kyriazakis, I. (2007c) *Int. J. Parasit*. **37**:1617.

Proc. Assoc. Advmt. Anim. Breed. Genet. 18:263-271

REVIEW OF GENETIC PARAMETERS FOR DISEASE RESISTANCE IN SHEEP IN NEW ZEALAND AND AUSTRALIA

C.A. Morris
AgResearch, Ruakura Research Centre, PB 3123, Hamilton 3240, New Zealand.

SUMMARY

Genetic factors, mainly heritability estimates, have been reviewed briefly for all major disease traits affecting sheep production in New Zealand and Australia. The traits reported included resistances to nematode parasites, liver fluke, flystrike and fleece-rot, lice, mycotoxic diseases (facial eczema, ryegrass staggers, heat stress caused by ergovaline, and infertility caused by zearalenone), mastitis, foot rot and pneumonia. Selection lines have been under study for at least six of these traits. In general, all the traits are characterised by heritabilities of sufficient size that selection progress can be made, if it is cost-effective to carry out the scoring and apply selection pressure. Given the difficulty of scoring many of the diseases under field conditions, the advent of DNA marker technology could provide a large boost in the near future to reducing incidence of these diseases by breeding. The quantitative trait locus studies carried out will provide pointers to candidate genes controlling the expression of disease resistance traits. It is likely to be beneficial to investigate the underlying genes controlling each resistance trait, with a view to developing additional possible forms of disease control.

INTRODUCTION

Interest in breeding sheep for resistance to diseases has grown considerably over the last few decades. The famous textbook by Drs Helen Turner and Sydney Young (1969) contains only one reference to disease genetics in its index (associated with inbreeding depression), although there is a specific mention of fleece-rot. Since the 1970s, considerable research effort has been put into estimating genetic parameters for disease resistance traits in sheep, particularly in New Zealand and Australia. This has often been followed by analyses of correlations with production traits, and in some cases by attempts to introduce a disease test for use in industry flocks. There could be ethical and animal welfare issues associated with offering such tests, but equally there are issues about doing nothing, or about relying on drug treatments and other management strategies to remain effective for future generations.

This brief review will summarise genetic parameters, mainly heritability estimates, for some of the disease traits in sheep in New Zealand and Australia, with consideration given to the following disease groups: nematodes (including *Nematodirus*), other parasites (liver fluke, flystrike, lice), mycotoxic diseases, and some other diseases including fleece-rot, mastitis, footrot and pneumonia. Some of these areas have been covered fully in the past, whereas others have hardly received any attention. As recently as 2005, a review of Australian genetic parameters for sheep (Safari *et al.* 2005) included just two diseases, host resistances to endoparasites and to fleece-rot. Raadsma *et al.* (1997) described contemporary measurements of sheep for resistances to endoparasites, footrot, fleece-rot and dermatophilosis, the major diseases affecting Merinos in Australia, and they reported heritabilities and genetic correlations, described below. This paper attempts to cover the genetics of a range of diseases, regardless of industry-wide prevalence, because individual diseases may be highly relevant to just a group of farmers in one region, and this may create the demand for ram breeders in that region to apply selection.

RESULTS AND DISCUSION

Parasitic diseases

Nematodes. Host differences in susceptibility to endoparasites have been the subject of genetic studies in sheep for over 30 years. Following examples from early cattle studies (e.g., Frisch 1981), research groups attempted to monitor host-genetic variation in nematode parasite burdens via breed differences, then via genetic variation among sire groups, and then exploiting it via experimental selection lines. Many reviews have been published on the genetics of resistance to nematode parasites in sheep in New Zealand and Australia. Morris (2000) summarised published data on selection lines for high or low faecal egg count (FEC) in Romneys and Perendales in New Zealand, and in Merinos in Australia. In New Zealand there has also been a Romney line selected for high resilience (defined as the time to first drench post-weaning, whilst under nematode challenge, with acceptable growth rate and with minimal breech soiling). The realised heritability of the variously transformed functions of FEC listed in that report averaged 0.32 (with s.e.s, by experiment, ranging from 0.03 to 0.14), whilst the heritability of the measure of resilience to nematodes was 0.14 ± 0.03. More recently, Safari *et al.* (2005) have reviewed heritability estimates from many sources (published over the 1992-2003 years), and found a weighted average for transformed FEC of 0.27 ± 0.02, from 16 experiments. Heritabilities tend to be greater, on average, in experimental flocks than in industry data (Morris *et al*. 1995c), partly because of the greater degree of control of management in experimental flocks, and perhaps because of higher degrees of challenge to the animals. On the AAABG website (http://www.gparm.csiro.au; accessed 7 June 2009), there is still no summary of genetic parameters for this disease (or any other) in sheep.

In practical terms, host resistance to endoparasitism is now known as a heritable trait and is used in industry programmes: WormFEC in New Zealand, in collaboration with the recording scheme Sheep Improvement Ltd (SIL), and Nemesis in Australia. Genetic progress in FEC is being made in flocks where selection is applied, with heritabilities of ~0.2 (depending on FEC-sample timing) in New Zealand (M. J. Young, personal communication, March 2009), and 0.22 in Australia (Eady 2009). The next phase is to offer a DNA marker test or marker-assisted selection to breeders wishing to select for improved resistance, and the 'WormSTAR™' test [http://www.catapultsystems.co.nz/products/55_wormstar.cfm; accessed 7 June 2009], marketed by Catapult Genetics of Pfizer Animal Health, is now available to New Zealand ram breeders (but, at the time of writing, it is not yet validated for Merinos). In New Zealand, the WormSTAR™ marker explains approximately 2.3-3.6% of the genetic variation for the FEC traits, 4.8-5.5% of the live weight traits, 3.7% for the wool traits and 6.2% for lean weight (McEwan *et al.* 2008). Several studies have identified quantitative trait loci (QTL) for host resistance in sheep (reviewed by Dominik 2005). The interferon-gamma gene and haplotypes for the major-histocompatibility-complex have been identified in some studies, but not all. However, Crawford *et al*. (2006) noted that "Our failure to discover more QTL suggests that most of the genes controlling this trait are of relatively small effect".

Estimates of genetic correlations between FEC or transformed FEC and production traits appear to vary with breed and country, particularly estimates in coarse-woolled vs Merino breeds (e.g., Morris *et al*. (1997, 2000) in New Zealand, and Safari *et al*. (2005) in Australia). It should be noted that 'breed' is confounded with management /grazing conditions across countries, and sometimes also with parasite species involved in the parasitism, and with method of challenge (artificial vs natural; single-species vs mixed-species). The New Zealand papers cited show evidence that FEC is genetically correlated unfavourably with lamb growth, and with fleece weights at all ages, whereas the Merino data reviewed by Safari *et al.* (2005) suggest no significant

genetic correlations of FEC with lamb growth or fleece weight.

By anti-parasite antibody studies, Green *et al*. (1999) have shown in New Zealand that mixed-species challenge during the genetic selection process has led to host resistance to various individual parasite species. In Australia, Eady (2009) has shown for the *Haemonchus* selection lines that there is effective cross-resistance to different parasite species.

Heritable resistance to *Nematodirus* species has also been reported (Morris *et al*., 2004), with heritability estimates of 0.15 ± 0.03 in lambs of 4 months of age and 0.26 ± 0.04 at 6 months of age, with genetic correlations of these with FEC data recorded at the same ages having a weighted average of 0.43.

Liver fluke. The epidemiology of fasciolosis, or infestation by the liver flukes *Fasciola hepatica* or *F. gigantica*, has been reviewed by Spithill *et al.* (1999). Early studies established that genetic factors (breed differences) were involved in host resistance to each species: whilst many sheep breeds were susceptible to *F. hepatica* (including Merinos), some were resistant to *F. gigantica.* The latter worm species is of primary concern in tropical countries, but the host's liver metabolism in response to *F. gigantica* is perhaps of wider interest, as presented below. Raadsma *et al*. (2008a) reported the development of a predictive index of *F. gigantica* worm burden, including the use of cathepsin L5 antibody titre, eosinophilia, and the activity levels of serum enzymes secreted during parasitic injury to the liver: glutamate dehydrogenase (GDH), and injury to the bile duct: gamma-glutamyltransferase (GGT). Studies with the Indonesian Thin Tail breed, which appears to carry a major gene for resistance to *F. gigantica* (Roberts *et al*. 1997), have since led to the identification of 12 QTL for host resistance using, in part, the liver enzyme indicators of host response (Raadsma *et al*. 2008b). Since the Indonesian Thin Tail breed also displays partial resistance to *Haemonchus contortus*, and this resistance appears to be influenced by same gene, Raadsma (2009) has suggested that there may be a "broad effect" of the gene on immune response, because *Haemonchus* and *Fasciola* are from roundworm and flatworm genera, respectively.

It is also notable that GDH and GGT are important indicators of liver and bile duct injury in facial eczema disease in New Zealand (see later), although they are probably downstream indicators of injury, rather than part of the genetic cause of the injury in susceptible animals.

Flystrike and fleece-rot. These two diseases will be discussed together, because of the high genetic correlation (>0.9) between them (Raadsma, 1991). Fleece-rot is a precursor to flystrike. In Merinos, susceptibility to flystrike (body strike) is heritable (e.g., Raadsma 1991) reported an estimate of 0.26 ± 0.12). On the underlying scale, consisting of a continuous grading of liability, his data led to a heritability estimate of 0.54 ± 0.25. The review by Safari *et al*. (2005) reported heritabilities of 0.17 ± 0.02 for fleece-rot incidence, and 0.23 ± 0.02 for it as a severity score, in fine-wool Merinos. Slightly higher values were reported by McGuirk and Atkins (1984), and they also estimated a heritability of 0.40 ± 0.11 for fleece-rot liability on the underlying scale. The main fly species causing flystrike in Australia is *Lucilia cuprina*, and the last 15 years have seen its immigration and spread across New Zealand (Heath and Bishop 1995), leading to more intensive fly damage to New Zealand sheep. Mortimer *et al*. (2001) have published evidence suggesting that a major gene may account for 20% of the phenotypic variance in fleece-rot and 15% of the variance in body strike in Merinos, as a result of studies of selection lines of sheep bred for resistance or susceptibility to fleece-rot and flystrike. One alternative approach when wool is of very limited value is to select for bare rumps; Scobie *et al.* (2007) reported a heritability of 0.33 ± 0.06 for breech bareness score, and 0.59 ± 0.06 for the length of bare skin under the tail.

Lice. In a four-year study at AgResearch, infestations by the louse, *Bovicola ovis,* were monitored in Romney lambs managed primarily for other purposes at Wallaceville Station, Upper Hutt, New

Zealand (Pfeffer *et al*. 2007). This study included natural and artificial infestations, providing data for heritability estimates for log-transformed louse score of 0.22 ± 0.06 in autumn (~6 months of age), 0.34 ± 0.08 in winter, and 0.44 ± 0.09 for a combined score. Cockle scores from exposed lambs were also recorded *in vivo* on skin below a closely shorn area, and on pelts *post mortem*, and heritability estimates for cockle score were 0.06 ± 0.04, 0.45 ± 0.09 and 0.40 ± 0.09 in autumn, winter, and combined, respectively. The genetic correlation between mean louse score and mean cockle score was 0.97 ± 0.04, and the genetic correlations between mean louse score and the levels of two different anti-louse wool antigens were 0.96 ± 0.08 and 0.95 ± 0.09. It was concluded that monitoring wool antigens may be a practical way of producing an objective score of susceptibility to lice.

Mycotoxic diseases. Four mycotoxic diseases will be discussed: facial eczema (FE), caused by the sporidesmin toxin, ryegrass staggers (RGS), caused by the lolitrem B toxin from the *Neotyphodium lolii* endophyte, heat stress caused by ergovaline, and infertility caused by zearalenone. FE receives wide publicity in New Zealand, because of its prevalence in the North Island, but greater awareness of mycotoxins in both countries might assist farmers in trying to avoid grazing conditions where there is severe risk of mycotoxic poisoning of stock.

To protect against each disease, the options are 1). To avoid grazing toxin-containing pasture, 2). To protect animals that do graze it, or 3). To breed resistant animals. Recent reviews by Bishop and Morris (2007) and Morris and Phua (2009) have covered the genetics of resistance to these diseases. For sheep in New Zealand, the heritability estimate for FE resistance (via GGT as an enzyme indicator) is 0.45 ± 0.05 (Morris *et al*. 1995a), and for RGS (0-9 score) it is 0.36 ± 0.04 (Morris *et al.* 2007) with a previous estimate of 0.13 ± 0.05 from a binomial scoring system (Morris *et al.* 1995b). Relationships between the indicator for FE and the liver injury caused by the disease itself have been reported by Morris *et al.* (2002). Heritability estimates are not available for the susceptibility of sheep to ergovaline, but American data suggest that differences in animal susceptibility to tall fescue toxicosis (which is also caused by ergovaline) are heritable in cattle (Lipsey *et al.,* 1992), and experimental selection for or against resistance to dietary ergovaline was successful in mice (Hohenboken and Blodgett 1997). For zearalenone, urinary breakdown products of zearalenone have been measured after controlled dosing, and heritability estimates have been obtained in experimental animals (0.32 ± 0.10) and in the field (0.19 ± 0.07) (Amyes and Morris 2008).

A phenotyping service for FE susceptibility,'Ramguard', has been offered to New Zealand ram breeders since 1984 (Morris *et al*. 1994), and this provides a sporidesmin-dosing procedure and GGT-enzyme response measure, carried out under controlled conditions with veterinary supervision. According to the medium/high heritabilities reported for resistance to these mycotoxic diseases in sheep, it should be feasible to select for resistance in industry flocks, if there is a financial incentive. Breeding resistant rams might only be relevant to the objectives of farmers within a limited geographical range, but performance of commercial sheep farmed in that range is likely to improve markedly as a result. The next stage for FE is to offer a DNA marker test or marker-assisted selection, to breeders wishing to select for greater resistance. Experimental flocks have been generated and managed at AgResearch Ruakura for resistance or susceptibility to FE since 1975 (Morris *et al*. 1995a), and these have been used to study the underlying biology and to search for DNA markers of resistance (Phua *et al*. 2009). Genetic correlations between FE and production traits in the FE selection lines have been published: for resistant-line animals, lamb weights were, on average, 5-6% lighter, and yearling greasy fleece weights 8% heavier than in susceptible-line animals, and for reproduction there was no significant difference (Morris *et al*. 1999).

For RGS, a demonstration of the potential for widespread toxic effects has come from Reed *et*

al. (2005) who reported mass deaths in 2002 from severe "perennial ryegrass toxicosis" on 224 Southern Australian farms (29,109 sheep and 448 cattle; up to a 30% mortality is some grazing groups); 2002 was one of three severe seasons for this in a 20-year period in Southern Australia. The clinical symptoms for RGS in Southern Australia appeared more serious than those for RGS in New Zealand, and may have resulted from the Australian endophyte producing a combination of two toxins, lolitrem B and ergovaline (Reed *et al.* 2005). High positive correlations have also been reported in the USA between lolitrem B and ergovaline concentrations in over 450 endophyte-infected perennial ryegrass samples (Hovermale and Craig 2001). Divergent selection has been applied successfully since 1993 at Ruakura for resistance or susceptibility to RGS (Morris *et al*. 2007).

Indicative genetic associations are now available among effects of the various mycotoxic disease traits: a). a genetic correlation of 0.31 between resistance to FE and to RGS (Morris *et al*. 1995b), a positive relationship between resistance to lolitrem B and to ergovaline (unpublished material cited by Morris and Phua 2009), and c). a positive relationship between resistance to sporidesmin and ergovaline, in mice (Hohenboken *et al*. 2000). It is known that these toxins are from different chemical families, and that their modes of action are different. Nevertheless, finding positive associations suggests that at least some parts of the detoxification pathways are common. In the case of zearalenone, one of its breakdown products (α-zearalenol) is thought to be more oestrogenic than zearalenone itself, at least in monogastrics, where there is also competition with the host's oestrogen receptors (Hagler *et al*. 1979). There is a suggestion of an unfavourable correlation between resistance to FE and to zearalenone (Smith and Morris 2006), which could be explained by this finding.

Some other diseases

Mastitis. Bacterial infection in the mammary gland of most lactating ewes may not be as obvious a problem as in machine-milked cows, but it is a particular problem in dairy sheep operations. A brief review by Bishop and Morris (2007) reported that somatic cell counts (SCC) can be used as a diagnostic of subclinical infection in ewes as in cows (though this doesn't necessarily follow across species, e.g., in goats). Most recent estimates of heritability for SCC in ewes range from 0.10 to 0.20 (e.g., Gonzalo *et al*., 2003; Legarra and Ugarte, 2005), as in dairy cattle. The sign of genetic correlations between SCC and milk yield in dairy ewes is not clear, as a result of quite variable estimates published so far. Results of a search for QTL for mastitis and other lactation traits in dairy sheep have been summarised by Barillet (2007).

Footrot. Footrot is a bacterial disease caused by *Dichelobacter (Bacteroides) nodosus*, and it causes lameness in all classes of sheep. Raadsma *et al.* (1994) reported considerable genetic variation in levels of host resistance to both natural and artificial challenge, using a scoring system modified from one originally developed by Egerton and Roberts (1971). Given the moderate repeatability of footrot scores, combining data and estimating heritabilities from mean scores has led to heritability estimates "which approached 0.30 for most indicators" (Raadsma, 2000). Heritability estimates reported from Britain (Nieuwhof *et al*. 2008) were slightly smaller (0.12 to 0.23), and depended on incidence. In New Zealand, successful breeding programmes for footrot resistance have been reported for Corriedales by Skerman and Moorhouse (1987), and for Merinos by Patterson and Patterson (1989). Skerman *et al*. (1988) reported heritability estimates of 0.28 for a binomial index of footscald and/or footrot, and 0.17 for footrot incidence alone. Associations between resistance and the major histocompatibility locus have been observed (Escayg *et al*. 1997; Raadsma *et al*. 1999). A specific association with the DQA2 gene was reported by Hickford *et al*. (2004).

Pneumoni., Subclinical pneumonia is common in New Zealand amongst lambs under stress, and can be caused by viruses, bacteria or parasites (Merck Veterinary Manual 2009). It has been reported with an average flock prevalence of 22-29% in the Canterbury, Gisborne and Manawatu districts of New Zealand; prevalence was higher in northern New Zealand than in Southland (Goodwin-Ray 2006). There was a suspected breed difference (based only on anecdotal data), with the Merino being less susceptible than coarse-woolled breeds in matched environments. A farm-level survey by Goodwin-Ray *et al*. (2008) showed that management factors affecting incidence included shearing lambs on weaning day, breeding ewe replacements on the farm, increasing the percentage of lambs sold between March (~5 months of age) and May, whereas fixed stocking-rate grazing, and protective vitamin B12 injection at weaning were two factors associated with lower incidences. In Australia, Abbott and Maxwell (2002) noted: "It is likely that, either alone or in combination with other disease conditions, respiratory diseases are a significant cause of loss to the Australian sheep industry." These include parasitic and microbiological conditions. Although no genetic studies appear to have been published on pneumonia in sheep, Snowder *et al.* (2006) published data in the USA from cattle suggesting a heritable component of 0.18 for a respiratory disease score (transformed to the underlying continuous scale), and Heringstad *et al*. (2008) in Norway reported a heritability estimate of 0.05 (95% confidence intervals, 0.02 to 0.09) for respiratory disease incidence in cattle, using an underlying scale. It is possible that, in studies so far, the genetic variance is swamped by uncontrolled environmental variance, and that the latter could be reduced in more detailed future studies. By analogy, it has already been noted that heritability estimates of nematode parasite resistance are higher under experimental conditions than in field data (Morris *et al*. 1995c).

EXPERIMENTAL SELECTION FLOCKS

For those disease traits in New Zealand and Australian sheep where selection has been applied in experimental lines, genetic progress has indeed been achieved in each, with estimates of realised heritabilities from the 12 sets of actual selection lines averaging 0.28 (Morris 2000), with a range from 0.13 to 0.45 (only three below 0.20). The lines reviewed involved selection for one of six single traits: resistance to nematode parasites, resilience to nematode parasites, and resistance to facial eczema, ryegrass staggers, body strike, and dermatophilosis. Realised annual responses reviewed by Morris (2000) averaged 0.073 phenotypic standard deviations per year.

CONCLUSIONS

For the limited number of sheep diseases that could be reviewed here, most heritabilities ranged from 0.2 to 0.4. It can be concluded that genetic progress can be achieved if it is economic to apply selection pressure, and this prediction is borne out by experience from the experimental sheep selection lines described. Given the difficulty of scoring many of the diseases under field conditions, the advent of DNA marker technology could provide a large boost in the near future to reducing incidence of these diseases by breeding, in particular taking advantage of the high density single nucleotide polymorphism approach just coming on stream in sheep [http://www.agresearch.co.nz/snp/snp-chip.asp; accessed 7 June 2009]. Without this, the factors (apart from economics) determining feasibility of disease-resistance selection under commercial conditions are developing a scoring system with the required accuracy, and having the flock manager believe in the objectives sufficiently to disease-challenge animals, in spite of an expectation that this will reduce their own performance.

REFERENCES

Abbott, K.A. and Maxwell, W.M.C. (2002) Faculty of Vet. Sci., Univ. Sydney, 2002. In "*Diseases of the respiratory system*". Vet. Education and Information Network, accessed 1 Apr. 2009, http://vein.library.usyd.edu.au/sheephealth/images/chapter20.pdf .

Amyes, N.C. and Morris, C.A. (2008) *Proc. NZ Soc. Anim. Prod.* **68**:122.

Barillet, F. (2007) *Small Ruminant Res*. **70**:60.

Bishop, S.C. and Morris, C.A. (2007) *Small Ruminant Res*. **70**:48.

Crawford, A.M., Paterson, K.A., Dodds, K.G., Diez Tascon, C., Williamson, P.A., Roberts Thomson, M., Bisset, S.A., Beattie, A.E., Greer, G.J., Green, R.S., Wheeler, R., Shaw, R.J., Knowler, K. and McEwan, J.C. (2006) *BMC Genomics* **7**:178.

Dominik, S. (2005) *Genet. Sel. Evol.* **37 (Suppl. 1)**:S83.

Eady, S.J. (2009) http://www.csiro.au/resources/pfb8.html; accessed 7 June 2009.

Egerton, J.R. and Roberts, D.S. (1971) *J. Comp. Pathol*. **81**:179.

Escayg, A.P., Hickford, J.G.H. and Bullock, D.W. (1997) *Res. Vet. Sci.* **63**:283.

Frisch, J.E. (1981) *J. Agric. Sci. (Camb.)* **96**:23.

Gonzalo, C., Ariznabarreta, A., Othmane, M.H., Carriedo, J.A., De la Fuente, L.F. and San Primitivo, F. (2003) *J. Anim. Breed Genet*. **120**:282.

Goodwin-Ray, K.A. (2006) "Pneumonia and pleurisy in sheep: Studies of the effect on lamb growth rate, prevalence, risk factors, vaccine efficiency and economic impact." Final Report, March 2006. Meat & Wool New Zealand – project 97AH/AG188.

Goodwin-Ray, K.A., Stevenson, M. and Heuer, C. (2008) *Prev. Vet. Med.* **85**:136.

Green, R.S., Morris, C.A., Douch, P.G.C., Wheeler, M., West, C.J. and Hickey, S.M. (1999) Livest. Prod. Sci. **58**:129.

Hagler, W.M., Mirocha, C.J., Pathre, S.V. and Behrens, J.C. **(**1979) *Appl. Env. Microbiol*. **37**:849.

Heath, A.C.G. and Bishop, D.M. (1995) *Surveillance* **22**:11.

Heringstad, B., Chang, Y.M., Gianola, D. and Osteras, O. (2008) *J. Dairy Sci.* **91**:367.

Hickford, J.G.H, Zhou, H., Slow, S. and Fang, Q. (2004) *J. Anim. Sci.* **82**:1553.

Hohenboken, W.D. and Blodgett, D.J. (1997) J. *Anim. Sci.***75**:2165.

Hohenboken, W.D., Robertson, J.L., Blodgett, D.J., Morris, C.A. and Towers, N.R. (2000) J. *Anim. Sci.* **78**:2157.

Hovermale, J.T. and Craig, A.M. (2001) *J. Vet. Diagnostic Investig.* **13**:323.

Legarra, A. and Ugarte, E. (2005) *J. Dairy Sci*. **88**:2238.

Lipsey, R.J., Vogt, D.W., Garner, G.B., Miles, L.L. and Cornell, C.N. (1992) *J. Anim. Sci.***70 (Suppl. 1)**:188 (Abstr.).

McEwan, J.C., Weston, N.K., Payne, G.M., O'Sullivan, N.H., Auvray, B.N.E.E. and Dodds, K.G. (2008) Patent: Ovine identification method number 556506, http://www.iponz.govt.nz

McGuirk, B.J. and Atkins, K.D. (1984) *Aust. J. Agric. Res.* **35**:423.

Merck Veterinary Manual (2009) viewed 6th Mar. 2009, http://www.merckvetmanual.com/mvm/index.jsp?cfile=htm/bc/121500.htm

Morris, C.A. (2000) Chapter 16, pp.343-55 in "Breeding for disease resistance in farm animals", 2nd edition, editors R.F.E. Axford, S.C. Bishop, F.W. Nicholas and J.B. Owen, CABI Publishing, Wallingford, UK.

Morris, C.A., Amyes, N.C. and Orchard, R.L. (2007) *Proc. NZ Soc. Anim. Prod.* **67**:204.

Morris, C.A., Amyes, N.C., Towers, N.R. and Wesselink, K. (1999) *NZ J. Agric. Res.* **42**:475.

Morris, C.A., Bisset, S.A., Vlassoff, A., Baker, R.L., Watson, T.G. and Wheeler, M. (1997) *Proc. Assoc. Advmt. Anim. Breed. Genet*. **12**:50.

Morris, C.A., Bisset, S.A., Vlassoff, A., West, C.J. and Wheeler, M. (2004) *Anim. Sci.* **79**:33.

Morris, C.A. and Phua, S.H. (2009) "Metabolic diseases in sheep and cattle". Chapter (in press) in "Breeding for disease resistance in farm animals", 3rd edition, editors R.F.E. Axford, J.B. Owen, F.W. Nicholas and S.C. Bishop, CABI Publishing, Wallingford, UK.

Morris, C.A., Smith, B.L. and Hickey, S.M. (2002) *NZ Vet. J.* **50**:14.

Morris, C.A., Towers, N.R, Wesselink, C. and Wheeler, M. (1994) *Proc. NZ Soc. Anim. Prod.* **54**:263.

Morris, C.A., Towers, N.R., Wheeler, M. and Wesselink, C. (1995a) *NZ J. Agric. Res.* **38**:211.

Morris, C.A., Towers, N.R., Wheeler, M. and Amyes, N.C. (1995b) *NZ J. Agric. Res.* **38**:367.

Morris, C.A., Vlassoff, A., Bisset, S.A., Baker, R.L., Watson, T.G., West, C.J. and Wheeler, M. (2000) *Anim. Sci.* **70**:17.

Morris, C.A., Watson, T.G., Bisset, S.A., Vlassoff, A. and Douch, P.G.C. (1995c) Chapter 5, pp. 77-98 in "Breeding sheep in New Zealand for resistance or resilience to nematode parasites", editors G.D. Gray, R.R. Woolaston and B.T. Eaton, ACIAR, Canberra, Australia.

Mortimer, S.I., Henshall, J.M. and Tier, B. (2001) *Proc. Assoc. Advmt. Anim. Breed. Genet.* **14**:171.

Nieuwhof, G.J., Conington, J., Bünger, L., Haresign, W. and Bishop, S.C. (2008) *Animal* **2**: 1289.

Patterson, R.G. and Patterson, H.M (1989) *J. NZ Mountain Lands Inst.* **46**:64.

Pfeffer, A., Morris, C.A., Green, R.S., Wheeler, M., Shu, D., Bisset, S.A. and Vlassoff, A. (2007) *Int. J. Parasitol.* **37**:1589.

Phua, S.H., Dodds, K.G., Morris, C.A., Henry, H.M., Beattie, A.E., Garmonsway, H.G., Towers, N.R. and Crawford, A.M. (2009) *Anim.Genet.* **40**:73.

Raadsma, H.W. (1991) *Aust. J. Agric. Res* **42**:279.

Raadsma, H.W. (2000) Chapter 10, pp. 219-41 in "Breeding for disease resistance in farm animals", 2nd edition, editors R.F.E. Axford, S.C. Bishop, F.W. Nicholas and J.B. Owen, CABI Publishing, Wallingford, UK.

Raadsma, H.W. (2009) "Genetic and immunological characterisation of high resistance to internal parasites in Indonesian Thin Tail Sheep", ACIAR report, # AS1/1997/027, viewed 4th Apr. 2009, http://www.aciar.gov.au/project/AS1/1997/027 .

Raadsma, H.W., Egerton J.R., Wood, D., Kristo, C. and Nicholas, F.W. (1994) *J. Anim. Breed. Genet.* **111**:367.

Raadsma, H.W., Fullard, K.J., Kingsford, N.M., Margawati, E.T., Estuningsih, E., Widjayanti, S., Subandriyo, Clairoux, N., Spithill, T.W. and Piedrafita, D. (2008b) Chapter 8, pp. 89-113 in "Genomics of Disease", editors J.P. Gustafson, J. Taylor and G. Stacey, Springer New York.

Raadsma, H.W., Kingsford, N.M., Suharyanta, Spithill, T.W. and Piedrafita, D. (2008a) *Vet. Parasitol.* **154**:250.

Raadsma, H.W., McEwan, J.C., Stear, M.J. and Crawford, A.M. (1999) *Vet. Immunol. Immunopathol.* **72**:219.

Raadsma, H.W., Nicholas, F.W. and Egerton, J.R. (1997) *Proc. Assoc. Advmt. Anim. Breed. Genet.* **12**:63.

Reed, K.F.M., Scrivener, C.J., Rainsford, K.A. and Walker, L.V. (2005) Chapter 1.4, pp. 43-54 in "*Neotyphodium* in cool-season grasses", editors C.A. Roberts, C.P. West and D.E. Spiers, Blackwell Publishing, UK.

Roberts, J.A., Widjayanti, S., Estuningsih, E. and Hetzel, D.J. (1997) *Vet. Parasitol.* **68**:309.

Safari, E., Fogarty, N.M. and Gilmour, A.R. (2005) *Livest. Prod. Sci.* **92**:271.

Scobie, D. R., O'Connell, D., Morris, C. A. and Hickey, S. M. (2007) Aust. J. Agric. Res. **58**:161.

Skerman, T.M., Johnson, D.L., Kane, D.W. and Clarke, J.N. (1988) *Austr J. Agric. Res.* **39**:907.

Skerman, T.M. and Moorhouse, S.R. (1987) *NZ Vet. J.* **35**:101.

Smith, J. F. and Morris, C. A. (2006) *Proc. NZ Soc. Anim. Prod.* **66**:306 and 319.

Snowder, G. D., Van Vleck, L.D., Cundiff, L.V. and Bennett, G.L. (2006). *J. Anim. Sci.* **84**:1999.

Spithill, T.W., Smooker, P.M. and Copeman, D.B. (1999) Chapter 15, pp. 465-525 in "*Fasciola gigantica*: epidemiology, control, immunology and molecular biology" editor J.P. Dalton, CABI Publishing, Wallingford, UK.

Turner, H.N. and Young, S.S.Y. (1969) In "Quantitative genetics in sheep breeding." Macmillan of Australia.

Proc. Assoc. Advmt. Anim. Breed. Genet. 18:272-278

OPPORTUNITIES TO BREED FOR RESISTANCE TO BREECH STRIKE IN MERINO SHEEP IN A MEDITERRANEAN ENVIRONMENT

J.C. Greeff and L.J.E. Karlsson

Department of Agriculture and Food Western Australia, South Perth, WA, 6151

SUMMARY

Three lines of Merino sheep that were selected on breech wrinkles, breech cover, dags, urine stain and wool colour from research and industry flocks were exposed to flies from birth to hogget age. Significant differences ($P<0.01$) existed between lines in breech strike from post-weaning to hogget shearing but not from birth to hogget shearing. The lines differed significantly ($P<0.05$) for the indicator traits except for dags. Highly significant differences ($P<0.01$) were found between sire progeny groups for incidence of breech strike. Dags was the dominant predisposing factor for breech strike in one of four years. Breech strike in unmulesed sheep had a heritability of 0.57 ± 0.28 which was genetically positively correlated to all the scored indicator traits but had the strongest relationship with dags ($r_g=0.86 \pm 0.17$).

INTRODUCTION

Flystrike became a serious disease for Merino sheep with the arrival of *Lucilia cuprina* and the introduction of the extremely wrinkly Vermont Merino from the USA in the late 1800's (Cameron 1999). Breech strike occurs when the wool and skin in the breech area becomes wet from capturing and holding rain, urine or fluid faeces during periods when the blowflies are active. This attracts the fly and creates an ideal environment for the gravid fly to lay its eggs, and for the eggs to hatch. When the hatched maggots develop into the third instar stage, they can cause serious damage and death to sheep, if not detected and treated early. In 1931, Seddon *et al*. (1931) showed that breech folds was a major determining factor. The limited treatment options at that time resulted in the development of the mules operation by Mr JWH Mules in 1931 in which the excess loose caudal skin folds are surgically removed. After healing this artificial extension of the bare skin area reduced breech strike as well as making crutching easier, hence its gradual adoption by the industry. However, changing public values has resulted in the questioning of this method from a welfare position and strong pressure is being applied onto the Australian Industry to phase out mulesing. A number of different alternatives are being pursued, but it is generally agreed that breeding is the only long term sustainable solution.

Many studies have been carried out to identify blowfly resistant animals (Atkins and McGuirk 1979; Mortimer 2001; Scobie *et al.* 2002). A number of studies has shown that wrinkles, wool colour and fibre diameter variability traits (James 2008; Raadsma and Rogan, 1987; Raadsma *et al*. 1997) contribute to body strike, and bare breech and dags (Scobie *et al*. 2002) to breech strike. Edwards *et al.* (2009) have confirmed that the bare breech trait is heritable, while Smith *et al.* (2009) have also shown that wrinkles and dags are correlated with the incidence of breech strikes. However, these indicator traits are generally correlated and confounded, and therefore this approach does not present a clear picture of the relationship between the indicator traits and breech strike on an individual basis. No study has been carried out on unmulesed Merino sheep where the visually scored indicator traits at weaning and hogget fleece traits, were jointly evaluated against breech strike. The aim of this study was to demonstrate the changes in breech strike in a Merino flock by selecting Merino sheep on the known indicator traits, *i.e.* breech wrinkles, breech cover, dags, urine stain and wool colour, to quantify their relative importance in contributing to breech strike in a Mediterranean environment and to determine whether genetic differences exist between sires in breech strike resistance.

MATERIALS AND METHODS

Animals. Three lines of 200 ewes each were established on the Mt Barker research station in Western Australia. This research station has an annual winter rainfall of approximately 700mm. The first line (BSR-A) represents a scenario in which a ram breeder would select both his rams and ewes intensely for the known indicator traits and using any breech strike information. The second line (BSR-B) represents a scenario in which a commercial producer buys rams from a BSR-A rambreeder and only culls replacement ewes using the indicator traits. The third flock consists of an unselected control flock (BSR-con).

To generate more reliable results the above mentioned three lines were selected from representative commercial populations by using existing knowledge on the known indicator traits. The foundation ewe population was established by screening ewe lambs on 10 different medium wool properties in Western Australia in 2005. On each property 22 ewe lambs were selected for each of the three lines at marking as follows. Firstly, ewe lambs were selected randomly for the control flock after which average animals were selected for line 2. As it was only possible to differentiate between lambs for wrinkles at this early age, the plainest lambs on each property were identified for line 1. A similar approached was used to screen mature ewes from three different medium wool research flocks of the Department of Agriculture and Food Western Australia. Potential rams from research and industry flocks were identified based on their performance on the key indicator traits of breech cover, wrinkles and dags. However, very little objective information was available on industry rams and most industry rams were used on the basis of their own wrinkle and breech cover scores.

The screened sheep were transported to the Mt Barker research station in December 2005. Mating took place in February/March with lambing in July/August from 2006 to 2008. Half of each sire's progeny were mulesed in the first 3 years, but this practise ceased in 2008. However, only the records of unmulesed sheep were used in this study. The progeny were scored for breech wrinkle (WSB), breech cover (BC), dags (DS), wool colour (WCOL) and urine stain (US) at weaning using the National Visual Sheep Scoring System (AWI 2007). No preventative fly treatments such as crutching and jetting were applied to the young sheep to ensure that the animals were appropriately challenged. Animals were allowed to be struck naturally until hogget shearing. Any struck animal was identified, treated with a short acting chemical and returned to the flock. The total number of fly strikes from birth to post-weaner shearing and from post-weaner shearing to hogget shearing, were recorded per site on all struck animals. Breechstrw (breech strike birth to post-weaning shearing) and breechstrh (breech strike post-weaning to hogget shearing) indicate the total number of flystrikes in the breech over these two periods, respectively. In 2008 the Rylington Merino (RM) flock (Karlsson and Greeff, 2006) was included in this trial because of the relatively small number of unmulesed sheep in this trial. The same traits were recorded and measured in the RM flock as in the BSR flock. However, the RM flock was managed separately to the BSR flock up to weaning, but after weaning the sexes were separated in each flock and combined across flocks.

Data. Two sets of data were used in this study. The first dataset was recorded on the BSR flock that consisted of 1882 records of which 710 animals were mules and 1172 unmulesed sheep. The RM flock contributed 505 records from the 2008 born lambs that were recorded up to December 2008, resulting in a combined dataset of 2387 records from 46 sires and 1061 dams of which 1677 progeny were from unmulesed sheep. The average sire progeny group size was 35 which varied between 11 and 73 animals. No pedigrees were available on animals sourced from industry flocks.

Statistical analysis. The breech strike counts from birth to post-weaning shearing (breechstrw) and from post-weaning shearing to hogget shearing (breechstrh) were pooled into breech strike (breechstr). This trait and the indicator traits on unmulesed sheep of the BSR lines were analysed using ASREML (Gilmour *et al.* 2006). A univariate analysis was carried out by fitting line (3 levels), property of origin (1-12), year of birth (2005, 2006, 2007, 2008), sex (male and female), birth type (singleton/multiple) and age of the dam (7 levels) as fixed factors. As breech strike was a count and the mean and variance was the same, a Poisson distribution with a log link function was fitted to breech strike. A multivariate analysis was then carried out by adding breech cover, breech wrinkle, dags, wool colour and urine stain at weaning and the fleece traits, greasy fleece weight (GFW), clean fleece weight (CFW) fibre diameter (FD), coefficient of variation of fibre diameter (CV), clean yield (YLD) and fibre curvature as covariates to the model to identify their relative contribution to breech strike. As all the traits were not available in every year, this was undertaken separately for each year for the RM flock.

The data from the BSR and RM flocks were then combined and the genetic parameters estimated by using the Gibbs sampling threshold program from the BLUPF90 suite of programs (Misztal *et al.* 2002). An animal model was fitted with flock, property of origin and the significant environmental factors as fixed factors. The sires' solutions were then regressed against the incidence of breech strike of their progeny groups.

Table 1. Percentage of animals that were struck by flies per site from 2005 to 2008 for the different lines. (M = Mulesed; UM = Unmulesed)

Birth	Line	Number		Body strike %		Pollstrike %		Breechstrw[b] %		Breechstrh[c] %	
Year		M	UM	M	UM	M	UM	M	UM	M	UM
2005	BSR_A	106	105	0.0	3.2	0.0	0.0	0.9	0.0	1.1	5.3
	BSR_B	102	102	0.0	3.4	0.0	1.1	0.0	2.9	2.1	10.3
	Control	102	103	0.0	0.0	0.0	0.0	1.0	1.0	4.6	15.9
2006	BSR_A	84	79	1.2	1.3	22.2	23.1	6.0	25.3	3.6	1.3
	BSR_B	63	62	6.3	4.8	38.5	31.8	3.2	24.2	6.3	1.7
	Control	76	74	5.3	9.5	48.1	48.1	7.9	32.4	4.2	11.1
2007	BSR_A	66	62	15.1	12.5	13.2	18.8	1.5	9.7	3.8	16.7
	BSR_B	66	72	18.8	20.9	16.7	18.6	1.5	11.1	4.2	58.1
	Control	56	59	10.7	22.7	10.7	4.5	3.6	6.8	7.1	50.0
2008	BSR_A	0	199		4.0		2.5		25.6		
	BSR_B	0	144		6.9		1.4		20.1		
	Control	0	113		10.6		2.7		40.7		
	RM[a]	0	505		4.0		0.4		23.0		

[a] Not comparable with other lines as RM flock was managed separately for most of the time
[b] Breech strike up to post-weaning shearing. [c] Breech strike from post-weaning to hogget shearing

RESULTS AND DISCUSSION

The percentage of animals that were struck on the body, poll and breech at weaning and in the breech between weaning and hogget shearing in mulesed and unmulesed sheep in the different lines over four years are shown in Table 1.

The total number of all types of strike gives an indication of the challenge that the animals received from flies. Twelve percent of the 2005 drop, 45% of the 2006 drop, 46% of the 2007 drop

and 26% of the 2008 drop animals (during first 5 months) were struck. About 65.4 % of all strikes were in the breech, 17.5% on the body and 17.1% on the head. Sixteen percent of the males and 22.1% of females were struck. Of the 1677 unmulesed animals, 356 (21.2%) animals were struck in the breech once, 36 (2.15%) were struck twice and 4 (0.24%) animals were struck 3 times from birth to hogget age. Breechstrw of the BSR-A line was not significantly different between the BSR lines but breechstrh (post-weaning to hogget shearing) was highly significantly (P<0.01) different between lines *i.e.* 7.7%, 20.7% vs. 25.6% for the BSR-A, BSR-B and BSR-C lines, respectively.

Table 2. Number of records, means and standard deviation (SD) of breech strike to hogget age and of the indicator traits at weaning for the breech strike (BSR) flock and the Rylington Merino (RM) flock (BSR-A = selection line 1; BSR-B = selection line 2 and BSR-C = control)

Trait	n	Mean	SD	Level of significance
Total number of breech strikes to hogget age				
BSR-A	296	0.22	0.48	NS
BSR-B	240	0.27	0.53	
BSR-C	209	0.32	0.53	
RM [a]	498	0.24	0.47	
Breech wrinkle score (WSB)				
BSR-A	296	1.48	0.75	P<0.01
BSR-B	240	1.80	1.09	
BSR-C	209	1.97	1.16	
RM [a]	498	1.03	0.18	
Breechcover (BC)				
BSR-A	296	3.49	0.56	P<0.05
BSR-B	240	3.48	0.51	
BSR-C	209	3.55	0.52	
RM [a]	498	3.62	0.43	
Dag score (DS)				
BSR-A	356	2.35	0.56	NS
BSR-B	309	2.39	0.68	
BSR-C	267	2.38	0.71	
RM [a]	498	2.64	0.66	
Wool colour (WCOL)				
BSR-A	296	2.29	0.46	P<0.01
BSR-B	240	2.33	0.43	
BSR-C	209	2.41	0.46	
RM [a]	498	2.52	0.32	
Urine stain at weaning (US-wean)				
BSR-A	356	1.51	0.75	P<0.01
BSR-B	309	1.60	0.86	
BSR-C	267	1.77	0.92	
RM [a]	498	1.34	0.77	

[a] RM line not directly comparable; NS = not significant (P>0.10)

Table 2 shows that the average number of total breech strikes from birth to hogget shearing in unmulesed sheep. Although a clear trend exists, the differences between the three BSR lines were not significant by fitting a Poisson model in spite of the fact that the overall incidence of breech strike from post-weaning to hogget shearing differed significantly between lines. The average breech strike incidence of the RM was 0.24 which is not different from the BSR lines.

Fixed Effects. Within the BSR lines, year of birth affected all traits significantly while property of origin had a significant effect on WCOL, WSB and BC. Females had significantly (P<0.01) higher breech cover scores while rams had significantly (P<0.01) more dags.. Significant differences (P<0.05) were found between the BSR lines for WSB, BCVR, WCOL and US (Table 2) while no significant differences were found between lines for DS. In general there was a progressive increase in all the indicator traits from BSR-A to BSR-B to the BSR-C line. The average scores and standard deviation of WSB, BC, DS, WCOL and US at weaning are shown in Table 2 for the different lines. The amount of variation for WSB was quite large (>50%) in the BSR line while the RM line showed much less variation because it was much plainer.

Table 3. Regression coefficients and standard errors (SE) of total number of breech strikes against the independent variables in different years for the BSR and RM flocks.

Covariate	2005 BSR	2006 BSR	2007 BSR	2008[a] BSR	2008 RM
Mu	0.72 (4.492)	-1.73 (5.359)	4.80 (7.543)	-0.10 (0.168)	0.00 (0.207)
WSB	0.09 (0.309)	-0.02 (0.227)	0.17 (0.287)	0.43 (0.489)	0.28 (0.235)
BC	0.07 (0.207)	0.23 (0.453)	**0.72 (0.404)**	0.05 (0.169)	0.00 (0.154)
DS	0.09 (0.523)	0.25 (0.196)	0.26 (0.227)	**0.36 (0.114)**	**0.24 (0.109)**
WCOL	0.13 (0.243)	--	--	0.01 (0.135)	0.10 (0.168)
US	0.11 (0.387)	--	--	0.12 (0.112)	0.16 (0.114)
GFW	-0.30 (1.271)	-0.11 (1.320)	-0.08 (2.197)	--	--
CFW	0.33 (1.734)	0.11 (1.889)	-0.02 (2.901)	--	--
FD	-0.00 (0.066)	0.03 (0.078)	-0.03 (0.071)	--	--
CV	0.00 (0.032)	0.02 (0.037)	-0.05 (0.063)	--	--
YLD	-0.01 (0.056)	0.01 (0.061)	-0.02 (0.096)	--	--
CRV	0.00 (0.007)	-0.00 (0.011)	-0.02 (0.012)	--	--
Deviance	-13.3	-21.1	-13.2	-57.0	-55.6
Degrees of freedom	217	122	72	429	456

Coefficients in bold are significant (P<0.05)

Table 3 shows the association between the incidence of breech strike and the indicator traits in different years and flocks. The models fitted the data reasonably well as shown by the deviance. In 2008 only DS was significantly (P<0.01) related to the incidence of breech strike in both the BSR and the RM flock, while only BC was significant (P<0.05) in 2007. No trait was significant in 2005 and 2006, and none of the wool traits contributed significantly to breech strike in any year.

Previous studies (Seddon, 1931) have shown a relationship between wrinkle score and breech strike and between dags and breech strike (Scobie *et al.* (2002). This study shows that in this environment dags is the most important predisposing factor, and it appears that under different conditions other traits may be more important.

Table 4 shows that breech strike and the indicator traits all have moderate heritability estimates, but the results are not very precise due to the large standard errors. Except for the negative phenotypic relationship between US and DS all the other indicator traits are genetically and phenotypically positively correlated with each other, and with breech strike. WCOL had the strongest phenotypic relationship with breech strike followed by DS, WSB. US and BC. Genetically the strongest relationship exists between DS and breech strike followed by US and breech strike. However the standard errors are again relatively large.

Table 4. Heritability of (diagonal), and genetic (below diagonal) and phenotypic (above diagonal) correlations between the indicator traits and incidence of breech strike

	WSB	BC	DS	WCOL	US	Breechstr
Vp	0.67	0.35	0.49	0.23	0.85	0.25
WSB	**0.45 (0.28)**	0.18	0.18	0.35	0.35	0.22
BC	0.19 (0.51)	**0.42 (0.32)**	0.06	0.15	0.25	0.17
DS	0.07 (0.53)	0.27 (0.65)	**0.55 (0.30)**	0.15	-0.35	0.23
WCOL	0.31 (0.18)	0.66 (0.25)	0.09 (0.55)	**0.49 (0.32)**	0.10	0.26
US	0.44 (0.22)	0.23 (0.23)	0.33 (0.18)	0.00 (0.10)	**0.49 (0.32)**	0.19
Breechstr	0.23 (0.63)	0.17 (0.19)	0.86 (0.17)	0.25 (0.24)	0.53 (0.22)	**0.57 (0.28)**

Figure 1 shows the raw incidence of breech strike of progeny groups of different sires over three years. Highly significant differences (P<0.01) were found between sires for incidence of breech strike that varied from about 2% to 55% animals struck per progeny group. The sire solutions from the Gibbs analyses were regressed against the raw breech strike incidence of the progeny groups . The relationship is shown in Figure 2 and the traits had a correlations of 0.64..

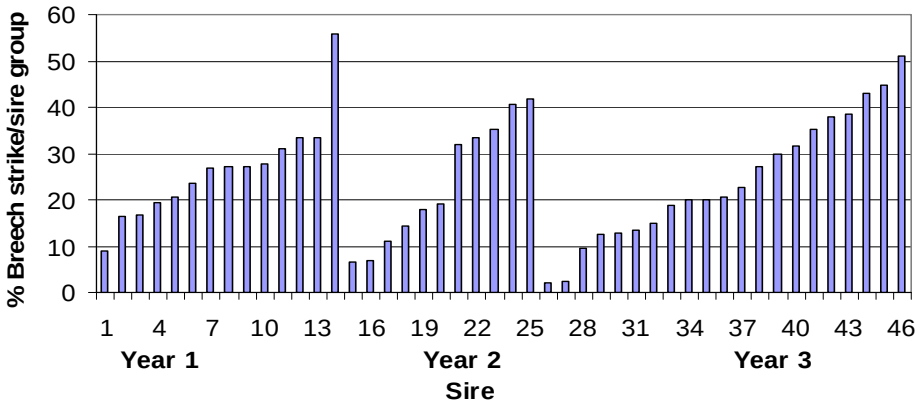


Figure 1 Differences in breech strike resistance between the progeny of 46 sires over 3 years.

CONCLUSIONS

The results show that considerable variation exists amongst sires that can be exploited to improve resistance to breech strike through breeding. While all the visually scored indicator traits are genetically correlated with breech strike, dags followed by breech cover were the most important indicator trait of breech strike in this environment. It appears that the relationships between the "known" indicator traits and breech strike are more complex and that these may vary between years. This indicates that environmental factors will be the key determining factor and more information is needed to elucidate these relationships and to estimate robust genetic and phenotypic parameters for the design of efficient breeding programs.

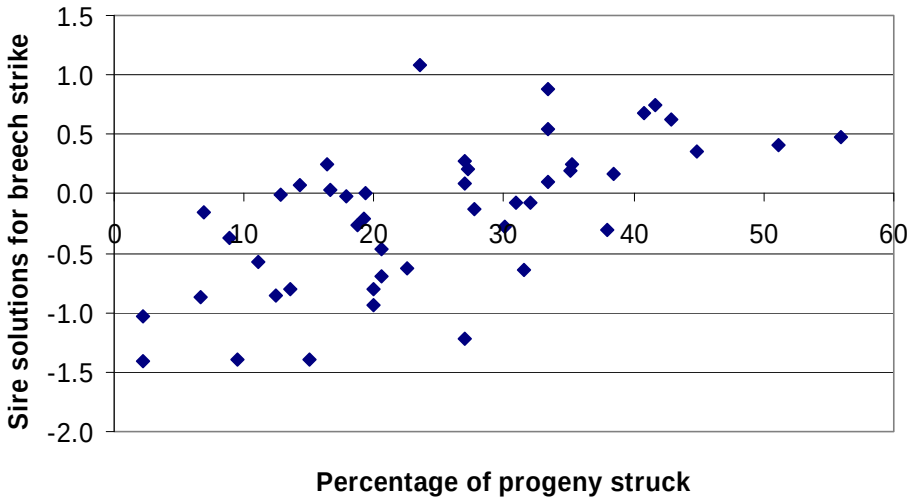


Figure 2. Relationship between sire solution and percentage progeny struck per sire progeny group

ACKNOWLEDGEMENTS

Funding for this study was provided by Australian Wool Producers through Australian Wool Innovation (Ltd) and the Department of Agriculture and Food Western Australia. Contributions from 10 sheep breeders are also gratefully acknowledged.

REFERENCES

AWI (2007) 'Visual Sheep Scores' AWI and MLA

Atkins K.D. and McGuirk B.J. (1979) *Wool Tech. Sheep Breed* **27**.

Cameron, A.W. (1999) Unwelcome partner. Hippo Books: New South Wales

James P.J. (2008) *Aust. J. Exp. Agric.* **46**:1.

Edwards N.M., Hebart M. and Hynd, P.I. (2009) *Anim. Prod. Sci.* **49**:56

Gilmour A.R., Gogel B.J., Cullis B.R., Thompson R. (2006) 'ASReml User Guide Release 2.0.' (VSN International Ltd: Hemel Hempstead).

Karlsson, L.J.E. and Greeff J.C. (2006) *Aust. J. Exp. Agric.* **46**:809.

Mortimer S.I. (2001) Proc. FLICS, flystrike and lice IPM conference. Ed S Champion. University of Tasmania, pp 406.

Misztal I., Tsuruta S., Strabel T., Auvray B., Druet T. and Lee D.H. (2002) In 7th World Conference on Genetics Applied to Livestock Production, Montpellier, 19-23 August 2002.

Raadsma H.W., Gray G.D. and Woolaston R.R. (1997) The Genetics of Sheep. Ed. L. Piper and A. Ruvinsky, CAB international.

Raadsma, H.W. and Rogan, I.M. (1987) In 'Merino Improvement Programs in Australia', p. 321, editor. B.J. McGuirk, Australian Wool Corporation, Melbourne

Scobie D.R., O'Connell D.O., Bray A.R. and Cunningham P. (2002) *Proc. Aust. Soc. Anim. Prod.* **24**:

Seddon H.R., Belschner H.G. and Mulhearn C.R. (1931). Dept. Agric. NSW Sci Bull. No 37.

Smith J.L., Brewer H.G. and Dyall T. (2009) *Proc. Assoc. Advmt. Anim. Breed. Genet.* **18**:334.

Proc. Assoc. Advmt. Anim. Breed. Genet. 18:278-282

SYSTEMS GENETICS ANALYSIS REVEALS GENE MODULES AND HERITABLE BIOMARKERS FOR SHEEP INTESTINAL PARASITE RESISTANCE

H. N. Kadarmideen[1], N. Andronicos[2] and N. S. Watson-Haigh[3]

CSIRO Livestock Industries, [1] JM Rendel Laboratories, Rockhampton QLD 4702, [2]FD McMaster Laboratories, Armidale NSW 2350, [3]Queensland Bioscience Precinct, St.Lucia QLD 4067

SUMMARY

Systems genetics methods were applied to microarray gene expression profiling data from a sheep gastrointestinal nematode (GIN) challenge experiment that was designed to detect genes associated with resistance to GIN. Analyses went beyond just detecting and annotating differentially expressed (DE) genes. This included detection of co-expressed (CE) gene modules associated with the duration of infection, essential hub genes, functional enrichment and pathway analyses. Results revealed that DE genes were highly enriched in functions such as cell-mediated and humoral immunological response to GIN. Further, heritabilities were estimated for expression phenotypes of such candidate biomarkers (range 0.05 to 0.9 with high s.e.) indicating their potential for expression-assisted selection. Hence, the systems genetics method is a key step in identifying biologically relevant and heritable genes/biomarkers amongst several sets of DE genes. This approach would provide specific targets for breeding and therapeutic interventions.

INTRODUCTION

This study is based on a sheep GIN challenge experiment that was designed to examine the microarray gene expression profiles obtained from sheep genetically resistant to parasitism to identify candidate genes involved in the resistance response. Identification of differentially expressed (DE) genes using standard methods do not reveal the complex interactions between individual genes (whether they are DE or not) in a given set of biological perturbations. An intuitive way would be to identify which genes are co-expressed (CE) instead of which genes are DE. From microarray data, a strong correlation of gene expression for a pair of genes implies that both genes act within a common functional group and are under similar transcriptional control. Genes with high level of connectivity with other genes in the CE network act as major hubs that are essential for important biological functions (e.g., Weston *et al.* 2008). With this as a basis, we applied a more general weighted gene co-expression network analysis (WGCNA) method for network construction following Zhang and Horvath (2005). The heritability (h^2) of gene expression is a powerful indicator to determine if expression profiles of regulatory hub or candidate genes in individuals can be used in *expression assisted selection* (Kadarmideen *et al*. 2006). In this study, we estimate heritabilities for candidate biomarkers or hub genes revealed by integrated analyses.

MATERIALS AND METHODS

Helminth naïve sheep (n=32) from the resistant line of the *T. colubriformis* (Tc) selection flock (TSF), which are genetically resistant to GIN infections were used in the experiment. They were given a single oral challenge of either 8000 *Haemonchus contortus* (Hc) or 20000 *Trichostrongylus colubriformis* (Tc) nematodes to examine the host gene expression response of resistant lambs to a primary challenge with pathogenic GIN. All animal experiments were approved of the CLI Animal Ethics Committee. For the Hc challenge, abomasal tissue (site of infection response; HcA) and blood (systemic response; HcB) were taken; and for the Tc challenge, jejunum tissue (site of infection response; TcG) and blood (systemic response; TcB)

were taken at 0, 3, 7 and 21 days post-challenge (length of infection, L) with 4 biological replicates for each sample (4 parasite-tissue groups x 4 time points x 4 replicates = 64 arrays total). RNA extracts were hybridised to Affymetrix GeneChip® Bovine Genome arrays.

Differential Gene Expression Analyses. Microarray data were processed using Affymetrix® Microarray Suite (MAS 5.0) software during normalisation and making detection calls. After editing based on 'Absent or Marginal' detection calls, there were 16936 transcripts for further analyses. All statistical and bioinformatic analyses were performed using the Bioconductor package, limma (v.2.10.5). DE genes at 1% and 5% adjusted False Discovery Rate (pFDR) for a range of biological contrasts were detected. Total number of genes that were DE for each contrast was reported, along with their mode of regulation and functional annotation.

Weighted gene co-expression network analyses (WGCNA). The WGCNA distinguishes from other unweighted network construction methods in that it assigns weights to each edge by soft rather than hard thresholding. The latter could result in some borderline wherein important genes drop-off the network due to cut–off threshold values. WGCNA begins with calculation of a co-expression measure, the absolute value of the Pearson correlation (ρ_{ij}) between every pair of genes. A network is represented by an adjacency matrix, **A** with elements a_{ij}. The **A** matrix is created by applying the adjacency function,β, to co-expression measure, $\rho_{ij,}$ as: $(\rho_{ij})^{\beta}$, which is a kernel of weighted co-expression gene networks (Zhang and Horvath, 2005). β must be specified in such a way that it corresponds to biological motivation. The adjacency matrix **A** is then used to calculate the topological overlap measure (TOM), which reflects the relative interconnectedness between every pair of genes. The TOM matrix was used as input to the average hierarchal linkage algorithm (unsupervised) to create a dendrogram. Modules were then defined using the dynamic hybrid tree cutting algorithm of the *dynamicTreeCut* R package (Langfelder *et al.* 2008). We applied the above WGCNA method for the 16936 transcripts across 16 (×4) arrays. The co-expression networks were constructed separately for Hc and Tc parasites, and for each one of the 3 tissues (ie., one each for HcA, HcB, TcG and TcB samples using 16 microarray samples each). As there are several modules in each network, significant correlation of gene modules with the trait, L (called "module significance") was calculated within each network to determine biological relevance with pathogenicity of Hc and Tc worms.

Systems biology of gene modules. Gene modules with high correlation with L>0.8, were subjected to a systems biology approach to provide a biological context. Within such significant modules, genes whose module membership was < 0.7 (ie., those that are not hub genes) were removed. Functional enrichment analyses were carried out by linking the filtered genes with external biological metadata using EASE (http://david.abcc.ncifcrf.gov/summary.jsp) and GOEAST (http://omicslab.genetics.ac.cn/GOEAST/) to choose candidate genes / biomarkers.

Heritability of DE and module hub genes. Expression profiles of the top 10 DE genes from contrasts and selected hub genes from 4 networks were then prepared for quantitative genetic analysis. A linear model fitted (10) sire effects as random and treatment effects such as time, tissue, flock as fixed, to log-transformed gene expression values (Kadarmideen 2008). This yielded sire variances (σ^2s) which were then used to calculate h^2 as $4*\sigma^2s / (\sigma^2s+\sigma^2e)$.

RESULTS AND DISCUSSIONS

Differential Gene Expression Analyses. Comparison between Hc and Tc were made using only the blood sample arrays as the other tissues are affected by Hc or Tc, but not both. Only results

from fitting linear contrasts for L, irrespective of Hc or Tc and contrasts for interaction between parasites and L are discussed here. Figure 1 (left) shows a Venn diagram containing sets of DE genes (Benjamini and Hochberg (BH) adjusted p-value ≤ 0.05 and fold change ≥ 2) for L3 (20 genes), L7 (132 genes) and L21 (23 genes) compared to uninfected samples (L = 0). Of these, there were 2, 113 and 14 DE genes specific to L3, L7 and L21, respectively. The interaction between parasite and L3, L7 and L21 provide 0, 213 and 54 DE genes respectively (Figure 1 right). Of these interaction genes, 200 were parasite:L7 specific and 41 were parasite:L21 specific. This shows that there are a large number of genes at L7 and L21 that respond differently between Hc and Tc infected sheep. Heat plots show hierarchical clustering of expression levels of the Parasite:L7 (Figure 2). The parasite: L7 heat plot shows 2 distinct gene clusters: 1) those more up regulated in Tc L7 and 2) those more up regulated in Hc L7. This indicates the presence of unique co-expressed gene modules for different parasites across different time points; which can be identified via WGCNA.

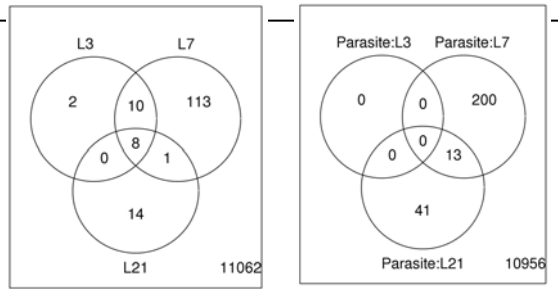


Figure 1. Venn Diagrams with number of DE genes for length of infection. Heatplots show hierarchical clustering of Parasite:L7 (top right) and Parasite:L21 genes (bottom right)

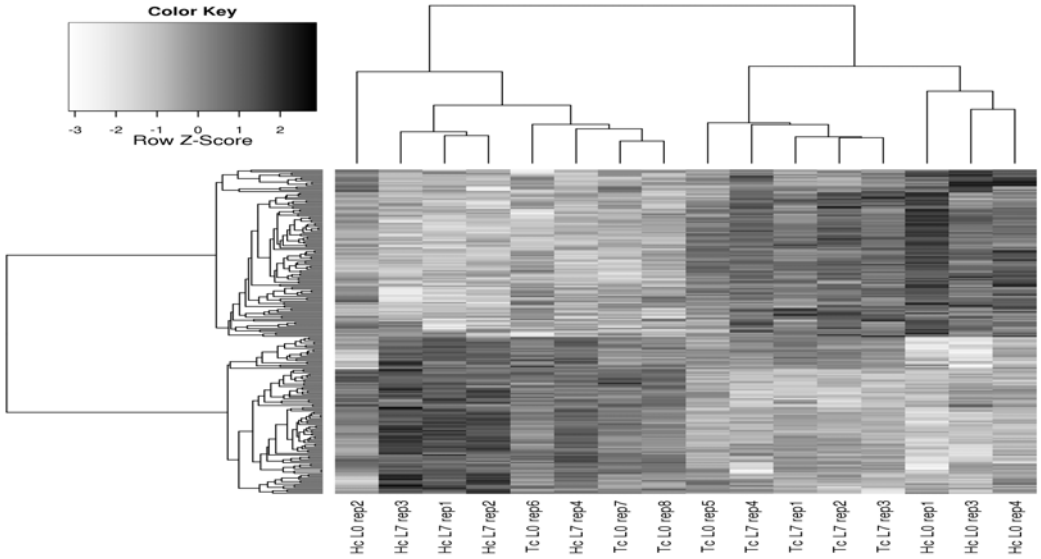


Figure 2. Heatplots show hierarchical clustering of Parasite:L7 interactions

WGCNA and Systems Biology. There were 11 different modules detected for HcA with two major modules (containing 41 and 94 genes, respectively) being the most significantly related to L. Similar results were available for other 3 networks (HcB, TcG and TcB). Systems biology results for different networks showed that they were involved in the activation of immune response, response to antigenic stimulus, cell mediated and humoral immune response, response to wounding, regulation of endocytosis, metabolic and catabolic processes, and signal transduction

among several dozens of other functions. Several of these genes were represented in Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways relating to regulation of autophagy, Notch signalling pathways and glycolysis. Using this systems biology approach we were able to target functionally relevant genes which are candidate drug targets, biomarkers as well as candidates for selection.

Heritability of DE and module hub genes. Heritabilities across a range of DE contrasts and hub genes from WGCNA indicated that about 50% of genes have no significant additive genetic variation in expression while the rest have the estimated h^2 ranging from 0.05 to 0.9 (with high s.e. range 0.09- 0.8). Figure 3 shows h^2 for the top 9 significant DE genes for 21 days of infection with Hc or Tc worms compared to uninfected samples (L0 v L21). Candidate genes with high h^2 of expression may be used in expression assisted selection (Kadarmideen et al., 2006).

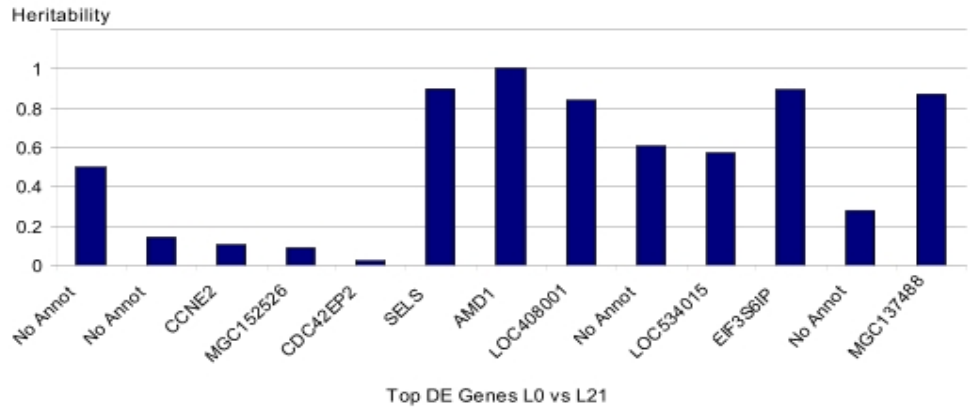


Figure 3. Heritability of Gene Expression of Top 10 DE genes after 21 days of worm infection

CONCLUSIONS

From these studies we conclude that the systems genetics method can be used to identify candidates for biologically relevant and heritable genes/biomarkers from a set of DE genes. This approach has provided a shortlist of a small number of genes for validation, with the eventual aim of developing targets to assist breeding decisions and to design bio-medical interventions.

ACKNOWLEDGMENTS

This project was financially supported by SheepGenomics, an initiative of Australian Wool Innovation Limited and Meat and Livestock Australia. The authors express appreciation to Ross Windon and Peter Hunt for their contribution to FG3 project design.

REFERENCES

Kadarmideen, H.N., von Rohr, P. and Janss, L.L.G. (2006) *Mamm. Genome* **17**:548.
Kadarmideen, H.N. (2008). *IET Systems Biology* **2:**423.
Langfelder, P., Zhang, B. and Horvath, S. (2008) *Bioinformatics*. **24**:719.
Weston, D.J., Gunter, L.E., Rogers, A. and Wullschleger, S.D. (2008). *BMC Systems biology,* ***2*:**16.
Zhang, B., and Horvath, S., (2005). *Stat. Appl. in Genetics and Mol.Biol*. **4**:17.

Proc. Assoc. Advmt. Anim. Breed. Genet. 18:283-286

ON THE EXPRESSION PROFILE OF CANDIDATE GENES CONFERRING RESISTANCE TO GASTRO-INTESTINAL NEMATODES IN SHEEP

Shivashankar H Nagaraj [1], Antonio Reverter [1], Moira Menzies [1], Nick Andronicos[2], and Aaron Ingham [1]

[1]CSIRO Livestock Industries, 306 Carmody Road, St. Lucia, Qld 4067, Australia
[2]CSIRO Livestock Industries, Armidale, NSW 2350, Australia

SUMMARY

We report on the integration of a number of RT-PCR expression datasets totalling 102 genes across 49 experimental conditions of relevance to gastro-intestinal nematodes in sheep. The experimental conditions include three challenge trials, six tissues, five genotypes and two parasites surveyed at multiple time points. After mining gene interaction data for these 102 genes, we have generated an interaction network that enabled us to study these genes in a biologically relevant, system context. Following a systematic investigation of this network, we have identified a number of genes encoding extracellular matrix proteins that may be useful biomarkers enabling identification of parasite resistant sheep.

INTRODUCTION

Selective breeding of sheep that are genetically resistant to nematode infection may be used to overcome problems associated with anthelmintic resistance (Waller, 2006). This strategy would be greatly enhanced if accurate tests were available for attributing resistance. Differential expression of genes or their encoded protein products between resistant (RES) and susceptible (SUS) animals provides one option for such tests. For this reason, we have generated gene expression profiles in RES and SUS animals before and after nematode challenge, demonstrating the impact of time, parasite species, multiplicity of challenge and tissue.

In this study, we combine expression data from various studies from our laboratory and apply data mining techniques to publicly available human data to generate a gene network that may be used to determine how the performance of one gene might be informed by others in the network and in so doing determine an optimal selection of genes that are predictive of phenotype. We also intend to determine the location of our key candidates within a previously determined protein interaction network. We overlay a series of gene attributes into the network, including regulatory and extracellular component, allowing us to efficiently identify targets that are likely to be measureable in the blood, thereby making good candidates for biomarker assays.

MATERIALS AND METHODS

We use the gene expression data from the set of 76 candidate genes related to nematode resistance in sheep reported by Ingham *et al.* (2007). The set comprised of genes differentially expressed (DE) between RES and SUS animals. To these data, an additional set of 26 genes was incorporated from studies in further trials with more biological replicates and tissues. In total, 102 genes across 49 experimental conditions were represented in the combined data set including 2 parasites, 3 trials, 5 flocks, 6 tissues, and 16 time points. The experimental layout of the resulting data set was that of an incomplete block design with only two genes (GAPDH and RPL) represented across the 49 conditions.

The entire data set was normalized by fitting a mixed ANOVA model with threshold cycle in the PCR reaction as the measure of expression abundance as the dependent variable. Main design effects were treated as fixed effects, while gene and gene by condition interaction were treated as

random effects. PermutMatrix (http://www.lirmm.fr/~caraux/PermutMatrix/EN/index.html) was used to perform cluster analysis of the normalized expression data across rows (genes) and columns (conditions).

Finally, we downloaded a set of 55,606 true positive interactions among 7,197 human genes that were defined from functional studies (Franke *et al.*, 2006). This interactions dataset was built including 2,788 confirmed, direct, physical protein-protein interactions derived from the Biomolecular Interaction Network Database (BIND; http://bind.ca), 18,176 confirmed human protein interactions from the Human Protein Reference Database (HPRD; http://www.hprd.org/), 22,012 direct functional interactions from the Kyoto Encyclopedia of Genes and Genomes (KEGG; http://www.genome.jp/kegg), and 16,295 interactions derived from Reactome (http://www.reactome.org).

RESULTS AND DISCUSSION

Figure 1 shows the result of the hierarchical cluster analysis of the expression profile of 102 genes across the 49 conditions. Expression results for 76 genes have been reported previously by Ingham *et al.* (2007). To this dataset we have added 6 additional genes studied in the same trial but not reported along with a further 20 genes analysed in third independent experiment. See supplementary data for details. To avoid biases in the generation of clusters, the average gene expression was imputed in the cells corresponding to genes not surveyed in a given condition. These are represented as black cells in Figure 1. The cluster of columns, corresponding to experimental conditions, reveals the tissues as being the biggest hierarchy in the clusters, followed by parasite and time point. The cluster of rows, corresponding to genes, reveals groups of genes enriched for specific functions. Across the rows, three main clusters are clearly distinguishable including toll-like receptors, cytokines, and proteases.

The mining of the interaction dataset resulted in a network with 703 nodes (genes) connected by 1,090 edges (interactions) that was visualized using Cytoscape (http://www.cytoscape.org). The entire view of the network is given in Figure 2. Interestingly, some of the clusters observed in Figure 1 and derived from gene expression data, retained their integrity in the network. For instance, TLR pathway members form a cluster in Figure 1 and a sub-network in Figure 2. The network file in Cytoscape format used to generate Figure 2 is available for download from our public website (http://www.livestockgenomics.csiro.au/courses/Shiv_AAABG09.html).

One limitation of using gene expression data as a biomarker is that the expression profile of the gene and subsequent location of its encoded product are often tissue specific. In this case, DE genes distinguishing RES and SUS animals in gut tissues might only be detectable in these tissues. Sampling gut tissue is highly invasive and therefore not appropriate for assaying in a routine manner. Instead, extracellular products are easily sampled in the blood. For this reason we performed a systematic investigation of this network and identified a number of genes encoding extracellular matrix protein candidates, associated with our DE candidates. The DE genes KCNJ15 and DYRK3 both distinguish RES and SUS animals in the gut, but these factors are intracellular. Through cluster analysis these genes were shown to interact with the extracellular factors IL16 and NT5E respectively, making these candidates potentially useful targets for biomarker development. Taken together, the results illustrate the benefits of integrating gene expression data, together with interaction networks to study genes involved in complex biological processes/signalling pathways associated with parasite resistance in sheep.

FUTURE DIRECTIONS AND CONCLUSIONS

Although gene expression data offers a list of DE genes, and interaction networks supply a list of dual connections, the nature of the information that can be drawn from either approach is relatively

limited. Transcript abundance alone reveals little about the mechanisms underlying the observed changes or details of different simultaneously occurring events.

In order to understand the genetic basis for ovine resistance to diseases, we need to gain insights into the dynamic regulatory nature of the immune response. Here we have shown how a network systems approach can be used to successfully inform the selection of potential biomarkers. Moving forward, we plan to capture the dynamism inherent in immune responses by

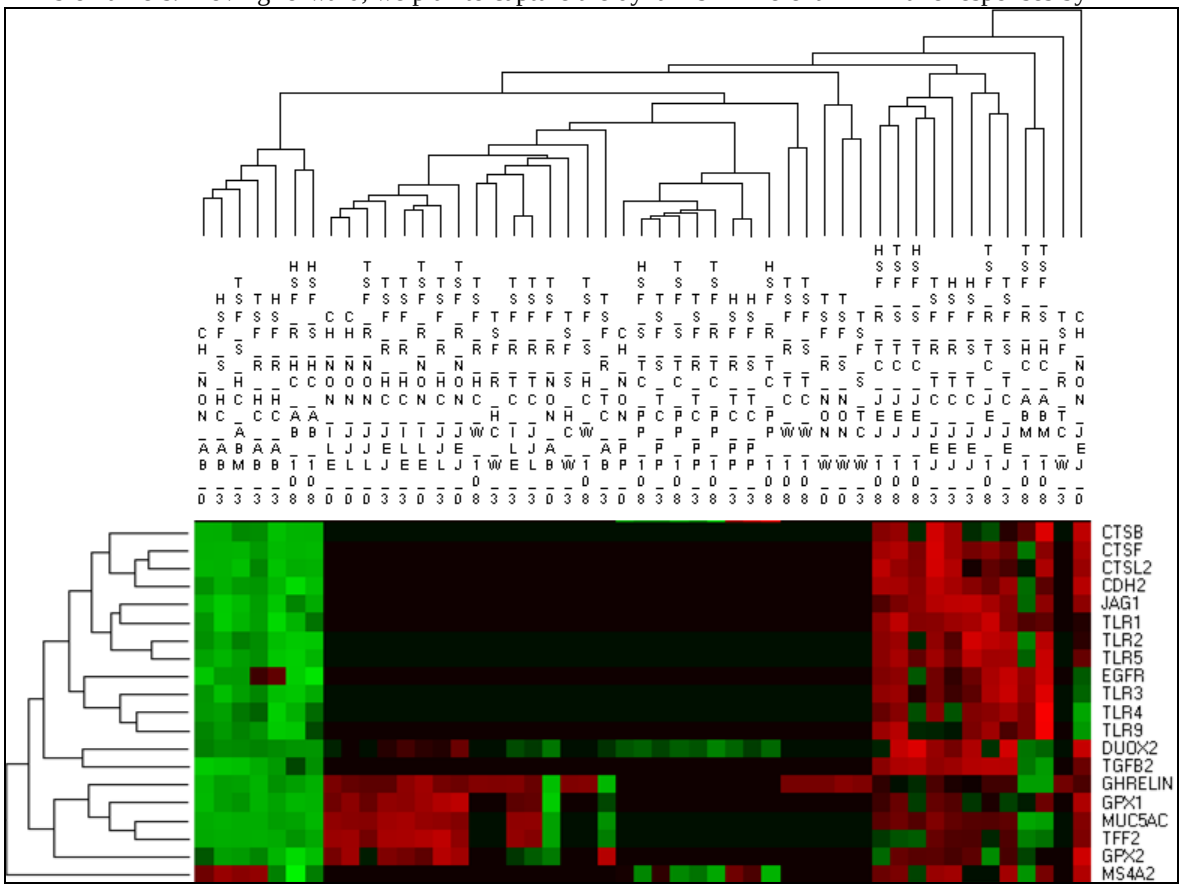


Figure 1. Hierarchical cluster analysis of 102 genes (rows) by 49 experimental conditions (columns) derived from parasite resistance studies with sheep.

A snapshot of representative cluster comprising TLR pathway genes has been shown; complete version of figure 1 containing all the clusters is available for download from our public website (http://www.livestockgenomics.csiro.au/courses/Shiv_AAABG09.html). The spectrum goes from green to red for low and high expression, respectively. Abbreviations used in column headings are as follows; Animals (TSF Trichostrongylus Selection Flock; HSF, Haemonchus Selection Flock; Chiswick, Outbred commercial sheep); Genetic Resistance (R, Resistant; S, Susceptible); Infection status (HC, Challenged with *Haemonchus contortus*; TC, Challenged with *Trichostrongylus colubriformis*; NONE, Sheep not challenged); Tissue sampled (ABOM, abomasum; ILE, ileum; JEJ, jejunum, WBC, white blood cells; JEJILE, junction of jejunum and ileum; PP, Peyer's Patch); Time of sample collection relevant to challenge (0, 3, or 108 days).

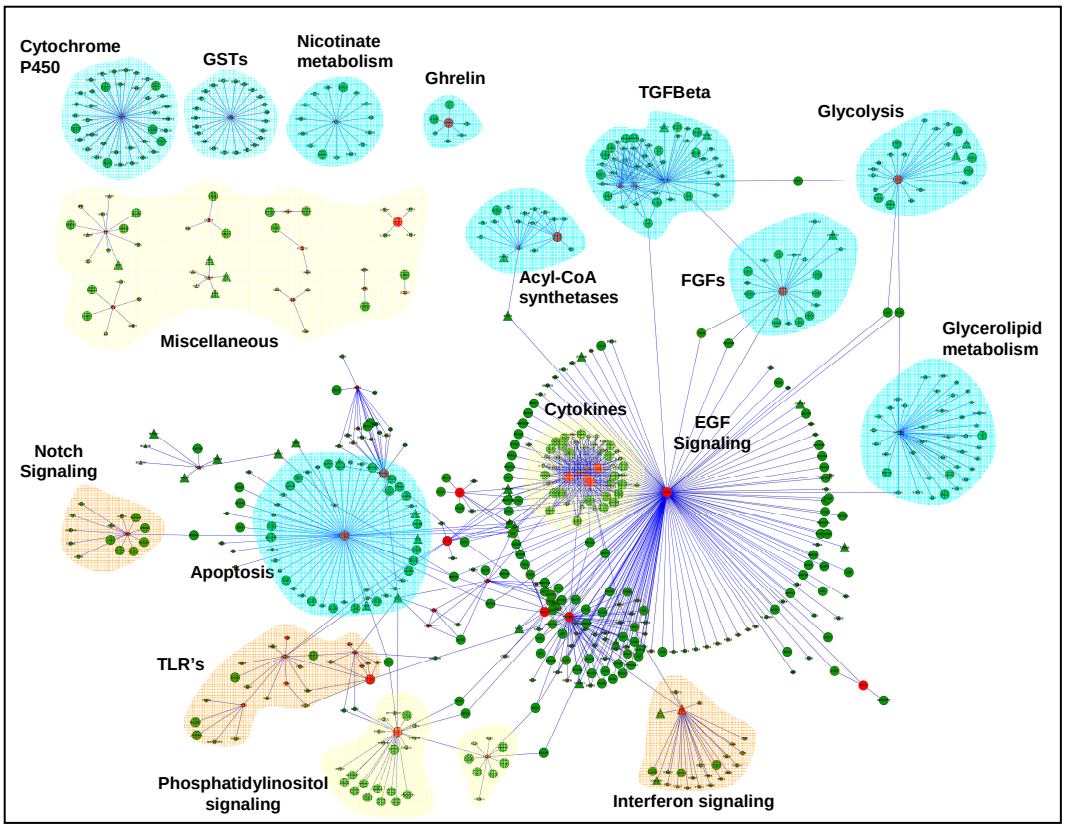


Figure 2: Network with 703 nodes (genes) connected by 1,090 edges (interactions). Visualised attributes included: genes surveyed in the RT-PCR assays were coloured red (otherwise green); transcription factors as triangles (otherwise, circles); big symbols were assigned to post-translational modifiers; red borders for secreted proteins; and sub-networks are highlighted as clouds.

studying gene co-expression networks overlaid with transcription regulation associated with regulatory events. These studies are crucial, given the current, serious resistance problems in parasites against most treatments, and residue problems in meat, milk and the environment.

ACKNOWLEDGEMENTS

SHN is grateful to CSIRO for the award of an OCE Post-Doctoral Fellowship. The financial support of the CSIRO Transformational Biology Initiative is gratefully acknowledged. The authors are grateful to SheepGenomics for financial support.

REFERENCES

Franke, L., van Bakel, H., Fokkens, L., de Jong, E.D., Egmont-Petersen, M. and Wijmenga, C. (2006) *Am. J. Hum. Genet.* **78**:1011.

Ingham, A., Reverter, A., Windon, R., Hunt, P. and Menzies M. (2007) *Int J Parasitol*. **38**:431

Waller, P.J. (2006) *Vet. Parasitol.* **139**:1

Proc. Assoc. Advmt. Anim. Breed. Genet. 18:287-295

CURRENT STATUS AND FUTURE PROSPECTS FOR REPRODUCTIVE TECHNOLOGIES IN SMALL RUMINANTS

W.M.C. Maxwell and G. Evans

Faculty of Veterinary Science, University of Sydney, NSW 2006

INTRODUCTION

Reproductive technologies such as artificial insemination (AI) and multiple ovulation and embryo transfer (MOET) have been technically feasible and commercially utilized in small ruminants for many years (reviewed by Betteridge 1981; Evans 1991; Salamon and Maxwell 2000). Sheep and goats are major contributors to the domestic economy of many nations, but neither AI nor MOET has been widely adopted into normal breeding programs as, for example, has been the case for AI in dairy cattle. This paper outlines the costs and benefits of applying reproductive technologies in small ruminants, assesses the current technology and explores potential practical developments. Recent technological developments regarding *in vitro* embryo production, particularly in juveniles, and the incorporation of sexed sperm into artificial breeding programs are described. Reference is made to other livestock species, particularly cattle, where contrast is appropriate, and brief conclusions are drawn on the most likely technologies to be used in the commercial sector in the near future.

THE COSTS AND BENEFITS OF USING REPRODUCTIVE TECHNOLOGIES

The slow commercial adoption of AI and MOET is not because the technologies do not have the potential to facilitate genetic improvement programs; indeed the advantages of AI (Abbott 1994; Windsor and van Bueren 1994; Nicholas 1996) and MOET (Smith 1986; Wray and Goddard, 1994a,b; Nicholas 1996) are well documented. The reasons are the relatively high cost of these technologies compared with the realisable short-term benefits (Windsor and van Bueren 1994). Moreover, in most countries where sheep and goats are maintained in large numbers, the relative value of individual animals is low compared with the cost of the technology. In Australia in the early 1990s, the average cost of breeding sheep to AI with fresh semen (cervical AI) was calculated as 3 times that of natural breeding, and with frozen semen and (laparoscopic) AI about 12 times (McClintock and Nicholas 1991; Abbott 1994); for breeding through MOET the cost was raised over 60 times. However, it may be possible to reduce these costs using new developments from MOET, such as juvenile *in vitro* embryo transfer (JIVET) and mature *in vitro* embryo transfer (MIVET), as these techniques can be applied to very young animals, can yield more embryos per donor than MOET, are repeatable more than once on each donor and, because they rely on *in vitro* fertilisation, can utilise expensive sperm very efficiently. Few commercial wool or prime lamb producers see value in using such technology but ram breeders and studs, which sell a more valuable product (breeding animals), are able to gain from use of expensive artificial breeding. This is particularly the case for advanced MOET technologies which are likely only to be applied at the elite level. Nevertheless, the MOET and AI technologies have been widely applied for the importation and dissemination of genetic material from new breeds, particularly in recent times for improvement of meat production from sheep and goats in Australia.

For sheep AI, fresh semen is simple and effective but there are obvious restrictions on its use over time and distance. In Australia, there has been a small but continuing interest in fresh semen as a cheap alternative to laparoscopic AI, since “in house” labour is used and discounted. More widespread use of valuable sires requires storage of semen. Liquid storage may extend the useful lifetime of semen, but ultimately frozen semen is required to fully

overcome restrictions of time and distance. The cost of using frozen semen successfully has, however, restricted its widespread use, and animal welfare concerns in some countries over use of laparoscopy have compounded the problem. Nevertheless, in situations where animals are of relatively high value, AI is used as the major breeding method. Such a situation occurs in Australia's ram breeding flocks and studs, and the genetic benefits are passed on to subsidiary flocks through sale of rams (Maxwell and Wilson 1990; Abbott 1994).

Similarly, the cost of using MOET in a structured breeding program has restricted its use to research purposes, for the introduction and dissemination of new breeds**,** and to a limited number of high value animals in elite flocks. As in the cattle industry, there is a limited number of examples where MOET has been used in sheep and goats as part of a breeding program leading to increased rates of genetic gain, as opposed to use for entrepreneurial reasons. To our knowledge, only one or two Merino stud flocks have incorporated MOET into a long-term breeding strategy, apparently with some success, but in recent times MIVET and JIVET have been applied on a number of terminal sire breed studs and at least two Merino studs; a number of artificial breeding companies now offer these technologies as part of their services and they are commonly mentioned as available techniques for incorporation in genetic improvement programs (for example, Meat and Livestock Australia 2004).

DEVELOPMENTS IN MOET, JIVET, MIVET AND EMBRYO VITRIFICATION

While the use of MOET in sheep and goats has increased in recent times, there is no evidence of better rates of success from commercial MOET since it was first developed in the 1960s. The only improvements adopted were laparoscopic AI (Killeen and Caffery 1982), which simplified the necessary intrauterine deposition of semen (Trounson and Moore 1974) and allowed the use of frozen semen (Salamon and Maxwell 1995), and pituitary FSH extracts for superovulation (Armstrong and Evans 1984), which overcame some of the problems of using PMSG alone as a superovulatory hormone (Evans and Robinson 1980). A cocktail of PMSG and FSH (Ryan *et al.* 1991), often combined with synthetic gonadotrophin releasing hormone, is commonly used as the most cost-effective regimen in Australia. Even so, choice of superovulatory regime, including various ovine and porcine FSH preparations, treatment schedules and doses, remains one of the more controversial issues in MOET technology. In most commercial situations, the choice of which hormone regime to use in small ruminants remains a matter of personal preference or trial-and-error.

While many factors contribute to the success of MOET, average success rates in properly controlled commercial situations remain disappointingly low, with between 3 and 4 lambs produced from fresh embryos per donor cycle (Maxwell *et al.* 1990; Evans 1991). This in itself does not preclude sufficient numbers of lambs being produced per donor per season, since each donor may undergo several treatment cycles. In this case, surgically-induced adhesions, particularly of the ovary and oviduct, could restrict the number of repeat operations but can be minimised using an embryo recovery technique which does not require handling or exposure of the ovaries and oviducts. Nevertheless, the number of times a donor can be used is limited. Attempts to overcome this problem using laparoscopic embryo recovery techniques (McKelvey *et al.* 1986) have not been widely adopted due to low success rates and the time taken to perform the operation with any degree of success.

A potentially simpler and more productive method of repeated recovery of gametes is by laparoscope-guided oocyte aspiration or ovum pick-up (OPU; reviewed by Tervit 1996; Baldassare *et al.* 2004). The combination of OPU with *in vitro* oocyte maturation (IVM) and fertilisation (IVF), and the *in vitro* culture of the resulting zygotes to transferable embryos, is termed *in vitro* embryo production (IVP) or more recently MIVET. This has been successfully applied to the large-scale production of offspring from transgenic goats (Baldassare and

Karatzas 2004) and is starting to be incorporated into commercial embryo transfer technology for cattle and sheep. MIVET embryos produced with standard IVM/IVF techniques result in normal viable lambs (Crozet *et al.* 1987) and kids (Keskintepe *et al.* 1994), and early problems with unusually large offspring have been largely overcome by use of appropriate culture media *in vitro* (Thompson *et al.* 1995). The efficiency of MIVET has considerably increased in recent years, with OPU repeatable in ewes 3 times or more, yielding 10-15 oocytes per donor per aspiration, with fertilization and blastocyst development rates of 70-80 and 60-80%, respectively (Morton *et al.* 2005). We have also produced lambs from the injection of single sperm into oocytes (intracytoplasmic sperm injection, or ICSI; Catt *et al.* 1996; Gomez *et al.* 1998), though more oocytes/presumptive embryos were lost *in vitro* compared with standard IVF, indicating some subsequent developmental problems associated with this procedure (Gomez *et al.* 1998).

Probably the most promising and exciting of the *in vitro* techniques for adoption into commercial practice is that of MIVET used in conjunction with juvenile donors (JIVET), since a reduction in the generation interval potentially increases the rate of genetic gain in comparison with the use of adult donors (Smith 1986). Early observations that lambs of 4 to 6 weeks of age underwent an unusual natural increase in folliculogenesis (Kennedy *et al.* 1974), which could be further stimulated by exogenous gonadotrophins (Worthington and Kennedy 1979), went unappreciated in a practical sense for many years until the phenomenon was used as a basis for JIVET. This involves superovulation of young calves, lambs, or kids, recovery of oocytes, fertilisation and culture *in vitro,* and transfer to adult surrogates (Armstrong *et al.* 1994; Earl *et al.* 1995). While the viability of prepubertal oocytes is relatively low compared with those obtained from adult animals (O'Brien *et al.* 1996, 1997b), the large number produced can ultimately result in a higher number of offspring than from superovulated adult donors (Armstrong *et al.* 1994). The viability of prepubertal oocytes can also be improved by pre-treatment of the lambs with ovarian steroids (O'Brien *et al.* 1997a). Recent improvements in juvenile donor treatments (Morton 2008) and *in vitro* oocyte maturation (Kelly *et al.* 2008) have brought JIVET close to commercial viability, with limited adoption for both sheep and cattle breeding in Australia. The combination of MIVET or JIVET with sexed sperm provides a powerful technology if single sex offspring are desired (Morton *et al.* 2004a).

One cloud on the JIVET horizon is the finding of alterations in the expression of several developmentally important genes in embryos derived from prepubertal cattle (Oropeza *et al.* 2004) and sheep (Leoni *et al.* 2006). Similarly, we have reported alterations in gene expression in bovine embryos derived from oocytes fertilised *in vitro* with sexed sperm (Morton *et al.* 2007). The implications of these aberrations are yet to be determined but they may contribute to epigenetic effects already reported in the literature (for example, review by Walker *et al.* 2000). On the positive side, two new techniques have great potential for combination with JIVET to enhance the rate of genetic gain. Further reductions in the generation interval could be achieved by incorporating foetal oocytes into IVP systems. However, to date foetal calf oocytes have had lower rates of maturation, fertilization and embryo development than those from adults (Chohan and Hunter 2004). Utilization of gametes from prepubertal males could further dramatically reduce generation interval on the male side. We have been able to initiate spermatogenesis in prepubertal ram lambs by administering PMSG (Morton *et al.* 2004b) and the injection of gametes from prepubertal ram lambs into IVM oocytes from adult and prepubertal lambs has yielded 16–32 cell stage embryos (Morton 2008) but not live offspring.

The most efficient use of MOET often involves cryopreservation of embryos for transport or long-term embryo banking for conservation. We have demonstrated that sheep embryos conventionally frozen for 13 years retained viability (Fogarty *et al.* 1999) indicating that frozen embryos are a safe method of long-term conservation of rare or valuable genetic strains

or breeds. Frozen storage is not routinely used in commercial practice other than for long-distance transport simply because conventional freezing, though successful, is time consuming and expensive. Vitrification of embryos (Rall 1987) is now the method of choice for cryopreservation, as it offers the advantages of simplicity and reduced equipment costs. It is successful for both *in vivo*-produced (Ali and Shelton 1993; Szell and Windsor 1994) and *in vitro*-produced sheep embryos (Evans *et al.* 1999), and for *in vitro* and *in vivo*-produced goat embryos (Traldi *et al.* 1998), though overall success rates in these early reports did not match those of conventional freezing. There has been much improvement in vitrification success over the past 5 years, resulting in commercially applicable methods for cryopreserving embryos in most species (reviewed by Vajta and Kuwayama 2006), including *in vitro* produced porcine oocytes (Liu *et al.* 2008) and embryos (notoriously difficult to freeze because of their high lipid content), even after cloning (Du *et al.* 2008). Vitrification of embryos, whether from MIVET or JIVET, should soon become a simple and routine component of MOET programs.

DEVELOPMENTS IN AI

When performed by experienced operators, AI with fresh ram semen deposited in either the vagina or cervix can usually result in acceptable levels of fertility comparable with that of natural mating (Maxwell and Butler 1984). The methods are simple, can be done "on farm" (Evans and Maxwell 1987) and can make more extensive use of a limited number of superior males than natural mating. However, for fuller appreciation of the benefits of AI, semen is ideally collected and distributed from highly selected males at an AI centre, preferably after careful selection through a sire referencing scheme linked to progeny tests (Windsor and van Bueren 1994; Nicholas 1996). In Australia and other parts of the world, semen is distributed from sheep and goat studs to breeder flocks in other parts of the country, and this requires that semen is stored for transportation, almost exclusively in frozen form, though liquid (chilled) storage is an option that has some advantage in the very short term. Unfortunately for the small ruminant industries, frozen, and to some extent chilled, sperm are not able to penetrate the cervix in sufficient numbers to bring about acceptable rates of fertility in most breeds (Maxwell and Hewitt 1986) and therefore stored semen requires an associated insemination technique which by-passes the cervix. Transcervical insemination is successful in about 60-70% of goats (Leboeuf *et al.* 2000) allowing the commercial utilization of frozen semen (Leboeuf *et al.* 2008). The advent of the laparoscopic AI technique for sheep in 1982 led to an increased use of AI of frozen ram semen in Australia (Evans 1991; Maxwell and Watson 1996) and other parts of the world. Though the proportion of the Australian national flock inseminated in this way is relatively small (1-2 %) it comprises a large proportion (40-50 %) of the ram-breeding ewe population. The small ruminant industries are crying out for improved success rates with cervical AI with frozen or liquid stored semen so that rates of genetic gain can be improved. This can only be achieved through dramatic improvement in semen processing methods or the development of a simple, practical method of non-surgical AI.

LIQUID STORAGE OF SEMEN

Freezing causes cryoinjury, but chilling (to 5°C) has the advantage of reducing the number of sperm deaths. Liquid (chilled) semen can retain acceptable fertility for 24 hours when inseminated via the cervix but thereafter fertility is low unless intrauterine AI is used (Maxwell and Salamon 1993). Since degradation of sperm may involve lipid peroxidation, we have attempted to prolong the life of liquid-stored sperm by addition of antioxidants, specifically superoxide dismutase (SOD) and catalase. In both sheep (Maxwell and Stojanov

1996) and goats (Pomares *et al.* 1995), antioxidants have prolonged the fertilising life of chilled semen to 14 days after intrauterine AI. It seems that chilled semen must be used quickly for cervical AI but laparoscopic AI may allow its utilisation after storage for up to 2 weeks. Liquid storage has been an efficient method for utilising ram semen storage in the past, and remains the method of choice in the New Zealand dairy cattle industry, mainly because of optimum longevity and viability of sperm during storage and after insemination. However, the need for sires to be used nationally or internationally over large numbers of females has meant that, for both sheep and cattle, frozen storage has been the method of choice for utilisation of semen in genetic improvement programs.

FROZEN STORAGE OF SEMEN

It has long been accepted that cervical AI of frozen-thawed semen in sheep cannot consistently produce acceptable fertility, no matter how many motile sperm are inseminated. Despite considerable expense and effort over the past 50 years or more, little progress has been made in developing methods of freezing semen which could result in acceptable fertility after cervical AI (Salamon and Maxwell 1995a). Intrauterine insemination via laparoscopy does, however, result in acceptable or "normal" fertility (Salamon and Maxwell 1995b). Since the advent of this AI method, there has been little progress, and indeed less incentive, to develop better methods of processing semen for cryopreservation. However, recent findings in our laboratory provide hope that frozen-thawed semen may one day be used successfully with cervical AI. The observation that seminal plasma proteins (SPP) could protect sperm through the trauma of flow cytometry (Catt *et al.* 1997) ultimately led us to add them back to diluted frozen-thawed semen, where we found that it arrested the progression of sperm through capacitation-like changes (McPhie *et al.* 1999) and improved motility *in vitro* (Gellatly *et al.* 1999). Ultimately we found that SPP, when added to frozen-thawed ram semen, could produce acceptable levels of fertility in ewes inseminated in the cervix (Maxwell *et al.* 1999), although the results are not consistently high (El-Hajj Ghaoui *et al.* 2007). The active component is a series of proteins that actively repair cryo-damaged sperm membranes (reviewed by Muiño-Blanco *et al.* 2008). Concentrations of these proteins vary with season, breed and among males (reviewed by Maxwell *et al.* 2007).

SEXED SPERM

Sperm sexing by high speed flow cytometry has been one of the most significant new technologies for artificial breeding of livestock developed in the twentieth century. It has been widely applied commercially in dairy cattle, with hundreds of thousands of healthy offspring born to date (Seidel 2009). It should be noted that the rapid adoption of this technology has not necessarily been for the purposes of genetic improvement. AI with sexed sperm has both management and marketing benefits, particularly in the dairy cattle industry, where it is an advantage to produce female replacement heifers from the best cows and male beef progeny from the rest of the herd.

Along with the new sexing technology has come a need to adapt and develop new semen storage and processing methods, so that the sexed sperm can be utilised for AI or for *in vivo* and *in vitro* embryo production. Limitations associated with the sex sorting apparatus (high cost, lack of portability and slow processing rates) have required that sorted sperm are also frozen before use. This has posed particular research challenges, because the sorting process itself subjects sperm to additional stressors, including very high dilution rates, extended time *in vitro*, staining of DNA, centrifugation, and exposure to ultraviolet light, high system pressures and electric charges. The sperm therefore need special protection to survive not only sex sorting but also the freeze-thaw procedures that follow. Despite the apparently harsh treatment of the gametes, the health and

normal reproductive capacity of the large number of offspring born to date attest to the safety of the sex-sorting procedure. Nevertheless, some caution and further research on possible effects of sperm treatments on their DNA are warranted, as increased rates of early and late pregnancy loss have been reported in cattle and pigs, but not sheep, after insemination with sex-sorted frozen-thawed spermatozoa (de Graaf *et al.* 2009) and altered gene expression has been detected in IVP embryos derived from sex-sorted bovine spermatozoa (Morton *et al.* 2007).

Much progress has been made in recent years on preservation of the viability of sperm through the sorting and freeze-thaw processes (reviewed by Maxwell *et al.* 2004; Rath and Johnson 2008), and there have been some unexpected findings, both practical and scientifically interesting, made along the way about the nature and function of ram and bull sperm (reviewed by de Graaf *et al.* 2009). One of these is the discovery that sex-sorting actually selects, apparently by accident, sperm with intact membranes that lack a marker protein for the acrosome reaction (Leahy *et al.* 2008), rendering the selected population more resilient to further processing than unsorted sperm.

In the early 1990s, we demonstrated that fertilisation could be achieved with oviductal insemination of as few as 10,000 sperm in sheep (Maxwell *et al.* 1993), and later obtained high fertilization rates after AI of superovulated ewes with sexed sperm (de Graaf *et al.* 2007a) as a way of maximising the use of their limited numbers. IVF also requires relatively low numbers of sperm and, if sexing is used in conjunction with JIVET, has potential to hasten the rate of genetic gain (Raadsma and Tammen 2005). Therefore, we have produced offspring by IVF with sex sorted-frozen-thawed sperm from *in vitro* matured abattoir-sourced peripubertal lamb oocytes and from oocytes aspirated from hormone-stimulated prepubertal lambs (Morton *et al.* 2004a).

One of the most exciting recent developments in our laboratory has been "reverse" sex-sorting technology for the utilization of frozen ram (Hollinshead *et al.* 2003) and bull semen (Underwood *et al.* 2009a). This allows high purity sorting of frozen-thawed sperm for re-cryopreservation and later use, without a reduction in the fertilizing capacity of the sperm after the two cycles of freezing and thawing. Lambs of predicted sex also have been born after the transfer of both fresh and vitrified IVP embryos, derived from "reverse sorted" ram sperm (O'Brien *et al.* 2004). Furthermore, our work on improved precision in synchronization of ovulation in sheep using GnRH (Reyna *et al.*, 2007) has allowed us to recently confirm the commercial viability of sex-sorted frozen-thawed sperm for artificial insemination in sheep, with a minimum effective AI dose of 1 million sperm (Beilby *et al.* 2009). Furthermore, over a large number of ewes, fertility was not different after insemination by laparoscopy of frozen-thawed (control), sex sorted-frozen-thawed or frozen-thawed-sex sorted-refrozen-thawed ("reverse sorted") sperm (de Graaf *et al.* 2007b). Another recent development has been the application of "reverse sorted" sperm to *in vitro* embryo production in sheep and cattle. In initial studies conducted under commercial conditions in the USA, blastocyst development rates were similar for embryos derived from *Bos indicus* OPU oocytes fertilized with frozen-thawed-sex sorted-refrozen-thawed and control (frozen-thawed) sperm (Underwood *et al.* 2009b). These results demonstrate that frozen-thawed ram and bull sperm can be sex-sorted for either immediate or future use in an IVF system after re-cryopreservation, and point the way to the commercial application of sexed sperm through JIVET or MIVET.

CONCLUDING REMARKS

Major new technologies in embryo production will take time to become commercially viable, though JIVET offers great promise to increase the rate of genetic gain, particularly when combined with sexed sperm and marker-assisted selection in sheep and cattle (Raadsma and Tammen 2005). As far as AI in sheep is concerned, the use of frozen sperm (sexed or unsexed) will be limited in the near future to laparoscopic AI, though our work with seminal

plasma proteins offers the first real prospect of dramatically increasing the use of frozen semen through cervical AI. We have briefly reviewed those technologies which have the most immediate likelihood of application in small ruminant breeding programs. Those omitted - such as somatic cell nuclear transfer (cloning), transgenesis and utilisation of stem cell spermatogonia - are either some years away or, for ethical, animal welfare or epigenetic/safety reasons, are likely to have delayed acceptance by the industry or consumers.

REFERENCES

Abbott, K.A. (1994) *Aust Vet J* **71**: 353.

Ali, J. and Shelton, J.N. (1993) *J Reprod Fert.* **99**:65.

Armstrong, D.T. and Evans, G. (1984) *Proc Int Congr Anim Reprod Artif Insem, Urbana, Ill.* **10**:VII-8.

Armstrong, D,T., Irvine, B. and Earl, C.R. (1994) *Proc Soc Study Reprod.* **27**:538.

Baldassarre, H. and Karatzas, C.N. (2004) *Anim Reprod Sci*. **82**:255.

Baldassare, H., Wang, B., Keefer, C.L., Lazaris, A. and Karatzas, C.N. (2004) *Reprod Fert Dev*. **16**:456.

Betteridge, K.J. (1981) *J Reprod Fert.* **62**:1.

Beilby, K.H., Grupen, C.G, Thomson, P.C., Maxwell, W.M.C. and Evans, G. (2009) *Theriogenology.* **71**:829.

Catt, S.L., Catt, J.W., Gomez, M.C., Maxwell, W.M.C. and Evans, G. (1996) *Vet Rec*. **139**:494.

Catt, S.L., O'Brien, J.K., Maxwell, W.M.C. and Evans, G. (1997) *Reprod Dom Anim*. **32**:251.

Chohan, K.R., Hunter, A.G. (2004) *Theriogenology.* **61**:373.

Crozet, N., Huneau, D., Desmedt, V., Theron, M-C., Szollosi, D., Torres, S. and Sevellec, C. (1987) *Gam Res.* **16**:159.

de Graaf, S.P., Beilby, K.H., O'Brien, J.K., Osborn, D., Downing, J.A., Maxwell, W.M.C. and Evans, G. (2007a) *Theriogenology.* **67**:550.

de Graaf, S.P., Evans, G., Maxwell, W.M.C., Cran, D.G. and O'Brien, J.K. (2007b) *Theriogenology.* **67**:391.

de Graaf. S.P., Beilby, K.H., Underwood, S.L., Evans, G. and Maxwell. W.M.C. (2009) *Theriogenology.* **71**:89.

Du, Y.T., Lin, L., Li, J., Kragh, P.M., Zhang, Y.H., Schmidt, M., Bogh. I.B., Zhang, X., Purup, S., Kuwayama, M., Jorgensen, A.L., Pedersen, A.M., Villemoes, K., Yang, H.M., Bolund, L. and Vajta, G. (2007) *Cloning and Stem Cells.* **9**:469.

Earl, C.R., Irvine, B.J., Kelly, J.M., Rowe, J.P. and Armstrong, D..T (1995) *Theriogenology*. **43**:203.

El-Hajj Ghaoui, R., Thomson, P.C., Leahy, T., Evans, G. and Maxwell, W.M.C. (2007) *Reprod Dom Anim.* **42**:541.

Evans, G. (1991) *Reprod Fertil Dev*. **3**:627.

Evans, G., Rao, K.M., Fogarty, N.M. and Maxwell, W.M.C. (1999) *Proc Aust Soc Reprod Biol*. **30**:130.

Evans, G. and Robinson, T.J. (1980) *J agric Sci Camb*. **94**:69.

Evans, G. and Maxwell, W.M.C. (1987) Salamon's artificial insemination of sheep and goats. Butterworths, Sydney.

Fogarty, N.M., Maxwell, W.M.C, Eppleston, J. and Evans, G. (1999) *Proc Aust Soc Reprod Biol*. **30**:129.

Gellatly, E.S., Evans, G., Maxwell, W.M.C. and Mortimer, S.T. (1999) *Proc Aust Soc Reprod Biol*. **30**:43.

Gomez, M.C., Catt, J.W., Evans, G. and Maxwell, W.M.C. (1998) *Theriogenology*. **49**: 1143.

Hollinshead, F.K., Gillan, L., O'Brien, J.K., Evans, G. and Maxwell, W.M.C. (2003) *Reprod Fertil Dev*. **15**:351.
Johnson, L.A. and Welch, G.N. (1999) *Theriogenology*. **52**:1323.
Kelly, J.M., Kleemann, D.O., Maxwell, W.M.C, and Walker, S.K. (2008) *Reprod Fertil Dev*. **20**:570.
Kennedy, J.P., Worthington, C. and Cole, E.R. (1974) *J Reprod Fert.* **36**:275.
Kestintepe, L., Darwish, G.M., Kenimer, A.T. and Brackett, B.G. (1994) *Theriogenology*. **42**:527.
Killeen, I.D. and Caffery, G.J. (1982) *Aust Vet J*. **59**:95.
Leahy, T., Marti, J.I., Evans, G. and Maxwell, W.M.C. (2008) *Reprod Dom Anim*. **43**:148.
Leboeuf, B., Restall, B. and Salamon, S. (2000) *Anim Reprod Sci*. **62**:113.
Leboeuf, B., Delgadillo, J.A., Manfredi, E., Piacere, A., Clement, V., Martin, P., Pellicer, M., Boue, P. and de Cremoux, R. (2008) *Reprod Dom Anim*. **43**:379.
Leoni, G.G., Rosati, I., Succu, S., Bogliolo, L., Bebbere, L., Berlinguer, F., Ledda, S., Naitana, S. (2007) *Reprod Dom Anim*. **42**:299.
Liu, Y., Du, Y.T., Lin, L., Li, J., Kragh, P.M., Kuwayama, M., Bolund, L., Yang, H.M. and Vajta, G. (2008) *Cryo letters*. **29**:315.
Maxwell, W.M.C. and Butler, L.G. (1984) *Proc Aust Ass Anim Breed Genetics, Adelaide*. **4**: 192.
Maxwell, W.M.C. and Hewitt, L.J. (1986) *J agric Sci Camb*. **106**:191.
Maxwell, W.M.C. and Wilson, H.R. (1990) *Proc World Merino Conf, Pretoria*. **1**,4.3.
Maxwell, W.M.C. and Salamon, S. (1993) *Reprod Fertil Dev*. **5**:613.
Maxwell, W.M.C. and Stojanov, T. (1996) *Reprod Fertil Dev*. **8**: I013.
Maxwell, W.M.C. and Watson, P.E. (1996) *Anim Reprod Sci*. **42**:55.
Maxwell, W.M.C., Szell, A., Hunton, J.P. and Ryan, J.P. (1990). In "Reproductive Physiology of Merino Sheep", p. 217, editors C.M. Oldham, G.B. Martin and I.W. Purvis, School of Agriculture, University of Western Australia, Perth.
Maxwell, W.M.C., Evans, G., Rhodes, S.L., Hillard, M.A. and Bindon, B.M. (1993) *Reprod Fertil Dev*. **5**:57.
Maxwell, W.M.C., Evans, G., Mortimer, S.T., Gillan, L., Gellatly, E.S. and McPhie, C.A. (1999) *Reprod Fertil Dev*. **11**:123.
Maxwell, W.M.C., Evans, G., Hollinshead, F.K., Bathgate, R., de Graaf, S.P., Eriksson, B.M., Gillan, L., Morton, K.M. and O'Brien, J.K. (2004) *Anim Reprod Sci*. **82**:79.
Maxwell, W.M.C., de Graaf, S.P., El-Hajj Ghaoui, R. and Evans, G. (2007) In "Reproduction in Domestic Ruminants", volume VI, p. 13, editors JI Juengel, JF Murray and MF Smith, Nottingham University Press, Nottingham, UK.
McKelvey, W.A.C., Robinson, J.J., Aitken, R.P. and Robertson, L.S .(1986) *Theriogenology*. **25**:855.
McClintock, A.E. and Nicholas, F.W. (1991). Project No. US016, Final Report, Australian Meat and Livestock Research and Development Corporation, Sydney, Australia.
McPhie, C.A., Evans, G. and Maxwell, W.M.C. (1999) *Proc Aust Soc Reprod Biol*. **30**:44.
Meat and Livestock Australia (2004). The Breeder's Guide. A breeder's guide to LAMBPLAN, Merino Genetic Services and KIDPLAN. Meat and Livestock Australia Limited, Walker Street, North Sydney.
Morton, K.M. (2008) *Reprod Dom Anim*. **43**:137.
Morton, K.M., Catt, S., Hollinshead, F.K., Maxwell, W.M.C. and Evans, G. (2004a) *Reprod Fert Dev*. **39**:454.
Morton, K.M., Maxwell, W.M.C., Evans, G. (2004b) Proc 15th Int Congr Anim Reprod, Porto Seguro, Brazil. **1**:233.

Morton, K.M., de Graaf, S.P., Campbell, A., Tomkins, L.M., Maxwell, W.M.C. and Evans, G. (2005) *Reprod Dom Anim.* **40**:422.
Morton, K.M., Herrmann, D., Sieg, B., Struckmann, C., Maxwell, W.M.C., Rath, D., Evans, G., Lucas-Hahn, A., Niemann, H. and Wrenzycki, C. (2007) *Mol Reprod Dev.* **74**:931.
Muiño-Blanco, T., Pérez-Pé, R. and Cebrián-Pérez, J.A. (2008) *Reprod Dom Anim*. **43**:18.
Nicholas, F.W. (1996) *Anim Reprod Sci*. **42**:205.
O'Brien, J.K., Dwarte, D., Ryan, J.P., Maxwell, W.M.C. and Evans, G. (1996) *Reprod Fertil Dev*. **8**:1029.
O'Brien, J.K., Beck, N.F.G., Evans, G. and Maxwell, W.M.C. (1997a) *Reprod Fertil Dev*. **9**:625.
O'Brien, J.K., Catt, S.L., Ireland, K., Maxwell, W.M.C. and Evans, G. (1997b) *Theriogenology*. **47**:1433.
O'Brien, J.K., Hollinshead, F.K., Evans, G. and Maxwell, W.M.C. (2004) *Reprod Fertil Dev.* **16**:286.
Oropeza, A., Wrenzycki, C., Herrmann, D., Hadeler, K-G. and Niemann, H. (2004) *Biol Reprod*. **70**:1634.
Pomares, C.C., Stojanov, S. and Maxwell, W.M.C. (1995) *Proc Aust Soc Reprod Biol*. **27**:52.
Raadsma, H.W. and Tammen, I. (2005) *Aust J Exp Agric*. **45**:1021.
Rall, W.E. (1987) *Cryobiology*. **24**:387.
Rath, D. and Johnson, L.A. (2008) *Repord Dom Anim*. **43**:338.
Reyna, J., Thomson, P.C., Evans, G. and Maxwell, W.M.C. (2007) *Reprod Dom Anim*. **42**:410.
Ryan, J.P., Hunton, J.R. and Maxwell, W.M.C. (1991) *Reprod Fert Dev*. **3**:551.
Salamon, S. and Maxwell, W.M.C. (1995a) *Anim Reprod Sci*. **37**:185.
Salamon, S. and Maxwell, W.M.C. (1995b) *Anim Reprod Sci*. **38**:1.
Salamon, S. and Maxwell, W.M.C. (2000) *Anim Reprod Sci*. **62**:77.
Seidel, G.E. (2009) *Theriogenology*. **71**:107.
Smith, C. (1986) *Anim Prod*. **42**:81.
Szell, A.Z. and Windsor, D.P. (1994) *Theriogenology*. **42**:881.
Tervit, H.R. (1996) *Anim Reprod Sci*. **42**:227.
Thompson, J.G., Gardner, D.K., Pugh, P.A., McMillan, W.H. and Tervit, H.R. (1995) *Biol Reprod*. **53**:1385.
Traldi, A.S., Leboeuf, B., Baril, G., Cognie, Y., Poulin, N., Evans, G. and Mermillod, P. (1998) *Proc Assoc Europ Transfert Embryons*. P. 258.
Trounson, A.O. and Moore, N.W. (1974) *Aust J Biol Sci*. **27**:301.
Underwood, S.L., Bathgate, R., Maxwell, W.M.C. and Evans, G. (2009a) *Reprod Dom Anim*. **42**:78.
Underwood, S.L., Bathgate, R., Pereira, D.C., Castro, A., Thomson, P.C., Maxwell, W.M.C. and Evans, G. (2009b) *Theriogenology*. in press.
Vajta, G. and Kuwayama, M. (2006) *Theriogenology*. **65**:236.
Walker, S.K., Hartwich, K.M. and Robinson, J.S. (2000) *Hum Reprod Update*. **6**:564.
Windsor, D.P. and van Bueren, M. (1994) *Wool Tech Sheep Breed*. **42**:253.
Worthington, C.A. and Kennedy, J.P. (1979) *Aust J Biol Sci*. **32**:91.
Wray, N.R. and Goddard, M.E. (1994a) *Anim Prod*. **59**:71.
Wray, N.R. and Goddard, M.E. (1994b) *Anim Prod*. **59**:87.

QUALITY CONTROL FOR OVINE SNP50 BEADCHIP GENOTYPES

K. G. Dodds, B. Auvray, N. Pickering and J. C. McEwan

AgResearch, Invermay Agricultural Centre, Mosgiel, New Zealand

SUMMARY

We use a dataset of Ovine SNP50 BeadChip genotypes to investigate quality control issues for genomic data. A number of criteria are investigated: % loci scored per animal, % animals scored per locus, deviations from Hardy Weinberg, comparison with animal information (gender, parentage, breed), reproducibility for replicate samples, and unusual allelic ratios. These checks can be used to clean the dataset for its endpoint analysis.

INTRODUCTION

An ovine single nucleotide polymorphism (SNP) chip able to assay over 50,000 SNPs has recently been developed by Illumina in collaboration with the International Sheep Genomics Consortium (ISGC; www.sheephapmap.org). Quality control is an essential step in the analyses of such data. This is especially important for early studies using a genotyping platform, as often the SNPs have not been independently verified following ascertainment from sequencing data. We discuss some quality control procedures and give examples of their application to Ovine SNP50 BeadChip data.

MATERIALS AND METHODS

Methods. A number of criteria are investigated. Individual criteria or combinations of them may lead to the rejection of subsets of the data.

Genotyping success rates. Simple summary statistics of success (a genotype reported) rates by animal and locus are calculated.

Consistency between animal replicates. Replicated assays of the same animal were compared.

Consistency with recorded gender. SNPs on the X chromosome should show a pattern consistent with recorded gender. Samples from females have many heterozygous calls across the X chromosome. Samples from males should not be heterozygous for loci in the non-pseudoautosomal region (the lower portion of the X chromosome), although allowance is made for genotyping errors and mis-positioned loci.

Consistency with recorded pedigree. Animals with one or both parents also genotyped were checked to see if their SNP results were consistent with those of their putative parents. The checks account for (or discard) SNPs which are inherited in an X-linked manner.

Consistency with recorded breed. Principal components were calculated from the genomic relationship matrix which in turn was calculated using the first method of VanRaden (2008). The principal components were plotted against each other with breed denoted.

Validation of SNP position by linkage mapping. Most of the SNPs have been positioned on v1.0 of the ovine sequence assembly (www.livestockgenomics.csiro.au/sheep/oar1.0.php). These assembly positions can be checked by linkage mapping the SNPs in an appropriate resource, in this case the international mapping flock (IMF; see below). A series of mapping steps was used to allow an initial validation of the SNP positions – these methods are likely to find only gross errors in position, e.g. assigned to the wrong chromosome. The first step was an approximate (for speed) linkage analysis against loci on the Maddox *et al.* (2001) map for the assigned chromosome. This analysis used only the last generation and assumed phases that gave the strongest linkage. Loci with a lod>2.5 were assumed to be correctly assigned to chromosome. For those that remained, the

same procedure was used against the Maddox *et al.* (2001) framework map loci for the other chromosomes, with a lod>4 being used as evidence for linkage. Remaining loci were then analysed in the full IMF pedigrees with Cri-map (Lander and Green 1987). Two-point lod scores were calculated for loci with more than 10 informative meioses (the others having insufficient information for detecting linkage) against loci on the assigned chromsome. Those with a lod>3 were assumed to be correctly assigned.

Deviations from Hardy-Weinberg equilibrium. A chi-squared test statistic for Hardy-Weinberg equilibrium was calculated for each SNP within each breed. An animal was assigned to a breed if it was recorded as being more than 75% of that breed. Quantile-quantile (QQ) plots were used to aid determining which loci showed extreme values, as these plots allow one to visually account for effects of population substructure and multiple testing.

Allelic ratios and relative intensity. Illumina report normalised intensities (denoted X and Y) for the two alleles assayed for a SNP. Plots of allele frequency (Y/(X+Y)) against genome position were created for each sample genotyped. The smoothed $\log_2$ relative intensity was also plotted, where the intensity was calculated as $r = \sqrt{X^2 + Y^2}$, and then calculated as the ratio to the mean value for all animals in the analysis for that locus.

Animals. The locus mapping procedure used the International Mapping Flock (IMF) pedigrees (Maddox *et al.* 2001). Other procedures are illustrated using all or parts of a multi-breed set of animals that are part of an Ovita-funded programme investigating the relationship between locus genotypes and traits of economic importance. This resource, comprising 2785 animals, was sourced from a number of research and breeder flocks in New Zealand. They were predominantly derived from Romney, Coopworth, Perendale and Texel breeds.

Genotypes. Genotyping was undertaken by Illumina for both sets of animals using their Ovine SNP50 Beadchip. The IMF animals were included as part of the HapMap project of the ISGC. The Ovita project involved 2865 samples, with 20 animals being run in duplicate and 60 Illumina controls. Forty animals failed the genotyping, including both samples from one that was duplicated. The chip assayed 59,454 potential loci, with genotype results being reported for 53,903 (90.7%) of these. A further 338 loci had intensity (X and Y) values reported, but not genotypes.

RESULTS AND DISCUSSION

Genotyping success rates. There were 48,944 (90.8%) loci scored for all 2839 successful samples. The distributions of the intensities of 3629 loci indicated that there may have been a nearby polymorphism, creating difficulties for scoring; 1606 of these had less than 95% samples scored. There were only 2 other loci scored this poorly. The success rates for each sample and for each locus, classified as above, are shown in Figure 1. Illumina also provide a quality score for each result, and these can be used to highlight potential genotype, locus or DNA sample problems.

Consistency between animal replicates. Of the 19 duplicated animals genotyped, there were no differences in the scored genotypes, and on average only 26 loci were scored in one sample and not in its duplicate. Gross inconsistencies may have indicated mislabelling or incorrect transfers between DNA stocks, plates and chips. Minor inconsistencies would reflect the repeatability of the genotyping process. Comparing inconsistencies by SNP may indicate problematic SNPs.

Consistency with recorded gender. Four animals (2 males and 2 females) had X chromosome genotypes inconsistent with their recorded gender, later found to be due to mislabelled samples.

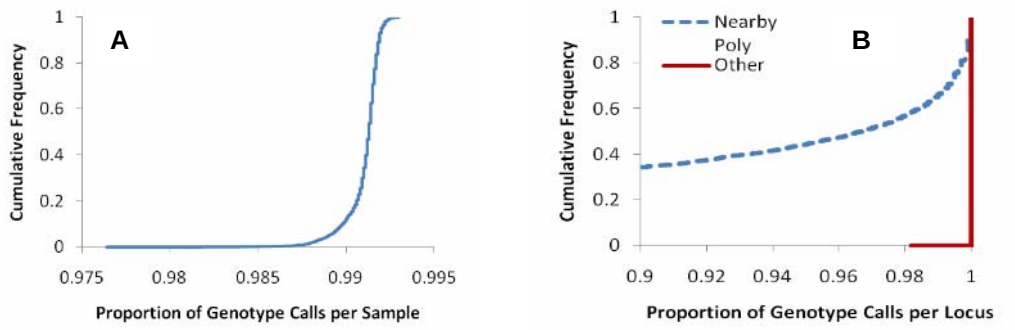


Figure 1. Genotype success rate: (A) per sample; (B) per locus.

Consistency with recorded pedigree. Of the 1302 animals with one parent also genotyped, 194 had fully concordant genotypes, 1025 had less than 30 discordant genotypes, while 83 had more than 1000 discordant genotypes. There were 8 animals with both parents genotyped; 5 of these had less than 20 discordant genotypes, while the other three had 450-500 discordant genotypes.

Consistency with recorded breed. The first two principal components are shown in Figure 2. The animals designated as being at least 90% of a particular breed tend to cluster in the same region of the figure.

Validation of SNP position by linkage mapping. Results from applying the mapping strategy to chromosome 26 are shown in Table 1. One of those that mapped elsewhere was linked to X chromosome markers. This locus was also noted to show an X-locus clustering pattern in allele intensities.

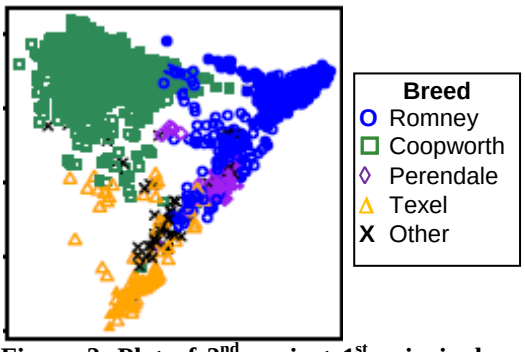


Figure 2. Plot of 2nd against 1st principal component. Closed symbols are used if the main breed component > 90%.

Table 1. Mapping Ovine SNP50 Beadchip loci from chromosome 26

	Number or %
Loci scored	917
Mapped to chromosome	90.3%
Low information	8.2%
Unmapped	1.3%
Mapped elsewhere	0.2%

Deviations from Hardy-Weinberg equilibrium. A QQ plot for the most common breed (Romney) is shown in Figure 3.SNPs with nearby polymorphism or appearing X-linked are denoted as class 2. Loci (not X-inherited) with high chi-squared values are candidates for further investigation.

Allelic ratios and relative intensity. Figure 4 shows a typical plot of allelic ratios and relative intensity for one chromosome of one animal. The frequency of one of the alleles is denoted by +, while the log relative intensity is shown by a solid line. These plots allow detection of chromosomal features of interest (Gibbs and Singleton 2006). Regions with low heterozygosity and normal intensity reflect identical by descent (inbred) regions (e.g., central region of Figure 4). Regions with low heterozygosity and low intensity reflect chromosomal deletions. Regions with high intensity and allele frequencies around 1/3 and 2/3, as well as 0 or 1, reflect chromosomal duplications. There were no obvious chromosomal abnormalities in these data.

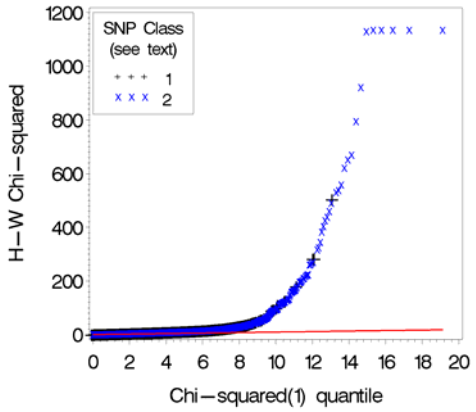


Figure 3. QQ Plot of Hardy-Weinberg chi-squared test statistics for Romney for those loci assigned to autosomes.

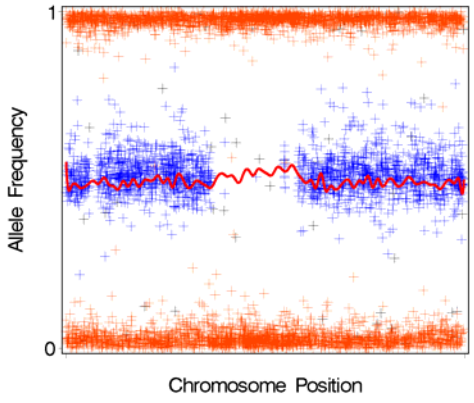


Figure 4. Plot of allelic ratios (+) and relative intensity (line) for one chromosome of one animal.

CONCLUSIONS

A variety of quality control procedures enable the screening of SNP-chip data before final analysis. These procedures often highlight data that require further checking. The process is often iterative between screening unsuitable markers and unsuitable animals.

ACKNOWLEDGMENTS

This research was supported by Ovita Ltd. We also thank Marylinn Munson and Kimberly Geizen (Illumina) for genotyping; ISGC for development of the ovine chip and for allowing us to use their genotyping results from the IMF; the many New Zealand breeders who provided DNA samples and Shannon Clarke, Dianne Hyndman and Nadia McLean for sample preparation.

REFERENCES

Gibbs, J.R. and Singleton, A. (2006) *PLoS Genetics* **2**:e150.

Lander, E. S., and Green, P. 1987. *Proc Natl Acad Sci U S A* **84:** 2363.

Maddox, J.F., Davies, K.P., Crawford, A.M., Hulme, D.J., Vaiman, D., Cribiu, E.P., Freking, B.A., Beh, K.J., Cockett, N.E., Kang, N., Riffkin, C.D., Drinkwater, R., Moore, S.S., Dodds, K.G., Lumsden, J.M., van Stijn, T.C., Phua, S.H., Adelson, D.L., Burkin, H.R., Broom, J.E., Buitkamp, J., Cambridge, L., Cushwa, W.T., Gerard, E., Galloway, S.M., Harrison, B., Hawken, R.J., Hiendleder, S., Henry, H.M., Medrano, J.F., Paterson, K.A., Schibler, L., Stone, R.T. and van Hest, B. (2001) *Genome Res.* **11**:1275.

VanRaden, P.M. (2008) *J. Dairy Sci.* **91**:4414.

BUILDING A DEER SNP CHIP

M.J. Bixley, J.F. Ward, R. Brauning, J.A. Archer and P.J. Fisher

AgResearch, Invermay Agricultural Centre, PB 50034, Mosgiel, New Zealand

SUMMARY

Advances in genomic sequencing programmes in livestock species such as cattle and sheep have enabled the building and application of large SNP (single nucleotide polymorphism) chips containing more than 50,000 markers. Applications include improved parentage and pedigree assignment as well as more accurate diversity and breed composition analysis. Further, genomewide selection (GWS) has been used to predict performance with the potential of increasing genetic gain. Other livestock species, including deer do not have a large scale genomic sequence, nor do they have other adequate supporting tools to enable trait to marker associations to be established. We have produced a reduced representational sequence of >160 million base pairs (Mbp), of which we mapped 44 Mbp to unique positions on the bovine genome. From this we selected 768 SNPs to be included in a Golden Gate (Illumina^{TM}) SNP chip. Further, we have assembled a mapping pedigree in order to quality control check these and other SNPs and to produce a genetic map. This mapping population will also be used to assess recombination rates and to reorder the deer sequence from bovine physical order to deer order. Other immediate outputs from this SNP chip will be new parentage assignment and breed composition panels. And we will investigate whether the chip will be informative for assessing within vs. across farm LD.

INTRODUCTION

This last year has seen the international dairy communities, including the dairy industry in New Zealand, adopting genome wide selection (Harris and Montgomery 2009). It is expected that this will lead to 50-70% greater genetic gain than was previously possible whilst maintaining flexibility and a multi-trait interest. One reason why this technology is valuable is because the generation interval for proving heavily used sires is dramatically reduced. Additionally there is the potential for early prediction and selection of commercial animals. Similarly, the sheep industry has developed a 60K ovine SNP chip and multiple research groups around the world have taken advantage of this new tool. The deer industry in New Zealand, with over a million animals in production, is keen to explore the options of improved genetic selection for desirable traits. We have begun phenotyping animal collections for seasonality (with conception date scanning), Johnes disease and some carcass and meat traits. To generate phenotype-marker associations, the upgrading of genomic tools (including a genetic mapping herd and larger marker sets) is also needed. The DNA from our previous (Pere-David x red) deer mapping resource (Slate et al., 2002) is nearly depleted and is not representative of the commercially important sub-species. A new 440-deer genetic mapping herd (halfsib and fullsib 3-generation families) was sourced but will not be discussed further in this paper. Here, we describe the process of sequencing for identification of SNPs and the use of bioinformatics to select markers for inclusion in a 768-SNP chip.

MATERIALS AND METHODS

Construction of Reduced Representation Libraries. The following procedure which selects a defined portion of the genome for sequencing is an adaptation of the method of Van Tassell *et al.* (2008). Breed composition using STRUCTURE^{TM} (Pritchard *et al.* 2000) was determined for Eastern Red Deer, Western Red Deer populations within New Zealand. Here, these populations will be referred to as different "breeds" although the term "sub-species" may be more appropriate.

From this analysis, pools of 25 animals, summarised in Table 1, were created by selecting those that above 70% for their respective breed. Wapiti were selected from animals that met the registration criteria of the New Zealand Elk and Wapiti Society (Asher *et al.* 2005).

Table 1. Summary of the 3 Deer breed-pools

Pool	n	Breed composition (%)	SD
Eastern	25	76.7	3.7
Western	25	85.7	3.0
Wapiti	25	N/A	

DNA was isolated from blood (Montgomery and Sise 1990) and quantified; 4µg of each sample was pooled (~100µg total). The pools were digested overnight with 2U/µg of *Rsa*I (New England Biolabs, Beverley, MA, USA) to ensure complete digestion of the genome and create blunt end fragments. Aliquots (150µl) of the digest were fractionated in 1.5% agarose gels (TBE Buffer) at 120V for 3 hours, then stained with Ethidium Bromide. Fragments in the range of 350 bp to 600 bp were excised from the gel (Figure 1) and purified using a Gel Purification kit (Qiagen Inc., Montgomery County, MD, USA).

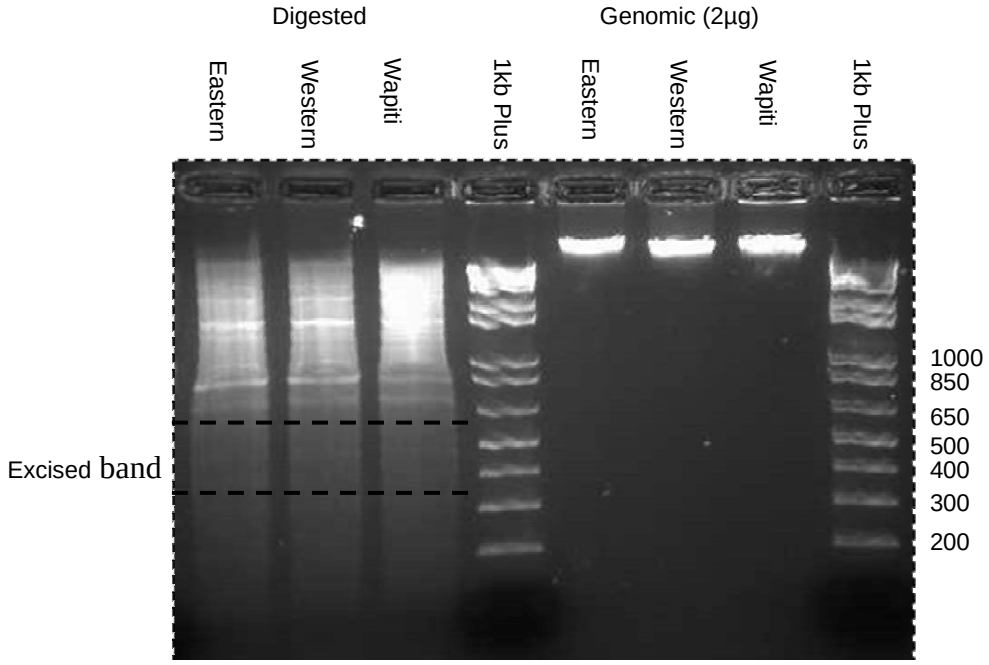


Figure 1. Digestion of pooled genomic DNA and region excised from gel

DNA Pyrosequencing. Each pool was run on a separate sequence plate section to distinguish which sequence reads belonged to which sub-species. Fragment libraries were sequenced at Otago University Sequencing facility using Roche-454 GS FLX Titanium sequence technology. This platform was used because it promised to yield long sequence reads (300-500 base pairs in length).

Genome Assembly. Pipelines originally built for assembling the ovine genomic sequence (http://www.sheephapmap.org/publications.php) were adapted for data cleaning. These procedures included masking of repetitive sequences (reads with <150 bp repeat-free DNA were removed from further analysis). The deer sequence reads were assembled, ordered and orientated using the bovine framework (Btau4 build). The assembly used the Newbler algorithm provided with the sequencing system. Alignment to the bovine genome was carried out using NCBI's BLAST tools (http://blast.ncbi.nlm.nih.gov/Blast.cgi). A GBrowse track was created of this "bovine-order deer genome", placing contigs and singleton reads on bovine chromosome and position (bp) from the chromosome start. Using this assembly, SNPs were identified and characterised based on the deer sequence reads taking into account the breed origin, sequence depth, and variability.

Deer Genome Annotation. SNP-based variants were observed and assigned each to a GBrowse track. SNP classification was based at two levels: (i) the quality of the SNP itself and (ii) the SNPs and flanking sequences were ranked technically for quality as (Golden Gate) SNP chip markers.

(i) Identification of SNP-based sequence variants

We identified variants for each sequence pool, forming six "allele x breed" categories per contig: allele 1 or 2 for each of the three (Western red, Eastern red or Wapiti) sequence runs. We selected putative SNPs from contigs where at least four reads were present but excluded SNPs where more than five reads were observed in two or more categories. We classified SNPs as "class A" if both alleles were present in each of two or more breed-pools or "class B" if we observed two alleles in one pool and if at least one other breed-pool had two or more reads of either allele.

(ii) SNP sequences and marker generation

SNP-containing sequences were sent to Illumina only if at least 50 bp of SNP-free flanking DNA was present on both sides of the SNP of interest. Further, at least 20 bp of the DNA immediately adjacent to either side of the targeted SNP had to be completely free of sequence variation.

Validation of SNPs. A panel of 70 SNPs were selected from across the genome and they have been assembled into 3 Sequenom™ mass-spectrometry (SNP) multiplexes; they will be analysed (against a 48-deer validation set) to confirm that the SNP sequence predictions are accurate.

RESULTS AND DISCUSSION

A total of 207 Mbp of raw genome sequence was produced (average read length = 267 bp, Table 2), but the distribution was bimodal (not shown). A sequencing artefact might have caused the smaller of the modes to peak at ~80 bp resulting in 41 Mbp of the reads (20%) being less than 120 bp. The largest modal peak was 340 bp or greater for all three sequence pools (Table 2); this long read length is very good for marker generation even after masking for repetitive motifs.

Table 2. Summary of the sequencing data

Breed pool	Reads (1000)	Million base pairs (Mbp)	Mean Read Length (bp)	Mode Read Length (bp)
Eastern Red	235	65.5	279	340
Western Red	356	95.8	269	340
Wapiti	179	45.5	254	380

A total of 43.9 Mbp was mapped to unique locations on Btau4 (where it was $>e^{-20}$ superior to the next best match) using BLAST tools. There were 127,266 singletons and 25,592 contigs (the term "contig" here refers to loci with more than one overlapping sequence read). The contigs'

upper size limits were seldom above 600 bp as the contigs were derived from 350-600 bp fragments that began or ended at the same enzyme cut-site. The average read depth was 1.44 (for all three breed-pools). From the exponential distribution (not shown) we calculated that >99% of the loci were expected to have fewer than six reads per allele x breed category. Loci with more than five reads in two or more categories were assumed to represent repetitive loci and were excluded. However if only one group was represented by more than five reads, we postulated that this might be due either to chance or to a run-specific sequence artefact; the contig was kept. Whilst the class A SNPs may be more likely to be genuine (i.e. not a sequence error), we also required class B SNPs because we intended to utilise breed-specific markers. We ended up with 2,170 class A or B SNPs, located on 830 contigs. The SNPs and flanking sequences were sent to Illumina for creation of a cervine 768-SNP chip. We will finish validating a 70-SNP validation set prior to final chip purchase; this testing is underway.

Our initial expectation was that we would be able to select 768 SNPs from >2,000 different contigs using a single 454 GS FLX (Titanium) sequencing run assuming a sequence yield of 350 Mbp. Further, we assumed that we could remove >97% of each genome from the gel. We were conservative in band excision, opting for a wider gel width than originally desired to ensure that the three pools contained similar loci. It was difficult to precisely determine the proportion of each genome and the sequence output was insufficient to estimate this subsequently, so it may be that > 3% was excised. Therefore it remains possible that a single Titanium sequencing run could yield >2,000 unique SNP-containing contigs (each represented by four or more reads), provided that measures are taken to increase the overall read depth from 1.44 to above 2.5.

The 768 SNPs are aligned to all 30 bovine chromosomes but their distribution is not uniform; the lowest density is on Bta24 (6 SNPs) and the highest density is on Bta28 (44 SNPs). They will be suitable for for parentage testing and breed composition panels as well as for diversity studies. We will also use the chip to genotype our new genetic mapping population for quality control of the SNPs, to determine linkage and to enable the reordering of the markers into deer-chromosomal order. And we will combine the SNPs with other markers to determine whether we can assess LD and/or haplotype blocks among samples sourced from multiple families and farms. The non-uniform SNP distribution might improve our chances of detecting LD because, for example there are 20 loci where >10 SNPs are present in less than 10 Mbp of bovine homologous sequence.

We have begun a large scale genomic sequence and SNP discovery programme to progress towards GWS in deer; whilst we expect to reveal many more than 50,000 SNPs from this project, we will optimise future work based on information and resources gathered in the described project.

ACKNOWLEDGEMENTS

This work was funded from the Foundation for Research Science and Technology, DEEResearch Ltd and Livestock Improvement Ltd.

REFERENCES

Asher, G.W., Archer, J.A., Scott, I.C., O'Neill, K.T., Ward, J.F. and Littlejohn, R.P. (2005) *Animal Reproduction Science* **90**:287.

Harris, B.L. and Montgomery, W.A. (2009) Interbull bulletin no. 39, www.interbull.org/bulletins

Montgomery G.W. & Sise J.A. (1990) *New Zealand Journal of Agricultural Research* **33**:437

Pritchard J.K., Stephens M., Donnelly P. (2000) *Genetics* **155**:945.

Slate J., Van Stijn T., Anderson R., McEwan K., Maqbool N., Mathias H., Bixley M., Stevens D., Molenaar A., Beever J., Galloway S. and Tate M. (2002) *Genetics* **160**:1587.

Van Tassell, C.P., Smith, T.P.L., Matukumalli, L.K., Taylor, J.F., Schnabel, R.D., Lawley, C.T., Haudenschild, C.D., Moore, S.S., Warren, W.C. and Sonstegard, T.S. (2008) *Nature Methods* **5:** 247.

SUMMARIZATION METHODS AND QUALITY PROBLEMS IN AFFYMETRIX MICROARRAYS

Cedric Gondro

The Institute for Genetics and Bioinformatics, University of New England, Armidale, NSW 2351

SUMMARY

The quality of RNA samples and slide hybridizations are paramount for gene expression studies. Here the focus is on hybridization quality in Affymetrix arrays and a comparison of the results of differential expression analysis for six summarization methods using slides of good and bad quality. The overlap of probes detected as differentially expressed across methods is very small even with good arrays. Slides of inferior quality also significantly change the generated list of differentially expressed genes. Extensive qualitative and quantitative quality control measures should be used prior to downstream analyses of data to ensure early detection and removal of problematic slides. Given the high variability of results across summarization methods it is recommended that different methods are used to conduct analyses and the intersecting results of these are used for further downstream analyses.

INTRODUCTION

Microarrays provide a simultaneous measurement of the expression level of thousands of genes from a biological sample. Their most common use is in comparative studies in which one condition is compared to another to identify relative changes between expression levels, i.e. the differentially expressed (DE) genes.

Microarrays are prone to exhibit high levels of experimental and systematic variability that are not related to the experimental contrasts. To ensure the best possible outcome it is critical that these effects are identified and adequately handled. Thus, the bulk of microarray analysis work lies in extensive pre-processing steps to determine the quality of the slides and calibration methods to remove spurious variation. Bad quality slides have unreliable intensity measures and can have a very large effect on final results. These slides should be identified and removed from the analysis. Slides deemed of adequate quality will then undergo calibration steps which generally consist of: (1) background correction to remove intensity measures that are not due to the target; (2) normalization of the probe intensities, which is achieved by adjusting the overall distribution intensities making them similar across slides (note that this step usually makes a dataset testable only within itself, if new slides are added to the experiment the entire set has to be renormalized); and (3) a summarization step which is specific to Affymetrix GeneChips, since these are unique in the use of a set of short oligos to target a transcript, usually 11 different pairs of 25mer oligos, with each pair consisting of a perfect match (PM) to the standard reference sequence and a mismatch (MM) with exactly the same sequence except for a mismatch at position 13 (in principle the MM should pick up cross hybridization noise). This *probe set* is summarized into a single intensity value for each target on each array.

Different methods have been developed for each of the above mentioned steps (Irizzary *et al.* 2006). However it is still unclear which approach is best, and it has been shown that the main source of variation between results is due to the choice of summarization method (Harrison *et al.* 2007). In this work we quantify the differences in probes detected as differentially expressed across six summarization methods using the Affymetrix GeneChip bovine genome array, and test the robustness of each method to technical hybridization problems.

MATERIALS AND METHODS

Data. Twenty Affymetrix GeneChip bovine genome arrays hybridized to RNA extracted from ovine blood samples were used. The data consists of 2 subsets of slides (taken from a larger experiment), one with 10 good quality arrays and the other with 10 bad quality arrays, each hybridized to the same RNA samples. The bad slides were due to technical hybridization problems that occurred in the original experiment while the good slides are simply a repeat with a new batch of slides using the same RNA samples. Each set is a simple control x treatment contrast with 5 slides per group.

Pre-processing. Qualitative and quantitative quality control (QC) measures were used prior to downstream analyses to detect problematic slides. These include image plots for detection of spatial effects, normalized unscaled standard errors (NUSE), relative log expression, MA plots, RNA degradation, call flags, match-mismatch intensities, slide correlations and principal component analysis (an overview of pre-processing is given in Gentleman *et al.* 2005).

Summarization methods. Six summarization methods were used to generate expression measures: MAS 5.0, RMA, GCRMA, PLIER, VSN and MBEI (methods are detailed in Gentleman *et al.* 2005).

Design. For each summarization method, differential expression of genes was tested using different combinations of ten slides. Initially analysis was conducted using only the good slides, and then successively repeated by replacing 2 good slides with their respective (same RNA source) bad slides until the analysis was run with only the bad slides (thus the number of good slides in each replicate was 10, 8, 6, 4, 2, 0). The procedure was repeated as a balanced replacement (bad slides shared equally between contrasts) and unbalanced (bad slides allocated in first instance to the control).

Analysis. Differential expression was tested on log2 expression intensities for the 6 summarization methods for each of the 10 datasets using a moderated t-statistic (Smyth 2004) after removing control probes and probes detected as marginal or absent across all arrays (probes with low intensities). Probes were selected as differentially expressed for a p-value of 0.01.

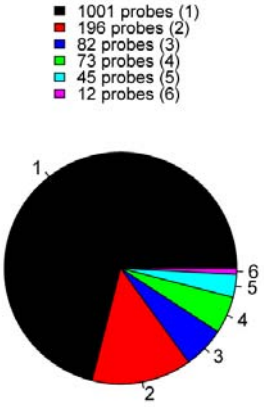


Figure 1. Intersect of DE probes across summarization methods for good slides. Parentheses indicate number of methods in which the probes were detected.

RESULTS AND DISCUSSION

The number of DE probes detected is summarized in table 1. Of immediate notice, and concern, is that the intersect of probes across all methods is extremely low irrespective of the quality of the slides. The intersect is moderately better for the good quality slides but still only 3.2%, that is only 12 probes out of the average 371DE probes per method (the total number of probes across methods was 2228 with 1409 unique). Most probes are detected in only one method (71%, see Figure 1). Even though some methods are methodologically close (e.g. same normalization is used), results tend to bear limited replicable correlation with different datasets. Hence if for a given analysis RMA and GCRMA yield similar results while MAS and PLIER are similar between them but further apart from the first two, this cannot be used to make decisions on a choice of summarization

method since in another experiment this order may be completely reversed. Figure 2 illustrates this issue with the good and bad arrays. In terms of numbers of DE probes there is a considerable spread across methods with GCRMA consistently showing the lowest numbers of DE probes and MBEI the highest, up to 10-fold differences can be seen in Table 1.

Table 1. Differentially expressed probes per summarization method. Numbers in parenthesis refer to the DE probes detected in the unbalanced designs. The last column shows the number of probes in common across all methods

arrays	mas	rma	gcrma	plier	vsn	mbei	intersect
good slides	292	347	126	386	306	771	12
8 good	158 (124)	127 (74)	52 (37)	58 (101)	55 (60)	300 (198)	2 (1)
6 good	65 (231)	20 (170)	19 (43)	22 (321)	18 (270)	38 (411)	1 (0)
4 good	40 (127)	25 (103)	22 (50)	9 (94)	17 (70)	41 (356)	0 (1)
2 good	91 (65)	143 (64)	31 (35)	32 (20)	43 (31)	75 (99)	0 (0)
bad slides	213	430	164	439	766	315	1

Taking the DE probes using the good slides as a *gold standard*, Figure 3 shows how even small numbers of bad slides can result in very different lists of DE probes. With two bad slides (one in each treatment) on average around 28% of the original DE probes are still detected. MAS is somewhat more robust at 43% overlap, whilst PLIER at the lower end shows only 16%. With four or more bad slides the numbers fall under the 20% line for all methods. Even with the lower numbers of DE probes in good/bad combinations there are still many new probes being detected which are just noise due to the poor slide quality. As would be expected the unbalanced designs show even greater disparities of results (Figure 3, right pane). The somewhat greater robustness of MAS is due to the method not normalizing between arrays but on a targeted predefined mean value. The downside of MAS is that it tends to overestimate the effects at low intensities.

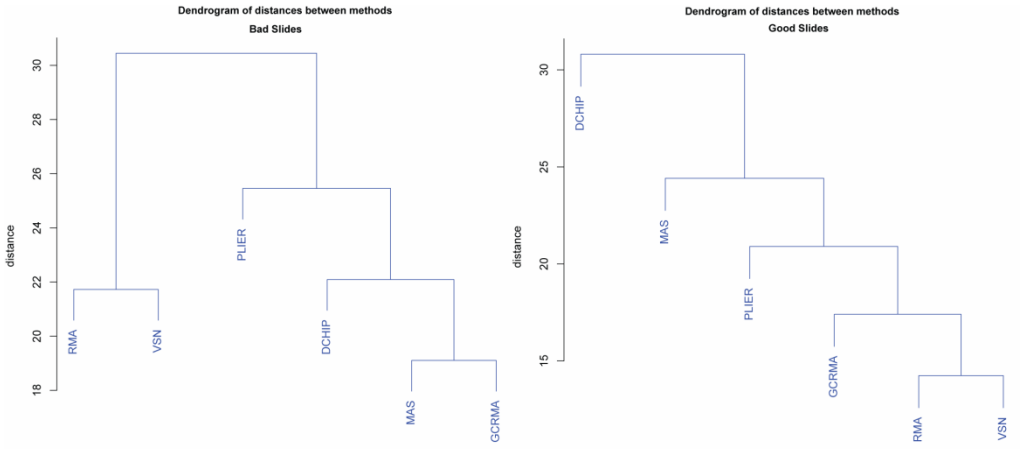


Figure 2. Distance plots of summarization methods for the bad and good slides.

The 12 intersect probes detected across all summarizations using the good slides are more robust to slide quality. With two bad slides the average across methods is close to 55% (between

25% for PLIER and 67% for MAS, RMA and MBEI) and 39% for unbalanced designs. Even using only bad slides the average overlap is still over 26%.

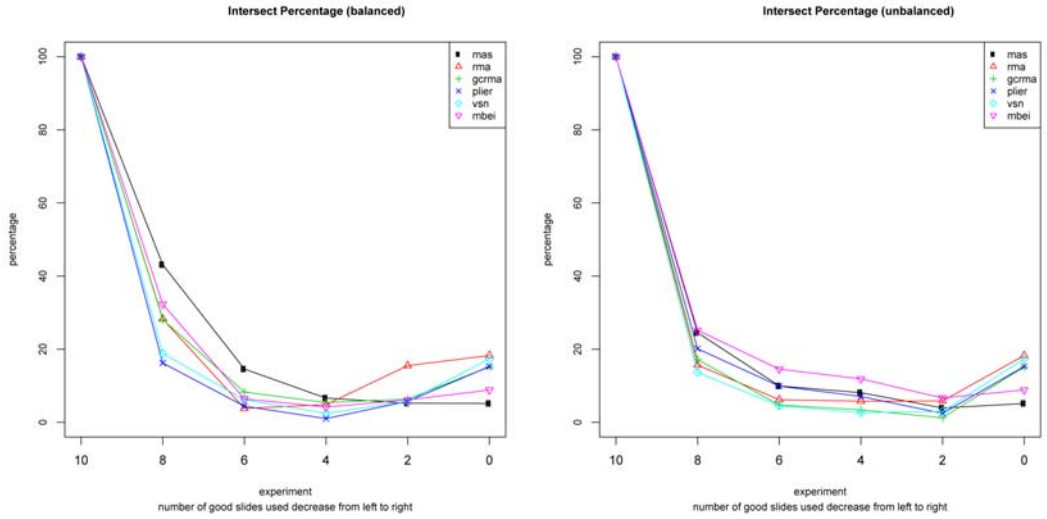


Figure 3. Similarity of DE probes for each summarization in relation to 10 good slides caused by replacing good slides with bad slides. Left pane balanced, right pane unbalanced.

CONCLUSIONS

The dataset used in this work was only available because of extensive quality control test which helped flag problems with the original hybridizations in the larger experiment. If these checks had not been performed the results would have been very different for whichever analysis methodology had been used. The choice of summarization method also has a large impact on final results and there is very little overlap between them. No individual method is highly robust to experimental noise but different summarizations combined can be more robust.

For positive outcomes from array experiments it is recommended that extensive quality control checks, such as those previously mentioned, are performed. If in doubt an array should be discarded or at least the analysis should be run with and without a questionable slide to quantify the effect it is having on results.

More than one summarization method should be used for the analyses. The intersect of results can help identify a more stable subset of results and in effect also help to correct for multiple testing problems.

ACKNOWLEDGEMENTS

This research was financially supported through SheepGenomics by Meat and Livestock Australia and Australian Wool Innovation Limited.

REFERENCES

Gentleman, R., Carey, V.J., Huber, W., Irizarry, R.A. and Dudoit, S (eds.) (2005). "Bioinformatics and Computational Biology Solutions Using R and Bioconductor" Springer, The Netherlands.

Harrison, A., Johnston, C. and Orengo, C. (2007). *BMC Bioinformatics* **8**:195.

Irizarry, R., Wu, Z. and Jaffee, H. (2006). *Bioinformatics* **22:**789.

Smyth, G.K. (2004). *Stat Appl Genet Mol Biol*. **3**:Article3.

A NEW STRATEGY TO IDENTIFY THE DISEASE CAUSING MUTATION FOR NEURONAL CEROID LIPOFUSCINOSIS IN SOUTH HAMSPHIRE SHEEP

I.F. Mohd Ismail[1], J.A.L Cavanagh[1], N.L. Mitchell[2], P.J. Houweling[1], D.N. Palmer[2] and I. Tammen[1]

[1]Reprogen, The University of Sydney, Camden, NSW, Australia
[2]Agriculture and Life Sciences Division, Lincoln University, New Zealand

SUMMARY

The New Zealand South Hampshire sheep have been well characterised as an animal model for variant late-infantile neuronal ceroid lipofuscinosis (vLINCL) in humans. The disease causing gene has been identified as CLN6, but so far no mutation has been identified. A sheep BAC containing CLN6 and flanking region (~120kb) was sequenced at 13.49-fold sequence coverage using the 454 sequencing™ method from Roche® to complement the existing but incomplete public domain ovine sequence for the region of interest. The 454 sequence was assembled to bovine genomic sequence on chromosome 10 (BTA10). For mutation screening 15 long-distance PCR products from affected and normal South Hampshire sheep covering CLN6 as well as substantial 3' and 5' flanking sequence are currently optimized to be sequenced. Regulatory elements and/or mutations identified in CLN6 non-coding regions are likely to indicate positions for disease causing mutations not only in the South Hampshire sheep but also in human uncharacterised vLINCL patients.

INTRODUCTION

The neuronal ceroid lipofuscinoses (NCL) are a group of inherited neurodegenerative diseases that occur in humans and several animal species. The NCLs are characterised by accumulation of fluorescent storage bodies in neurons and other cells with at least 8 different causatives genes identified so far. In South Hamsphire sheep the mode of inheritance is autosomal recessive and onset of clinical signs is about 8 months of age. Linkage mapping established South Hampshire (SH) sheep as a valuable animal model for human variant late-infantile NCL (vLINCL) (Broom *et al.* 1998), which is caused by mutations in the CLN6 gene (Wheeler *et al.* 2002; Sharp *et al.* 2003). Reduced expression of CLN6 in affected SH sheep tissues and absence of a disease causing mutation in the coding sequence suggests a mutation in a regulatory element (Tammen *et al.* 2006). These elements are found mostly in non-coding regions within the gene or regions upstream and downstream.

When compared to mouse models, large animal models have been found to be particularly valuable for these neurodegenerative diseases, as they resemble humans more closely in relation to brain size and structure, clinical signs and progression, and life expectancy.

MATERIALS AND METHODS

South Hampshire sheep. South Hampshire (SH) sheep DNA samples were provided from the research flock at Lincoln University, Christchurch, New Zealand (Jolly *et al.* 1980, Tammen *et al.* 2006). The sheep were diagnosed as either normal, carrier or affected with NCL using clinical signs (blindness, seizures and behavioural changes), histology (fluorescent storage material in neurons and other cells) and pedigree information, as well as an indirect DNA test based on an A/G polymorphic allele in exon 7 of CLN6 (Tammen *et al.* 2006).

Identification of conserved non-coding sequences. The programs Vista® (Frazer *et al.* 2004) and GeneDoc® (Nicholas *et al.*1997) were used to align genomic and coding sequence of CLN6 in 9 species (human, cattle, dog, macaque, opossum, mouse, rat, chicken and Fugu fish) with the aim to identify conserved non-coding sequences (CNCS) with potential regulatory functions.

Bacterial Artificial Chromosome (BAC) sequencing. The sheep BAC clone 270H8 was provided by Dr. Daniel Vaiman (INRA, Jouy-en-Josas, France) after PCR screening of the sheep BAC library (Vaiman *et al.* 1999) with CLN6 specific primers. Further characterisation of the BAC revealed presence of the CALML 4 gene downstream of the CLN6 gene as well as a CNCS approximately 7kb upstream of CLN6 using primers designed from publically available sheep sequences from the International Sheep Genomics Consortium (ISGC) or CNCS sequences identified as part of this project.

The BAC DNA was purified using the QIAGEN Large-Construct® kit according to manufacturer's protocol. This method involves an ATP-dependent exonuclease digestion step for the selective removal of bacterial genomic DNA. A total of 3μg of purified BAC DNA at a concentration of 80ng/μl and an OD_{260}/OD_{280} of 1.65 was submitted for 454 sequencing™ (Goldberg *et al.* 2006) to University of Otago, High Throughput DNA Sequencing Unit, New Zealand. **A** sequencing library was constructed and sequenced in a $1/16^{th}$ standard FLX plate equivalent.

Sequencing raw data was returned and contig assembly and individual reads were obtained in FASTA format using the whole genome shotgun and EST sequence assembler program MIRA (Chevreux *et al.* 2004). BLAST (Altschul *et al.* 1990) batch analysis was performed at the National Center for Biotechnology Information (NCBI) to identify *E.coli* and pBeloBACII contamination. The remaining contigs were aligned to ISGC sheep sequence version 1.5 using GeneDoc® (Nicholas *et al.*1997) as well as to bovine chromosome (BTA) 10 using the Ensembl Genome Browser (http://www.ensembl.org/).

Targeted sequencing for mutation screening. A total of 15 PCRs covering approximately 30kb of ovine genomic DNA are currently optimised for targeted sequencing and mutation screening in affected and normal South Hampshire sheep. This region covers genomic sequence including and surrounding CLN6 - starting from the CNCS 7kb 5' of CLN6 and ending in the 3' region of CALML4. Primers were designed from published CLN6 cDNA (Tammen *et al.* 2006) and ISGC sequence. PCR products from affected, carrier and normal South Hampshire sheep will be submitted for 454 sequencing™ (Goldberg *et al.* 2006).

RESULTS AND DISCUSSION

Initial bioinformatic analysis. Bioinformatic analysis of VISTA® across 9 species revealed that parts of the 5'UTR, 3'UTR and intron 1, 2 and 3 of the CLN6 contain highly conserved regions (CNCS) with potential regulatory function (Mohd Ismail *et al.* 2006). Sequence information from regions conserved in cattle and human were used to design PCR primers for amplification of sheep specific sequence predicted to be located 7-10kb upstream of *CLN6* Exon 1. Furthermore, comparative information was used to predict the identity and distance of genes flanking ovine CLN6: FEM1-b was predicted to be located 40kb upstream and CALML 4 about 7kb downstream of CLN6.

Sequencing of the sheep BAC. 454 sequencing™ of the sheep BAC clone 270H8 resulted in approximately 2 million base pair (bp) sequence reads. 1248 asterix symbols were found throughout the 454 sequences which represent homopolymers. These homopolymers are

considered sequencing errors and they are randomly located with only 20% occurring before or after stretches of A/T.

Contig assembly from the raw data provided by the University of Otago sequencing unit was performed using the assembler program MIRA. Sequences were assembled into 114 contigs ranging from 42 to 18863bp each with an average read of 80 reads per contig and a 13.49 fold coverage as calculated by MIRA. BLAST analysis identified that approximately 30% of the total 114 sequence contigs contained BAC vector or bacterial genomic DNA sequences. However, these were smaller in length with an average size of 200 to 400 bp. BAC vector and *E.coli*- free contigs were aligned to the cattle genome (Btau4.0) at position 14,821,789-14,942,374 bp using the Ensembl Genome Browser to produce a visual representation as illustrated in Figure 1.

Additional editing based on ISGS sequence information as well as previously derived CLN6 sequences appears to allow bridging of most gaps but is not finalised. The sequence includes CLN6, CALML4, FEM1-b as well as parts of PIAS1. Two long terminal repeat (LTR) sequences ranging from 500bp to 1kb were matched to 10 to 15 regions of the cattle chromosome 10 which made assembly difficult.

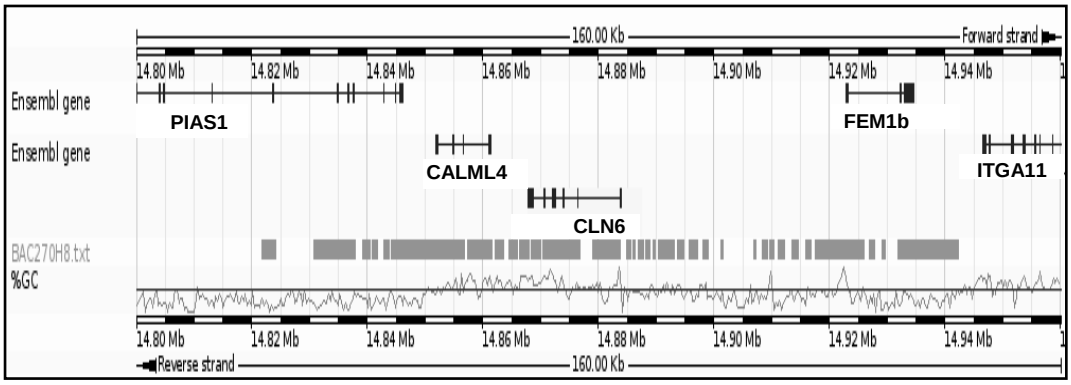


Figure 2. Alignment of 454 sheep BAC 270H8 contigs against cattle chromosome 10: 14,800,000-14,960,000 (assembly Btau 4.0) using the Ensembl browser.

Targeted sequencing for mutation screening. Ten out of 15 long distance PCR products have been optimised for amplification on sheep genomic DNA. The newly generated BAC sequence will allow for design of new primers for the remaining regions, which failed so far to amplify single band PCR products of predicted sizes. The resulting products will be individually multiplex identifier (MID) tagged and 454 sequenced.

CONCLUSION

Whole genome sequencing has not been completed for the sheep and the publically available sequence information for our region of interest had large gaps. Traditional Sanger sequencing of the genomic CLN6 sequence, particularly in the very GC rich 5' region of the gene, has been very cumbersome (Tammen *et al.* 2006) and hindered the identification of a disease causing mutation for NCL in South Hampshire sheep. Identification of such a mutation is crucial, as these sheep are the most extensively characterised model for NCLs in general, and for vLINCL in particular. Identification of the disease causing mutation will not only assist greatly in the management of the research flock, but provide additional support for our claim that NCL in these sheep is caused by CLN6 down regulation, and thus be of great importance in relation to the evaluation of therapeutic approaches such as gene therapy.

The 454 sequencing of the sheep BAC containing CLN6 was found to be cost-effective and efficient for the generation of good coverage sequence for our region of interest. Initial contig assembly left some gaps, but these can be mostly bridged using existing sequence information or were caused due to differences between sheep and cattle sequences. This sequence provides the required backbone for the sequencing of long-distance PCR products that will be used for mutation screening. Depending on the types, number and locations of mutations identified by the proposed mutation screening approach, further experiments will be designed to assess the impact of mutations on gene expression. A number of human cases of vLINCL linked to the CLN6 gene appear to also have mutations in non-coding regions of CLN6. The identification of regulatory elements for CLN6 in sheep will hopefully indicate positions for further analysis in these patients and thus improve DNA diagnostics in patients and their relatives.

ACKNOWLEDGMENTS

This work has been funded by the BDSRA and NIH Grant (R01 NSO53559-01A1). We would like to thank Dr Kyall Zenger and Dr. Matthew Hobbs for assistance with bioinformatic analysis, Dr. John McEwan and the ISGC for early release of partial OAR7 sequence v1.5, and Dr. Daniel Vaiman (INRA, France) for providing the sheep BAC clone. I.F. Mohd Ismail is a sponsor student of Malaysia's Ministry of Higher Education and Universiti Putra Malaysia.

REFERENCES

Altschul, S.F., Gish, W., Miller, W., Myers, E.W. and Lipman, D.J. (1990) *J. Mol. Biol.* **215**:403.

Broom, M. F., Zhou, C., Broom, J. E., Barwell, K. J., Jolly, R. D., and Hill, D. F. (1998) *Neuropathol. Appl. Neurobiol.* **35:**717.

Chevreux, B., Pfisterer, T., Drescher, B., Driesel, A. J., Mueller, W. E., Wetter, T. and Suhai, S. (2004) *Genome Res.* **14**:1147

Frazer, K.A., Pachter, L., Poliakov, A., Rubin, E.M., and Dubchak, I. (2004) *Nucleic Acids Res.* **32**:W273.

Goldberg, S.M**.,** Johnson, J**.,** Busam, D**.,** Feldblyum, T**.,** Ferriera,, S**.,** Friedman, R**.,** Halpern, A**.,** Khouri, H**.,** Kravitz, S.A**.,** Lauro, F.M**.,** Li. K**.,** Rogers, Y.H**.,** Strausberg, R**.,** Sutton, G**.,** Tallon, L**.,** Thomas, T**.,** Venter, E**.,** Frazier, M**.,** and Venter, J.C**.** (2006) *Proc. Natl. Acad. Sci. USA* **24:**103.

Jolly, R.D., Janmaat, A., West, D.M., and Morrison, I. (1980) *Neuropathol. Appl. Neurobiol.* **6:**195.

Mohd Ismail, I.F., Palmer, D.N., and Tammen, I. (2006) ICE-EM Summer Symposium in Bioinformatics, Canberra, 5-8.12.2006.

Nicholas, K.B, Nicholas, H.B. Jr., and Deerfield, D.W. (2007) *EMBNEW.NEWS* **4**:14.

Sharp, J.D., Wheeler, R.B., Parker, K.A., Gardiner, R.M., Williams, R.E., and Mole, S.E. (2003) *Hum. Mutat.* **22**:35.

Tammen, I., Houweling, P.J., Frugier, T., Mitchell, N.L., Kay, G.W., Cavanagh, J.A., Cook, R.W., Raadsma, H.W. and Palmer, D.N. (2006) *Biochim. Biophys. Acta* **1762**:898.

Vaiman, D., Billault, A., Tabet-Aoul, K., Schibler, L., Vilette, D., Oustry-Vaiman, A., Soravito, C., and Cribiu, E. P. (1999) *Mamm. Genome* **10:**585.

Wheeler, R. B., Sharp, J. D., Schultz, R. A., Joslin, J. M., Williams, R. E., and Mole, S. E. (2002) *Am. J. Hum. Genet.* **70**:537.

AN INTEGRATED GENOMICS APPROACH TO IMPROVING WOOL PRODUCTIVITY AND QUALITY

P. I. Hynd[1], C.S. Bawden[2], N.W. Rufaut[3], B.J. Norris[4], M. McDowal1[l], A.J. Nixon[3], Z. Yu[3], A.J. Pearson[3], G.S. Nattrass[2], C. Gordon-Thompson[5], G.P. Moore[5], S.M. Dunn[2], N.M. Edwards[1], D. Smith[2] and C.J. McLaughlan[2]

[1] School of Animal and Veterinary Science The University of Adelaide, Roseworthy Campus Roseworthy SA 5371; [2]South Australian Research and Development Institute Roseworthy Campus Roseworthy SA 5371; [3]Growth and Development Section, AgResearch Ruakura, Private Bag 3123, Hamilton 3214, NZ, [4]CSIRO Livestock Industries, Queensland Bioscience Precinct, St. Lucia 4067 Qld, [5]School of Science, University of Western Sydney, Kingswood, NSW, Australia

SUMMARY

This paper summarises the approach taken in a national research program designed to "identify and utilise genes of importance in the sheep industries". The sheepgenomics program as it was known, comprised meat, wool and parasite subprograms with an underpinning core technology subprogram. The wool subprogram used a combination of gene association and functional biology studies to identify genes and gene networks amenable to manipulation or selection to improve wool production and quality. Significant progress was made in identifying genes involved in wool follicle initiation, hair cycle regulation, recessive black pigmentation and fleece rot. Manipulation of key windows of foetal development resulted in lifetime positive changes in wool production, an important proof of concept in functional, developmental genomics.

INTRODUCTION

The Australian wool industry operates in a highly-competitive, global, textile fibre market in which it currently captures a small and diminishing share of the consumer's expenditure on apparel clothing. To remain competitive in this market, the industry must address a number of pressing issues which are limiting productivity, profitability and consumer acceptance of the products. These include mulesing, dark fibre contamination, fleece rot, flystrike, anthelminthic resistance, relatively coarse fibres, weak fibres, prickle in garments, poor easy-care attributes and high price relative to competitors. At the time of inception of the sheepGENOMICS program (2004), molecular genetics applied to animal breeding was in its infancy, with great expectations attached to the discovery of quantitative trait loci (QTL) for difficult-to-measure traits. Since then, high throughput single nucleotide polymorphism (SNP) genotyping has paved the way for whole genome selection and more targeted SNP marker identification. Developments in bioinformatics, and in particular, networked pathway analyses, now allow more functionally-relevant interpretation of gene expression studies. The wool subprogram of the sheepGENOMICS initiative developed an integrated, functional genomics approach to dissect the molecular and cellular events involved in the critical periods of development of the follicle population in the skin of developing sheep foetuses. A suite of techniques for gene detection, gene expression, gene localisation, gene transfection, *in vitro* cell functional assays, gene network analysis and biochemical manipulations were targeted at key windows of skin development and at the long-standing problems of fleece rot and recessive black fibre pigmentation. These techniques and a summary of progress to date are the subject of this paper.

THE EXPERIMENTAL APPROACHES TAKEN IN sheepGENOMICS.

Every aspect of lifetime wool production (clean fleece weight, wool growth efficiency), quality (mean diameter, variance in diameter, length growth rate of fibre, crimp characteristics, staple

strength, style, colour, fleece rot incidence), and even aspects of ectoparasite attraction (suint and wax contents), are established during the critical windows of development of skin and its appendages. The wool subprogram of SheepGENOMICS therefore concentrated on cellular and molecular events occurring during skin development, and particularly during critical windows of follicle initiation. The foetal development of wool follicles and their associated glands (sweat and sebaceous) is called neogenesis. The morphological changes that take place throughout the post-natal hair cycle are thought to be a recapitulation of these neogenic events (indicated in Fig 1 on the right side of the 'figure-of-six' diagram of Stenn and Paus, 2001).

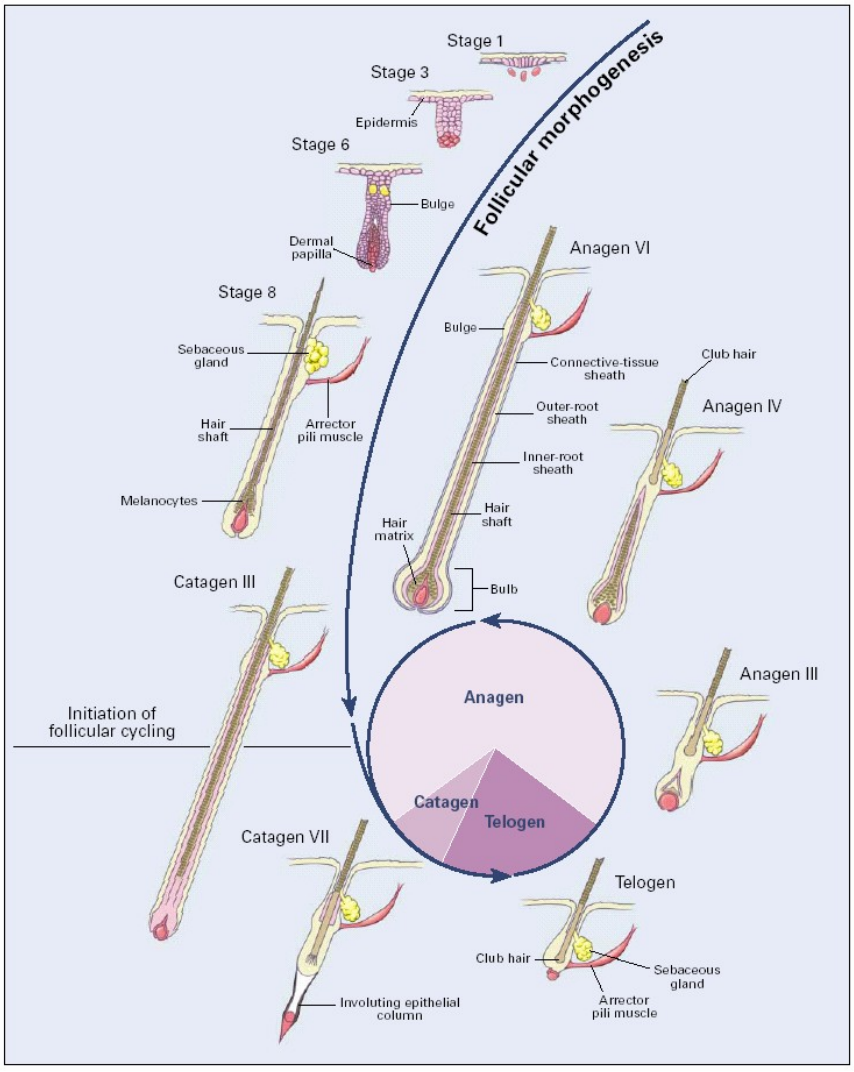


Figure 1. Targets of investigation of gene expression in the Wool subprogram. The left side of this figure-of-six' diagram represents the neogenesis of hair follicles and associated structures. The right side represents the on-going morphogenic events during the hair cycle in established follicles (after Stenn and Paus 2001).

In other words, the signalling events involved in the re-establishment of the follicle after the resting phase of the cycle (telogen) were hypothesised to be the same as those operating in neogenesis.

The strategy taken in the Wool subprogram of the SheepGENOMICS program was deliberately multi-pronged (from gene detection, gene sequencing, gene function and pathway modulations), because at that time, there were few examples of successful application of molecular genetics to applied livestock production. We took several approaches to identifying the key, and possibly unique, gene networks operating in follicle neogenesis and post-natal cycling: (1) expression studies based on candidate genes known to be involved in neogenesis in other epithelial tissues; (2) identification of genes and gene networks operating at critical phases of neogenesis and morphogenesis, using expression micorarrays; (3) targeting of pathways known to be involved in tissue neogenesis by direct biochemical/endocrinological intervention during critical periods of skin and appendage formation; and (4) identification of gene function by transfection of keratinocytes and assessment of their behaviour using *in vitro* assays. The candidate gene approach was also taken to identify the hitherto intractable problem of the agouti locus (see below), while the microarray global screen approach was taken to identify potential markers for fleece rot incidence. Our philosophy was that this multi-pronged candidate gene, targeted manipulation, and gene discovery approach would be a powerful model for dissecting functional genomics in its broadest sense. The results obtained using this approach are outlined below.

IDENTIFICATION OF TARGET GENES AND PATHWAYS IN FOLLICLE INITIATION

The major resource developed in SGP for the study of gene expression during foetal and early postpartum skin development was a tissue library of skin samples derived from lambs from artificially-inseminated, synchronised ewes. Samples were collected from cohorts of 3 foetalfoetal lambs sacrificed at 2-4-day intervals from days 35 to 143 post conception and from 8 lambs at 4-7 day intervals from days 2 to 100 postpartum. This resource, created by researchers from the South Australian Research and Development Institute and The University of Adelaide (CS Bawden, G. Nattrass, C.J. McLaughlan, H. McGrice, S.M. Dunn), allowed detailed identification of the histological stage of development of the skin and its appendages, localisation of mRNA expression (by *in situ* hybridisation) and quantification of mRNA expression by quantitative PCR.

A candidate gene approach was taken to characterise the expression profiles of approximately 100 genes, identified from literature related to the ontogeny of other organs and tissues which rely on similar ectodermal/mesenchymal interactions as skin. The genes included those involved in early epidermal/dermal signalling (eg members of the Wnt/βcatenin, FGF, EDA, BMP, Notch/Delta and Shh signalling pathways); mid- to late-signals in organ development (eg VEGF, VE-cadherin, Ephrin, NGF); structural proteins in the follicle, fibre, skin (eg trichohyalin, keratin and keratin-associated proteins); cell signalling molecules (eg β-Activin, Follistatin, SCF, HGF); transcription factors (eg Lef-1, Gli1, Prox1, HIF1); proteins involved in tissue remodelling (eg E-cadherin, matrix metalloproteinases); proteins involved in organ branching events (eg FGF10, BMP4, Sprouty, Netrin); and those involved in the formation and function of accessory glands (eg matrix metalloproteinases, cytokeratin 7 (CK7) and 1,5-alpha reductase (SRD5A1), Androgen receptor (AR) and Carcinoembryonic antigen cell adhesion molecule 1 / Biliary glycoprotein (CEACAM-1 / BGP)."

For each of these genes, spatiotemporal expression profiles were painstakingly quantified by qPCR and *in situ* hybridisation throughout the developmental time series. Expression profiles such as that depicted in Figure 2 were thereby obtained, allowing coincidence of changes in gene expression to developmental event to be determined. In this example the expression of the developmental protein, sonic hedgehog, begins expression at day 57 the time of primary follicle

initiation. Gene expression then peaks at day 83 when the large secondary population is initiated, with a secondary peak at day 101 when the secondary derived population is initiated.

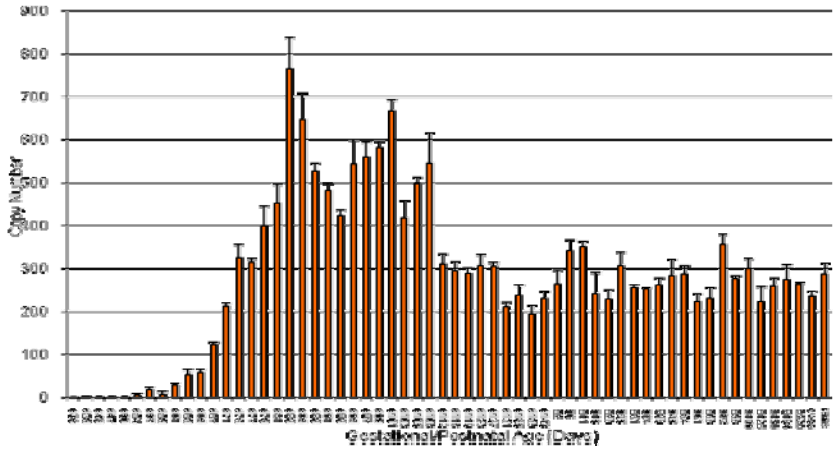


Figure 2 : qPCR assay data showing prevalence of transcripts encoding sonic hedgehog transcripts in midside skin during ovine development. Mean transcript prevalence for each developmental time point is plotted +/- the standard deviation (McLaughlan, Bawden, Nattrass, Dunn and McGrice unpubl.)

The challenge now facing the researchers is to coordinate this vast array of spatial and temporal information into an integrated pathway analysis which will focus targeted manipulations or identification of genes within which we might search for SNP mutations.. Ingenuity Pathway Analysis® is one means of integrating the data into a consolidated output, and is currently being employed to identify the major pathways related to developmental events.

One such pathway that appears to play a major role in tissue development in general and skin and appendage development in particular is the Delta/Notch signalling pathway. Philip Moore, Peter Wynn, Claire Gordon-Thompson and Stephanie Xavier found members of the Notch signalling pathway are indeed present in the cells that form the dermal condensate (the precursor to the dermal papilla of the follicle). Differences in the level of expression of Notch-1 and Delta-1 proteins were apparent between days 56 and 70 post-conception, and the ratio of Delta to Notch in dermal condensate cells was greater in Merino foetuses than Tukidale foetuses. A targeted biochemical approach was taken in a series of experiments initiated to determine the viability of manipulating the Delta/Notch pathway using an inhibitor of γ-secretase, DAPT ((N-[N-(3,5-difluorophenacetyl)-L-alanyl]-(S)-phenylglycine t-butyl ester). The results of an *in vitro* trial of DAPT on cell aggregation in dermal papilla cells indicated that the inhibition of the Delta/Notch pathway resulted in complete inhibition of cellular aggregation (Figure 3).

It is well-established that properties of the dermal tissue exert a strong effect on the final structure of developing follicles *in vivo*. Thus the size of the dermal condensate and papilla seem to determine the size of the mature follicle and the diameter of the fibre it produces. The aggregation of papilla cells *in vitro* is a model for this morphogenetic process. The work of Nick Rufaut (AgResearch NZ), described below, encompasses the development of an *in vitro* bioassay which uses papilla cell aggregation to evaluating the impact of potential therapeutics on the specification of follicle size. Philip Moore's work suggests that the Delta/Notch signalling

pathway regulates the formation of dermal cell condensates. It is yet to be demonstrated that targeted manipulations of this pathway can result in increased follicle density and decreased fibre diameter *in vivo*, the desired outcome. Nevertheless, manipulation of this pathway has potential for *in utero* therapeutic manipulation of follicle formation and lifetime fibre growth. Moreover identification of SNPs in the Delta/Notch pathways may provide novel markers of wool growth.

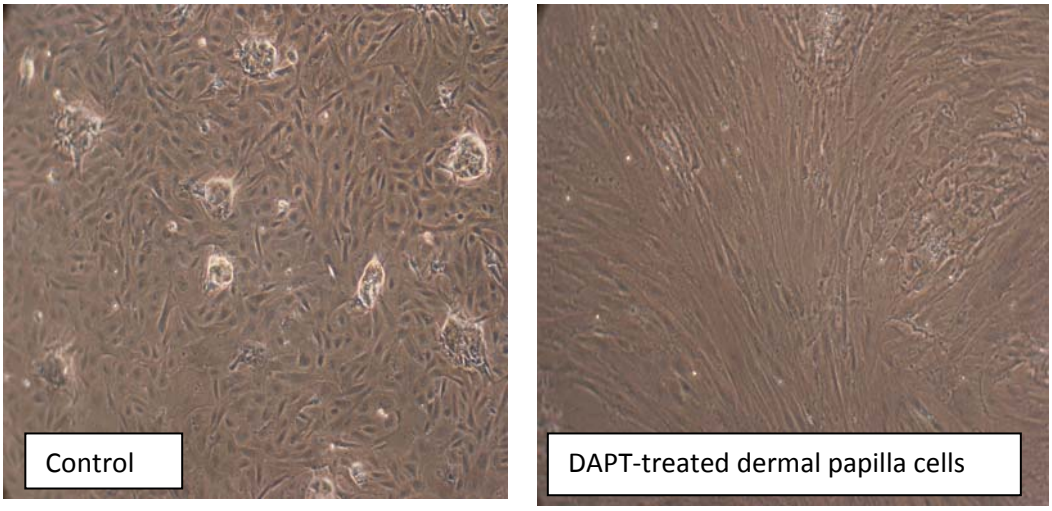


Figure 3. Blockage of Notch signalling by DAPT inhibits dermal papilla cell aggregation

DEVELOPMENT OF RAPID SCREENING METHODS FOR ASSESSING THE EFFECTS OF TREATMENTS ON SKIN DEVELOPMENT

As illustrated by the Delta/Notch work, it was clear from the outset of the sheepGENOMICS program that the ambitious goal of identifying gene pathways and therapeutic molecules with potential for beneficial manipulation of follicle and skin appendage formation would require test-bed systems which would allow screening of numerous candidates. We also needed a skin cell transfection system which would allow evaluation of the impact of gene changes on cell behaviour. Several models were developed in the wool subprogram to address these issues.

CELL CULTURE METHODS FOR SCREENING THERAPEUTICS WITH EFFECTS ON FOLLICLE INITIATION AND FIBRE GROWTH

Three cell culture methods were developed as screening tools: (a) a keratinocyte proliferation assay (Rufaut *et al.*, 2007); (b) a keratinocyte apoptosis assay (Rufaut *et al.*, 2007); and (c) the dermal papilla cell aggregation assay (Goldthorpe *et al.*, 2008). The latter relies on the unusual behaviour of papilla cells, which, on reaching confluence, begin 'clumping' in 3-dimensional aggregates spaced approximately equidistant from one another. As noted above, this behaviour is thought to recapitulate aspects of morphogenesis *in vivo*. Culture conditions were optimised to facilitate robust aggregation, and image analysis methods were developed for quantification of aggregate size and number. The value of this assay was evaluated by imposing a number of treatments which alter signalling pathways known to be important in morphogenesis *in vivo*, including the BMP, FGF and Wnt pathways. For example, lithium chloride is an inhibitor of GSK3B which mediates β-catenin degradation in the Wnt signalling pathway. Treatment with lithium diminished the size of DP cell aggregates in a dose-dependent fashion (Figure 4). This supports the gene expression results which indicated an important role for the Wnt signalling

pathway in follicle formation and implies a role for Wnt signalling in specifying follicle size and fibre diameter. Similarly, inhibition of BMP or FGF signalling by small molecules that target the receptors reduced aggregate size (Fig 5). Co-treatment with these inhibitors plus lithium chloride produced an additive effect.

Stratifin, a molecule involved in multiple cell signalling events, was identified in the microarray studies of hair cycle regulation. Apoptosis was induced in keratinocytes when stratifin expression was suppressed by RNAi, mediated by transfection with short interfering RNAs (Fig 6),. Stratifin and related molecules are potential candidates in wool growth regulation and should be the subject of further research in relation to regulation of wool growth (eg bioharvesting molecules).

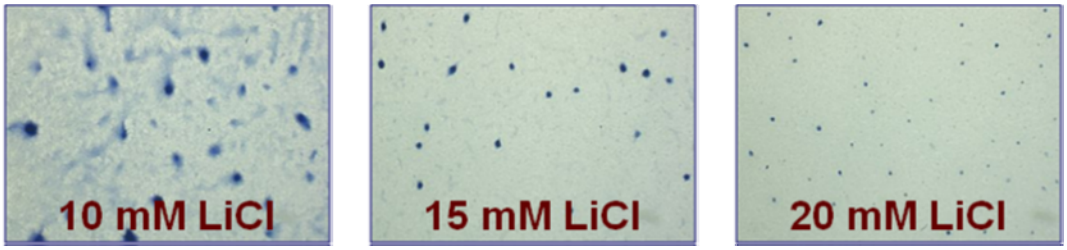


Figure 4. Lithium chloride, an agonist of the Wnt signalling pathway, affects dermal cell aggregate size in a dose-dependent manner. The aggregates of dermal papilla cells appear as dark blue clumps.

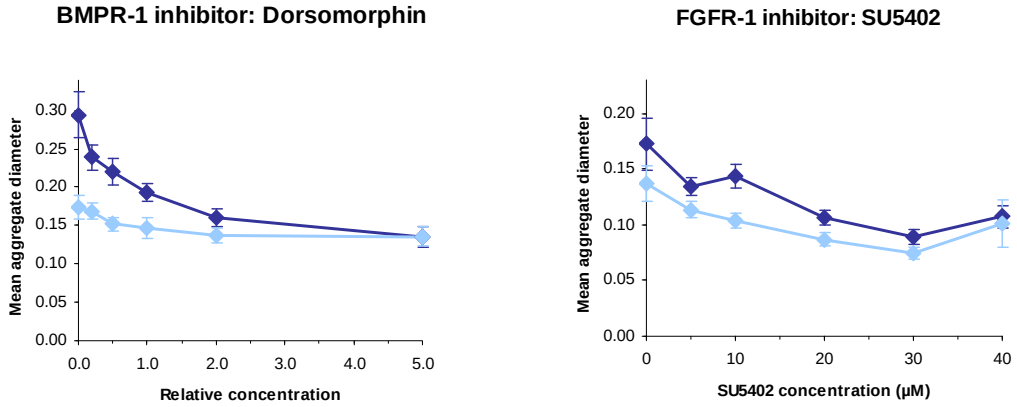


Figure 5. Inhibition of BMP and FGF receptors reduces aggregate size. Ovine dermal papilla cells were cultured in varying concentrations of dorsomorphin, an inhibitor of BMPR-1, and SU5402, an inhibitor of FGFR-1 (dark blue lines). Cells were also co-treated with the inhibitors and 10 mM lithium chloride (pale blue lines). Error bars show SEM for aggregate diameter. The cell culture system shows quantitative responses to changes in the signalling of developmental regulators that are likely to determine wool fibre diameter.

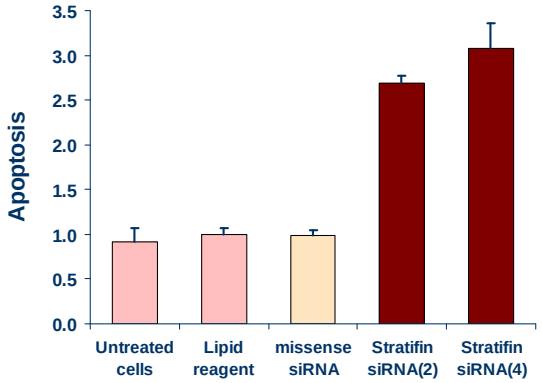


Figure 6. Treatment of ovine follicle keratinocytes with siRNAs targeting stratifin (columns 4-5) enhance apoptosis compared with control treatments (columns 1-3).

IN VITRO CULTURE OF EMBRYONIC MOUSE SKIN AS A TEST-BED FOR THERAPEUTIC MOLECULES TO ALTER WOOL FOLLICLE FORMATION

A mouse embryonic skin culture model was developed which allowed rapid screening of candidate molecules in a defined and quantifiable system. The advantages over the pregnant sheep model (see below) are that small quantities of therapeutic molecules can be tested, the gestation period is 20 days cf 147 days in sheep, the litter sizes are large (6-12), the strains of mice are inbred and therefore genetically-uniform, and the pattern of hair follicle formation is not unlike that of the developing foetal lamb. The method developed in the sheepGENOMICS program was based on that of Kashiwagi *et al.* (1997) with some modifications that allowed complete follicle formation to the point of hair growth. This model was used to determine the effects of manipulating the TGFβ superfamily signalling pathway. Two key components of that pathway are the activins and follistatins (see review by McDowall *et al.* 2008). Addition of exogenous follistatin stimulated follicle initiation *in vitro* by as much as 13%, while activinA induced a 20% reduction in follicle initiation. The culture of murine embryonic skin appears to be a responsive and robust system for testing the effects of signalling molecules on developmental events.

MANIPULATION OF FOETAL SKIN DEVELOPMENT USING A PREGNANT MOUSE MODEL

This model has many of the advantages of the murine embryonic skin culture system described above, but with the added advantage that the foetuses can be maintained into postnatal life to ascertain the long-term effects of treatment. Large litter sizes again provide a powerful experimental model. We used this model to investigate the effects of manipulation of the pentose phosphate pathway and retinoic acid signalling pathways. Manipulation of these pathways significantly influenced follicle initiation but a somewhat surprising result was that the vehicle control solution (which contained glucose), also significantly influenced follicle initiation. Further studies have validated this finding. The recent discovery that glucose availability can affect histone acetylation (Wellen *et al.* 2009), provides a potential mechanism whereby glucose could influence epigenetic events in developing foetal skin.

MANIPULATION OF FOETAL DEVELOPMENT IN PREGNANT SHEEP

Two methods of manipulating the development of foetal sheep skin *in vivo* were developed in the wool sheepGENOMICS program,: (1) by provision of therapeutics to the pregnant ewe (effective for manipulation of early stages of cutaneous development- days 0-50 pc). In this first

trimester most molecules supplied to the mother are delivered to the foetus via the foetal blood supply. Beyond day 50 the placental barrier forms and exposure of the foetus to the substances depends on the nature of the substance and its mode of delivery across placental tissues; (2) by provision of therapeutics directly to the foetus via intra-amniotic injections guided by ultrasound. This approach was used for pregnancy beyond day 50 and for substances known to be regulated by placental carrier systems.

An initial experiment designed to influence the cortisol axis in pregnant ewes at days 55-65pc (the time of primary follicle initiation) resulted in marked changes in the birthcoat scores of the resulting lambs. Metyrapone, which reduces cortisol biosynthesis, produced lambs which were hairier at birth ($P<0.003$) than control lambs, and those treated with betamethasone (Fig 7). The metyrapone lambs produced wool staples that were 10% longer than controls throughout life (up to 3 annual adult shearings) with no differences in staple strength or fibre diameter.

This is a major novel finding and one that has implications for all livestock species. It is the first 'proof of concept' that brief 'windows' of development exist in which therapeutic manipulations can induce lifelong changes in production traits. Further work has been conducted to identify the major gene network changes associated with this manipulation. Microarray data from the metyrapone experiment were analysed within a systems biology framework using Weighted Gene Co-expression Network Analysis (WGCNA). Four networks were created to determine those genes involved in metyrapone-mediated improvement of wool parameters. Using the WGCNA approach, we were able to detect co-expressed gene modules associated with metyrapone treatment. Gene ontology enrichment analysis of the genes comprising these modules identified networks associated with tube branching morphogenesis and with the BMP signalling pathway (Watson-Haigh, Kadarmideen, G. Nattrass, M. McDowall, N.M. Edwards and P.I. Hynd, unpubl. data). The known involvement of the BMPs and other members of the TGFβ-superfamily in skin development supports the notion that microarray analysis using this approach is a robust method. Interestingly the BMPs also featured strongly in the spatio-temporal expression studies referred to above. Furthermore, increased BMP signalling in association with a hairier birthcoat is consistent with the miniaturisation of papilla cell aggregates induced by BMP inhibition *in vitro*..

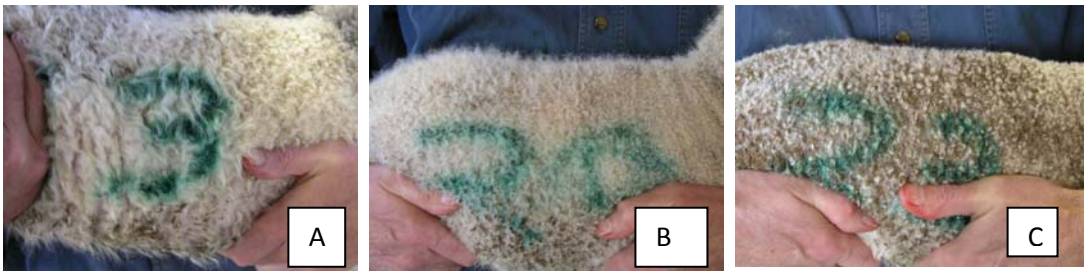


Figure 7. Lambs born after a 10-day treatment (days 55-65) with metyrapone (A), Vehicle (B) and betamethasone (C) differed in birthcoat score (BCS) and lifetime wool production. Lambs shown represent the mean BCS of each treatment (Mean± sem: A= 5.3±1.0; B= 3.7±1.3; C=2.6±1.9)

INVESTIGATIONS OF RECESSIVE BLACK (AGOUTI)

The white coat phenotype of domestic sheep breeds shows an autosomal dominant inheritance and has reached a high frequency in certain breeds as a result of selection for white fibres in an attempt to eliminate coloured sheep and coloured fibres. Self-colour black and badgerface are recessive pigmentation phenotypes of domestic sheep breeds caused by alleles at the agouti locus.

The dominant white or tan (A^{Wt}) agouti allele is responsible for the white wool phenotype in modern sheep breeds while the most recessive allele, non-agouti (A^a), results in black/brown wool (self-colour black). Another agouti allele, badgerface (A^b) is characterised by a pale dorsal, and darker ventral pattern; it is recessive to A^{Wt} and dominant to A^a. This project, lead by Dr Belinda Norris and colleagues at CSIRO Livestock Industries, has unravelled the mystery of the agouti locus, which has eluded researchers to date (Norris and Whan 2008). The principal findings are that the sheep dominant white allele (A^{Wt}) is characterised as having one or more extra copies of the gene at the agouti locus; that each point at which an additional agouti gene is inserted in the A^{Wt} allele (junction point) can be identified by a unique genomic sequence that spans the Junction Point; and finally that the recessive black alleles (A^a and A^b) each contain a single agouti gene with a dysfunctional promoter. The project team has developed an assay for counting Junction Points. However, the Junction Point assay is diagnostic of Carriers (ie heterozygotes A^{Wt}/A^a and A^{Wt}/A^b) only where the A^{Wt} allele contains a single junction point. Presently, for example, in our assays that detect two Junction Points, we cannot distinguish between a homozygous white animal in which each A^{Wt} allele has one junction point ($A^{Wt.1}/A^{Wt.1}$) and a Carrier with an allele with two Junction Points ($A^a/A^{Wt.2}$). Also for animals with higher numbers of junction points, there is a level of uncertainty in the junction point counts. Fortunately, we estimate that ~55% of Merino Carriers have only one junction point. We estimate that approximately 33.5% of Merino Carriers have 2 junction points (a triplicated agouti allele) and 11.5 % 3 junction points (a quadruplicated agouti allele) (Figure 1). These multiple junction point carriers cannot be classified as such by the junction point assay alone. Unfortunately a proposal to test for markers (SNP) adjacent to the duplicated region 3' breakpoints of recessive black and dominant white alleles to identify particular profiles (haplotypes) characteristic of different copy number carrier animals, failed to find an association between haplotype and agouti gene copy number or coat colour phenotype. Nevertheless the test has been developed to include a front-end, user friendly interface, which allows results to be readily interpreted based on parallel consideration of pedigree information.

GENE EXPRESSION THROUGHOUT THE HAIR GROWTH CYCLE

Drs Allan Nixon, Allan Pearson, Zhidong Yu and Nick Rufaut at AgResearch in New Zealand developed a wool shedding model based on the photoperiod-induced wool loss of Wiltshire Horn sheep (Nixon *et al.*, 2002). Gene expression profiling using an ovine cDNA microarray covering foetal initiation and postnatal hair cycling phases, revealed hundreds of differentially-expressed genes (Yu *et al.*, 2007). As hypothesised, many pathways are shared between follicle initiation in the foetus and hair cycling in the adult. Using Ingenuity Pathway Analysis® the differentially-expressed genes have been ordered into known gene networks. The opportunity now exists to identify inhibitors or stimulators of these pathways to control both foetal follicle formation and adult fleece growth.

IDENTIFICATION OF FUNCTIONAL GENES IN MUTANT SHEEP

Naturally-occurring or induced mutations, particularly in mice, have proven to be a valuable resource for studies of human disease. We reasoned that given the size of the Australian merino sheep population and a typical rate of spontaneous mutation, there should be a large number of phenotypes which would be valuable in dissecting functional genes involved in skin and follicle development. The media (including international outlets) were particularly interested in this project and it received widespread publicity, resulting in 31 unusual phenotypes being identified. These included a cohort of felting lustre mutants, sheep with very low follicle density, sheep with periodic shedding of fibres across the body, sheep with excessive skin wrinkle, sheep born with no wool follicles (so-called 'bald at birth'), sheep with unusual bare patches, and the previously described hypotrichosis phenotype. DNA, wool and skin samples have been taken from these

animals for analysis and cell lines or germplasm stored to enable rederivation of the phenotypes. To date, examination of gene expression and follicle and fibre structure in the felting lustre mutants has identified differences informative in defining the origin of fibre curvature and crimp (Li *et al.*., 2009).

GENES ASSOCIATED WITH RESISTANCE AND SUSCEPTIBILITY TO FLEECE ROT

During 2003-2004 genetically diverse resource flocks were measured for fleece rot resistance in artificial rain wetting shed trials at DPI NSW (Trangie) and CSIRO (Armidale). This provided fleece rot scores, blood and skin samples from two different genetic backgrounds (Merino and Merino x Romney) and genetic extremes (20 resistant and 20 susceptible animals from each genetic background). Samples of skin were collected to construct subtracted cDNA libraries and investigate gene expression patterns during the development of and recovery from fleece rot. Using skin samples from two resistant and two susceptible animals, six subtracted normalised cDNA libraries were made. From each library ~960 anonymous ESTs were printed together with CSIRO bovine and ovine skin ESTs in a 24K cDNA skin array. A microarray experiment of 31 array hybridisations was conducted. Analysis resulted in 297 differentially expressed (DE) array elements, the majority (72%) of which were anonymous ESTs from the six fleece rot subtracted cDNA libraries. SNPs within candidate genes amongst the group of differentially-expressed genes are currently being sought.

CONCLUSIONS

The Wool subprogram of the sheepGENOMICS program used a combination of molecular, biochemical and cellular approaches to dissect key spatio-temporal events in skin development. Identification of important genes, gene pathways and potential targets for therapeutic manipulation was attempted using cDNA micorarrays, qPCR of candidate genes, transfection of cells with genes of interest and assessment of their subsequent function, and 'best-guess' biochemical interventions. This approach has successfully identified potential targets for therapeutic intervention in skin development and has demonstrated that such intervention over brief 'windows' of foetal development can alter lifetime productivity.

REFERENCES

Chapman, R. E., Hopkins, P. S. and Thorburn, G. D. (1974). *Journal of Anatomy* **117**: 419.

Goldthorpe NT, Nixon AJ, Rufaut NW (2008). *Provisional patent, New Zealand (Assignees: AgReseach, Ltd. Australian Wool Innovation, Ltd).*

Kashiwagi M., Kuroki, T. and Huh, N. (1997). *Dev Biol.* **189:**22.

Li, S.W., Ouyang, H.S., Rogers, G.E., Bawden, C.S. (2009) *Experimental Dermatology* **18:**134.

McDowall M., Edwards N. M., Jahoda C. A., Hynd P. I. (2008) *Cytokine Growth Factor Rev* **19:** 415.

Nixon A. J.; Ford C. A.; Wildermoth J. E.; Craven A. J.; Ashby M. G. and Pearson A. J. (2002). *Journal of Endocrinology* **172:** 605.

Norris, BJ and Whan, VA (2008). *Genome Res.* **18:**1282**.**

Rufaut NW, Goldthorpe NT, Craven AJ, Wallace OA, Wildermoth JE, Pedersen G, *et al.*. (2007). *Proceedings of the New Zealand Society of Animal Production* **67:**326.

Stenn, K and Paus, R. (2001). *Physiological reviews* **81:**449.

Wellen KE, Hatzivassiliou G, Sachdeva UM, Bui TV, Cross JR, Thompson CB. (2009) *Science* **324:**1076.

Yu, Z.D.; Bawden, C.S.; Henderson, H.V.; Nixon, A.J.; Gordon, S.W.; Pearson, A.J. (2006) *Proceedings of the New Zealand Society of Animal Production* **66:** 129.

WEIGHTED CO-EXPRESSION NETWORKS SHED LIGHT ON THE MOLECULAR MECHANISM OF ACTION OF METYRAPONE ON WOOL FOLLICLE DEVELOPMENT

N.S. Watson-Haigh[1], H.N. Kadarmideen[2], M. McDowall[3], G.S. Nattrass[4], H.A. McGrice[3] and P.I. Hynd[3]

[1] CSIRO Livestock Industries, St Lucia [2] CSIRO Livestock Industries, Rockhampton; [3] The University of Adelaide and [4] South Australian Research & Development Institute (SARDI)

SUMMARY

The density of Merino wool follicles is established early in fetal development. This commercially important trait dictates wool fibre diameter which is the key driver of the price paid for wool. Merino lambs exposed to metyrapone (an inhibitor of cortisol synthesis) *in utero* show a lifetime alteration in wool growth parameters. Microarray data from a metyrapone treatment experiment were analysed within a systems biology framework using Weighted Gene Co-expression Network Analysis (WGCNA). Four networks were created to determine those genes involved in metyrapone mediated improvement of wool growth parameters. Using the WGCNA approach, we were able to detect co-expressed gene modules associated with metyrapone treatment. Gene ontology enrichment analysis of the genes comprising these modules identified a Bone Morphogenetic Protein (BMP4), a ligand known to be involved in hair/wool follicle development and expressed at the time of branching of secondary-derived follicles in Merino sheep.

INTRODUCTION

Merino sheep have a characteristically high follicle density of up to 60/mm^2 and a total of 10-100 million follicles compared to the estimated 5 million follicles in human skin. Primary follicles are the first to form, followed by secondary follicles and then secondary-derived follicles that branch from the secondary follicles. It is this high concentration of secondary-derived follicles which is a distinctive feature of Merino sheep. Branching of the secondary original follicles is essential in the Merino fleece as it is the major source of fine fibres (Hardy and Lyne 1956; Adelson *et al.* 2004). More detail is provided in a recent review (Rogers 2006). Fleece density, quality and length are important commercial traits, so improvements in these are beneficial.

A previous study showed that lambs exposed to metryrapone, an inhibitor of cortisol synthesis, *in utero* between day 55 and 65 of gestation have been shown to possess lifetime improvements in their fleeces (in prep.). The identification of genes responsible for this fleece improvement would facilitate the development of other pharmaceutical intervention strategies to improve wool growth. A microarray experiment was performed on foetal skin samples in an attempt to determine those genes involved in metyrapone mediated improvement of wool parameters. We follow a general framework for constructing gene co-expression networks (Zhang and Horvath 2005) and used the WGCNA R package (Langfelder and Horvath 2008). For details of network concepts and terminology used see (Langfelder and Horvath 2008; Dong and Horvath 2007).

MATERIALS AND METHODS

Experimental Design. Twenty time-mated (synchronised with progesterone sponges and then artificially inseminated) pregnant Merino ewes were allocated to 4 equally sized treatment groups receiving daily intramuscular injections of a control (5ml of 0.6 M tartaric acid in 5% Tween 20) or metyrapone (5 ml of 1.76g metyrapone in 0.6 M tartaric acid in 5% Tween 20) between day 55

and 65 of gestation. 2cm midside foetal skin samples were collected at two time points (Table 1). Samples were snap frozen in liquid nitrogen prior to RNA extraction.

Microarray Data. The RNA from 16 foetal skin samples was hybridised to Affymetrix GeneChip® Bovine Genome Array. Data were subjected to quality control procedures to ensure that: 1) hybridisations were uniform as assessed by pseudo-array images produced by the Bioconductor (Gentleman *et al.* 2004) affyPLM package and 2) RNA was not degraded as assessed by the Bioconductor affy package (Gautier et al. 2004).

Gene expression values were generated using the median-polish summarisation method following GC-RMA background correction of PM probes and quantile normalisation.

Table 1. Experimental design showing treatment groups

Group	Treatment	Treatment period (day of gestation)	Sample collected (day of gestation)	Number of single pregnancies
1	Control	55-59	60	4
2	Metyrapone	55-59	60	5
3	Control	55-65	67	4
4	Metyrapone	55-65	67	3

Weighted Network Construction. Four networks were constructed using sample groups as follows: network A) group 1 and 3; network B) group 2 and 4; network C) group 1 and 2; and network D) group 3 and 4. For each network non-changing genes with low mean levels of expression ($< 10^{0.805}$) and low variance ($< 10^{-6}$) across all the arrays were excluded to reduce computational complexity in later steps. The power adjacency function was applied to the co-expression measurement, the absolute Pearson correlation coefficient, to derive the adjacency matrix: $a_{ij} = |cor(x_i, x_j)|$. The value of the power function exponent, β, was chosen based on the accuracy of the model fit, R^2 (Zhang and Horvath 2005). We chose β in the interval [0,30] which maximised R^2 while maintaining a high level of mean connectivity. Using the gene connectivity measure, we constructed a network based on the top 8,000 most connected genes.

Modules were defined for the top 8,000 most connected genes using the dynamic hybrid tree cutting algorithm of the dynamicTreeCut R package (Langfelder *et al.* 2008). Interesting modules were defined as those which had a high Pearson correlation (≥ 0.8) with the day of gestation.

Gene Ontology Enrichment Analysis. Gene lists were validated for biological relevance using GOEAST (Zheng and Wang 2008), a gene ontology (GO) enrichment analysis tool. GO enrichment tools are used to identify GO terms which are statistically overrepresented in a list of genes and provide a means to identify biological functions and processes at play.

RESULTS AND DISCUSSION

Microarray Data. Pseudo-array images showed that several of the arrays contained small hybridisation artefacts (data not shown) but nothing so unusual as to warrant their removal (Bolstad *et al.* 2005). The slopes and profiles of our RNA degradation plot indicated no major issues with RNA quality (data not shown).

Weighted Co-Expression Network Analysis. Figure 1 shows a Topological Overlap Measure (TOM) plot for network B with module membership indicated below and to the right of the dendrograms. It shows a high level of topological overlap between genes within a module

(indicated by the dark squares along the diagonal), with only a small amount of inter-module overlap (indicated by mostly pale off-diagonal regions).

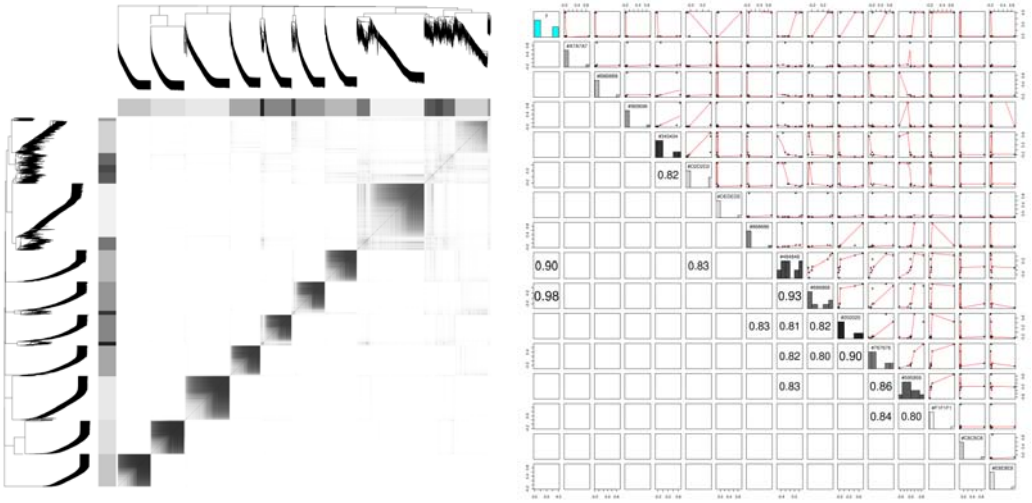


Figure 1. WGCNA results for network B. Left) TOM plot with module membership indicated next to dendrograms. Right) Pairwise correlations plot between module eigengenes and experimental parameter, day (1st column/row). Correlation values ≥ 0.8 are shown.

Effective time-mating resulted in no modules highly correlated with experimental parameter day, as expected, in Network A. On the other hand, network B provided 2 modules, "#484848" and "#686868" with correlations of 0.9 and 0.98 with experimental parameter, day (Figure 1). This indicates differential gene expression between day 60 and 67 in metyrapone samples which is not seen in controls. Module eigengenes, the first principle component of a module, for the two modules show higher levels of gene expression in the metyrapone day 67 samples compared to the metyrapone day 60 samples (Figure 2).

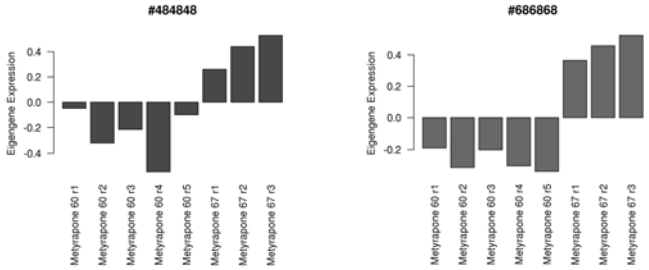


Figure 2. Eigengene expression for modules "#484848" and "#686868".

Gene Ontology Enrichment Analysis. The "#686868" module (267 genes) showed enrichment for 58 Biological Process GO terms (a selection shown in Table 2), 33 Molecular Function GO terms and 9 Cellular Component GO terms. Among several interesting tube/patterning terms were GO:0048754 (branching morphogenesis of a tube) and GO:0030509 (BMP signalling pathway) which comprise of Bone Morphogenetic Protein 4 (BMP4) and Chemokine (C-X-C motif) Receptor 4 genes. BMPs are members of the transforming growth factor-β (TGF-β) superfamily of ligands, with important roles in a myriad of biological activities (Waite and Eng 2003; Shi and Massagué 2003). They have been shown to be involved in the regulation of hair follicle initiation

and development, including the branching of secondary follicles (Menzies *et al.* 2009) which is important in Merino sheep.

Table 2. A selection of Biological Process GO terms enriched in the "#686868" module

GOID	Definition	No. of genes	P-value
GO:0051056	regulation of small GTPase mediated signal transduction	9	0.004
GO:0007389	pattern specification process	3	0.018
GO:0010646	regulation of cell communication	11	0.018
GO:0001763	morphogenesis of a branching structure	2	0.028
GO:0048754	branching morphogenesis of a tube	2	0.028
GO:0030509	BMP signaling pathway	1	0.051
GO:0001569	patterning of blood vessels	1	0.051
GO:0009880	embryonic pattern specification	1	0.051
GO:0035239	tube morphogenesis	2	0.056
GO:0009799	determination of symmetry	1	0.070

CONCLUSIONS

Metyrapone treatment of pregnant ewes between 55 and 65 days of gestation result in pronounced differences in gene co-expression patterns. Using a WGCNA approach we were able to recover modules, which are highly correlated to experimental parameters of interest, which contain several hundred genes. Using a GO enrichment analysis we showed that these modules contain genes known to be involved in the regulation of hair follicle initiation and development. This information can now be used to refine the timing of drug administration or used to identify compounds with similar or more potent effects on wool growth.

ACKNOWLEDGMENTS

The project was financially supported by Sheep Genomics, an initiative of Australian Wool Innovation Limited and Meat and Livestock Australia. The Systems Biology Project is supported by CSIRO OCE Postdoctoral Fellowship and Transformational Biology Capability Platform.

REFERENCES

Adelson, D.L., Cam, G.R., DeSilva, U. and Franklin, I.R. (2004) *Genomics*. **83**:95

Bolstad, B.M., Collin, F., Brettschneider, J., Simpson, K. (2005) In *Bioinformatics and Computational Biology Solutions Using R and Bioconductor (Statistics for Biology and Health)*.

Dong, J. and Horvath, S. (2007) *BMC Syst. Biol.* **1**:24.

Gautier, L., Cope, L., Bolstad, B.M. and Irizarry, R.A. (2004) *Bioinformatics*. **20**:307

Gentleman, R.C., Carey, V.J., Bates, D.M. and Bolstad, B. (2004) *Genome Biology*. **5**:R80.

Hardy, M. and Lyne, A. (1956) *Aust. J. Biol. Sci.* **6**:423

Langfelder, P. and Horvath, S. (2008) *BMC Bioinformatics*. **9**:559.

Langfelder, P., Zhang, B. and Horvath, S. (2008) *Bioinformatics*. **24**:719.

Menzies, M., Stockwell, S., Brownlee, A. and Cam, G. (2009) *Exp. Dermatol.*:In Press.

Rogers, G.E. (2006) *Exp. Dermatol.* **15**:931

Shi, Y. and Massagué, J. (2003) *Cell*. **113**:685

Waite, K.A. and Eng, C. (2003) *Nat. Rev. Genet.* **4**:763

Zhang, B. and Horvath, S. (2005) *Stat. Appl. Genet. Mol. Biol.* **4**:Article 17.

Zheng, Q. and Wang, X. (2008) *Nucl. Acids Res.* **36**(suppl_2):W358.

GENETIC PROGRESS IN THE AUSTRALIAN SHEEP INDUSTRY

A.A. Swan[1], D.J. Brown[1] and R.G. Banks[2]

[1]Animal Genetics and Breeding Unit[*], University of New England, Armidale, NSW 2351
[2]Meat and Livestock Australia, C/- University of New England, Armidale, NSW 2351

SUMMARY

Genetic progress was estimated in the major breed groups of the Sheep Genetics database as a means of monitoring changes in the productivity of sheep enterprises. There has been substantial improvement in productivity between 1990 and 2005 based on representative breeding objectives for each breed, ranging from $10 per ewe for Border Leicesters (0.7 standard deviations of the breeding objective) to $17 per ewe for Terminal Sire breeds (2.9 standard deviations). Rates of progress have increased significantly since 2000 for the Terminal Sire, Border Leicester and Coopworth breeds, while Merinos have maintained a relatively constant and favourable rate of progress over the whole time period. Compared to simple breeding programs simulated for each breed group, Terminal Sires are exceeding the simulated potential rate of gain, Border Leicesters and Coopworths are approaching the potential gain, while Merinos are achieving only one third of the potential gain.

INTRODUCTION

Genetic progress is a key profit driver for the Australian sheep industry, and as a consequence there has been significant industry investment in performance recording and genetic evaluation systems since the late 1980's. The meat and dual-purpose maternal breeds have had access to Estimated Breeding Values (EBV's) through the LAMPLAN system since 1989, with across flock evaluations becoming available in the mid 1990's. In the wool sector, across flock evaluation began in the early 1990's through the Central Test Sire Evaluation program. Larger evaluations using on-farm data began in the late 1990's with the advent of Merino Benchmark and Merino Genetic Services. These systems, including LAMBPLAN, were merged under the banner of Sheep Genetics in 2005 (Brown *et al.* 2007). In this paper, we compare the genetic progress in predicted profitability since 1990 in the main breed groups serviced by Sheep Genetics.

MATERIALS AND METHODS

Breeding objectives. Estimates of the change in profitability were based on breeding objectives for each of the four main Sheep Genetics breed group databases: Terminal Sire, Border Leicester, Coopworth, and Merino. Breeding objectives were calculated using SheepObject (Swan *et al.* 2007), and selection indexes based on these were derived. The advantage of using SheepObject indexes is that they are expressed in dollar terms. Although these are not currently used by breeders, index values on individual sires are highly correlated with the de-facto standard indexes shown in Table 1.

Estimated genetic trends. Genetic trends were estimated using results from the December 2008 LAMBPLAN and MERINOSELECT evaluations. EBV's for sires were used to calculate SheepObject index values, and these were averaged by year of birth ranging from 1990 to 2005. Trends were expressed both as dollars per ewe per year (abbreviated to $ per ewe hereafter), and scaled by the standard deviation of the breeding objective.

[*] AGBU is a joint venture of NSW Department of Primary Industries and The University of New England

Table 1. Number of sires in each breed group, industry index, and correlation between SheepObject index and industry index values for sires

Breed group	Sires	Industry index	Correlation
Border Leicester	3,468	Border$	0.82
Coopworth	1,828	Coopworth$	0.90
Merino	9,382	M10SS	0.94
Terminal Sire	18,118	Carcass+	0.99

Predicted genetic trends. These realised industry trends were compared with predicted trends derived using the SelAction computer program (Rutten *et al.* 2002) which calculates theoretical progress achievable based on key breeding program parameters. Response to index selection on the SheepObject breeding objectives was calculated for a relatively simple breeding program, with a flock size of 500 ewes and 10 rams mated annually. There were five age classes for ewes and one for rams, with parents selected at one year and having their first progeny at two years of age. All information from relatives was assumed to be available to estimate breeding values, as would be the case for flocks recording full sire and dam pedigrees. Truncation selection across ewe age classes was practiced.

Simple sets of selection criteria were used, including: yearling clean fleece weight, weaning weight, and post-weaning weight, fat, and muscle depth for Border Leicesters and Coopworths; yearling weight, clean fleece weight, fibre diameter, and CV of fibre diameter for Merinos; and birth weight, weaning weight, and post-weaning weight, fat, and muscle depth for Terminal Sire breeds. These traits are easily and relatively inexpensively measured early in an animal's lifetime. It was assumed that all selection criteria were measured on all relatives.

RESULTS

Estimated genetic trends are shown in Figure 1. Between 1990 and 2005, the Terminal Sire and Coopworth breeds have improved by around $17 per ewe, while Border Leicesters and Merinos have improved by approximately $10 per ewe. However, when expressed in terms of the standard deviation of the objective, the Terminal Sires were well ahead of the other breeds, showing an improvement of almost 3 standard deviations. Coopworths have improved by 1.3 standard deviations, and Border Leicesters and Merinos by 0.7 to 0.8 standard deviations. The rate of progress shows an increase from 2000 in the Terminal Sire breeds, Coopworths and Border Leicesters, while the Merinos show relatively constant improvement over the entire period.

Standard deviations of objectives and indexes calculated in SelAction are shown in Table 2, together with index accuracy and the standard deviation observed in sire indexes. The latter are approximately double the predicted values due to a combination of higher accuracy of progeny test information included in the sire indexes, and the influence of substantial genetic differences between flocks in these databases.

Table 2. Standard deviations (SD) of objective and index and accuracy from selection index predictions, and observed standard deviation of actual sire indexes

Breed group	SD objective	SD index	Accuracy	SD sire indexes
Border Leicester	14.4	3.8	0.27	7.2
Coopworth	13.7	4.6	0.34	8.6
Merino	13.5	4.5	0.33	7.9
Terminal Sire	6.1	2.7	0.45	7.5

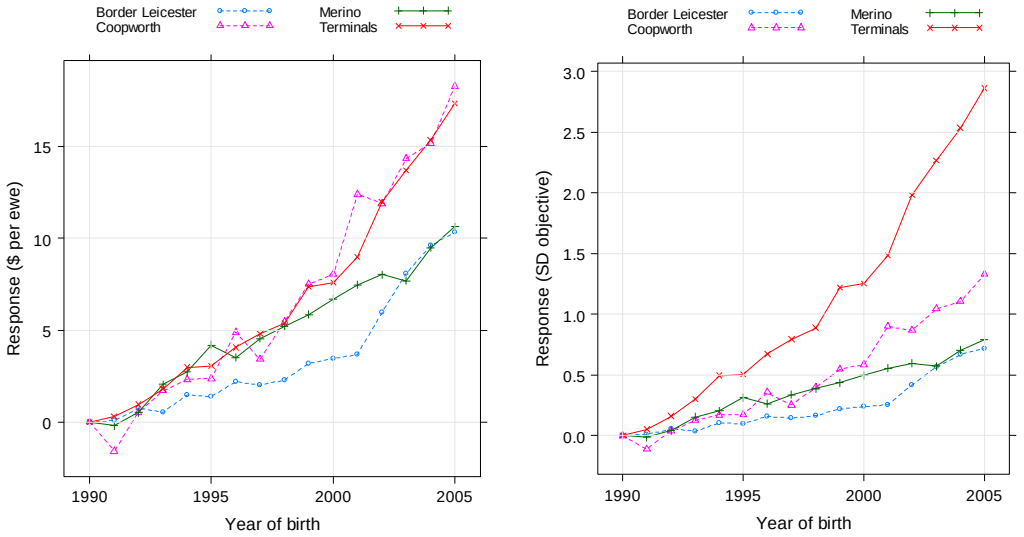


Figure 1. Genetic progress in the main Sheep Genetics breed groups between 1990 and 2005. The left panel expresses response in dollars per ewe per year, while the right has been scaled by the standard deviation of the respective breeding objectives for each breed.

Predicted annual responses from SelAction and realised annual responses post 2000 are shown in Table 3. Under the within flock breeding programs modelled, predicted annual responses ranged from 1.8 to 2.4 dollars per ewe, or 0.14 to 0.30 objective standard deviations. Compared to these figures, the realised response in industry ranged from 30% of the predicted response for Merinos to 111% for Terminal Sire breeds. The Border Leicesters and Coopworths were intermediate achieving approximately 80% of the potential response.

Table 3. Predicted and post 2000 realised annual response in industry

Breed group	Annual response ($ per ewe)			Annual response (SD objective)		
	Predicted	Realised	Ratio (%)	Predicted	Realised	Ratio (%)
Border Leicester	2.0	1.7	85	0.14	0.11	79
Coopworth	2.4	1.8	75	0.17	0.13	76
Merino	2.3	0.7	30	0.15	0.05	33
Terminal Sire	1.8	2.0	111	0.30	0.33	110

DISCUSSION

There has been significant albeit variable genetic progress across the major breed groups in the Australian Sheep industry since 1990. This progress has led to substantial improvements in productivity, with our estimates suggesting cumulated increases of $10 to $17 per ewe depending on breed.

All breed groups made steady progress through the 1990's, more so the Terminal Sires. From 2000 on, the rate of gain in the Terminal Sire, Border Leicester, and Coopworth breeds has increased significantly. There are several possible reasons for this increase, including the introduction of young sire programs, introduction of Carcass+ and maternal dollar indexes, more widespread use of fat and muscle scanning, a greater focus on data quality, and in the case of Terminal Sires, the move to a fully across flock and across breed analysis in 1999. In Border

Leicesters and Coopworths, greater awareness of the variation available for genetic improvement was stimulated by the Maternal Central Progeny Test program (Fogarty *et al.* 2001).

These developments have allowed the Border Leicester and Coopworth breeds to approach the rates of gain predicted in our simulated breeding program, and for the Terminal Sire breeds to exceed the prediction. The latter is entirely feasible because the simulation did not model the higher selection intensities achievable when using young sire programs, or across flock and across breed effects.

Merinos on the other hand have made consistent but slow progress over the entire period of evaluation, with the rate of gain being only one third of the rate predicted in our simulation. It could be argued that the Merino is a very diverse breed, with a wide range of breeding objectives in use across flocks, and that this might be a limitation when estimating progress across the whole breed. However, apart from at the extremes, say ultra-fine compared to dual purpose sheep, the majority of breeding objectives are highly correlated. A second argument to explain slower progress is the perception that the traits are more difficult to select on, but as can be inferred from the predicted annual gain of $2.3 per ewe in Table 3, this is not the case. Although there are economic antagonisms between traits including fleece weight, fibre diameter and staple strength, these can be overcome with an appropriate breeding objective and selection index.

There are three more likely limitations to genetic progress in Merinos. First is that generation intervals are typically longer compared to the other breeds: the average age of Merino sires in Sheep Genetics is currently 3.3 years, while Terminal Sire breeds average 2.6 years. Second is that there has been a lower level of pedigree recording in the Merino: in 2000 only 30% of Merino progeny had both sire and dam recorded, and although there has been an increase to more than 50% in recent years, other breed groups are currently recording full pedigrees on more than 95% of progeny. This lower level of pedigree recording in Merinos limits progress through lower selection accuracy. The third limitation is that Merino breeders traditionally place more emphasis on traits outside of the objective, and in particular visual traits. All of these limitations can be overcome, and there is evidence that leading Merino breeders are making the necessary changes to their breeding programs.

CONCLUSIONS

Rates of genetic progress in the Australian sheep industry since 1990 have been impressive. While the Terminal Sire breeds have been the stand out performers, Border Leicesters and Coopworths have also increased their rate of gain since 2000. Merino breeders need to make a similar jump in progress.

ACKNOWLEGMENTS

This research was supported by the Australian sheep industry and funded by Meat Livestock Australia and Australian Wool Innovations.

REFERENCES

Brown, D.J., Huisman, A.E., Swan, A.A., Graser, H-U., Woolaston, R.R., Ball, A.J., Atkins, K.D. and Banks, R.G. (2007) *Proc. Assoc. Advmt. Anim. Breed. Genet.* **17**:187.

Fogarty, N. M., Cummins,L., Gaunt, G., Hocking Edwards, J., Edwards, N., Lees, K. and Morgan, J. (2001). *Proc. Assoc. Advmt. Anim. Breed. Genet*. **14:**123.

Rutten, M.J.M., Bijma, P., Woolliams, J.A., van Arendonk, J.A.M. (2002) *Journal of Heredity* **93 (6)**:456.

Swan, A.A., van der Werf, J.H.J., and Atkins, K.D. (2007) *Proc. Assoc. Advmt. Anim. Breed. Genet.* **17**:483.

BREEDING PLAIN-BODIED FINE WOOLS - NO PROBLEM!

S. Hatcher, K.D. Atkins and K.J. Thornberry

NSW Dept of Primary Industries, Orange Agricultural Institute, Forest Rd Orange NSW 2800

SUMMARY

The inheritance of wrinkle in fine wool sheep was investigated in a mixed bloodline flock run in western NSW. Both repeatability (0.6) and heritability (0.4) estimates agreed with published estimates from broader flocks as did the phenotypic and genetic correlations with a suite of assessed and measured traits. The correlations suggest that selection for plain-bodied easy care fine wool sheep can be achieved while maintaining wool production and quality despite medium to strong antagonistic relationships with wool production and staple strength. Furthermore the genetic relationships with liveweight were favourable, albeit small.

INTRODUCTION

The very early (to 1885) history of the Australian sheep industry was characterised by "a passionate focus on wool quality, so long as the wool produced was 80's or 90's nobody bothered much what the sheep cut" (Austin 1943). The early fine wools were said to carry "a beautiful, long staple of high yielding wool on a completely plain body" (Dun and Eastoe 1970). The development of the Peppin strain in the 1860s saw the focus shift to production based on the concept of "return per head" (Austin 1943). The introduction of the heavily wrinkled Vermont Merino about 20 years later was an attempt to increase the wool cut of fine wool, again to increase the return per head. Despite the subsequent downfall of the Vermont in the early 1900s, many stud breeders to this day believe that a moderate degree of skin wrinkle is required to maintain wool weight through its association with fleece density (Crook and James 1991; Sutton *et al.* 1995). The wool production focus continued for 130 years until the mid 1990s when consumer preferences for lightweight clothing and the requirement for increased processing efficiency drove a trend towards the production of finer wool (Swan *et al.* 2008). Fine wool producers are now seeking to breed 'easy care' plainer bodied animals with acceptable carcase traits while maintaining or improving their wool quality and production (A. Casey pers comm.). This paper reports on the genetic relationships between wrinkle score and assessed and measured traits in a mixed bloodline flock run in western New South Wales.

MATERIALS AND METHODS

The sheep used in this study were wethers born between 1991 and 1996 in the CSIRO Fine Wool Project flock at Armidale (Swan *et al.* 2000) and transferred to the Condobolin Agricultural Research and Advisory Station (ARAS) following their hogget shearing. Prior to transfer, each wether was assessed for neck and body wrinkle using a 1-6 score (1 = plainest and 6 = wrinkliest) (Turner *et al.* 1953) at 10 months of age. All wethers were mulsed when 2 months old so breech wrinkle was not scored. Total wrinkle was calculated by adding the neck and body wrinkle scores. Repeat assessments were made on the 1991 drop at 3 and 4 years of age and the 1994, 1995 and 1996 drops at 5, 4 and 3 years of age respectively. The wethers remained at Condobolin for 4 consecutive shearings. At each annual shearing a suite of measurements and assessments were taken including wool production, wool quality, liveweight (Hatcher *et al.* 2005). The wethers were assessed for disease resistance using a scoring system for the incidence of fleece rot (1-8 score), bacterial stain (1-5 score), dermatitis (1-9 score) and flystrike (1-7) where low scores indicate an absence of the disease. The occurrence of disease within the flock was low, ranging from 0.3% for bacterial stain to 6.4% for fleecerot.

Approximately 1,800 individual animals, each with an identified sire, were involved in this analysis and represented 11 bloodlines (6 superfine, 3 fine and 2 medium wool). ASReml (Gilmour *et al.* 2006) was used to estimate variance components using a general linear mixed model by residual maximum likelihood. A univariate analysis was undertaken for each wrinkle score (neck, body and total). The model included the fixed effects of drop (5 levels: 1991, 1993 to 1996), year (9 levels: 1993 to 2001) and flock (11 levels) together with significant interactions. Random effects were estimated for sire (σ^2_s), animals within sire (σ^2_b), within animals (σ^2_w) and between flocks (σ^2_f). Repeatability ($\sigma^2_s + \sigma^2_b$)/ ($\sigma^2_s + \sigma^2_b + \sigma^2_w$), heritability ($4\sigma^2_s$)/($\sigma^2_s + \sigma^2_b + \sigma^2_w$) and their standard errors were calculated from the univariate analyses. Genetic and phenotypic covariances were estimated using a series of bivariate analyses involving the wrinkle scores and each of the other traits. Fixed effects and interactions were fitted as appropriate from the univariate analyses. Genetic and phenotypic correlations with standard errors were estimated from the appropriate covariances in ASReml.

RESULTS AND DISCUSSION

The average wrinkle scores were 3.15 ± 0.21, 2.22 ± 0.20 and 5.36 ± 0.38 for neck, body and total wrinkle respectively. Differences between flocks in wrinkle score were not large. Flock means ranged from 2.70 - 3.54 for neck wrinkle, 2.31 - 2.94 for body wrinkle and 5.00 - 6.46 for total wrinkle. Between and within animal variation in wrinkle score together accounted for 90, 91 and 88 % of the phenotypic variance for neck, body and total wrinkle respectively (Table 1). Between sire variation was the least important source for each of the 3 wrinkle scores and between flock variation was equally low. The 3 wrinkle scores were highly repeatable (0.50-0.61) with high heritability (0.35-0.44). Both the repeatability (Beattie 1961; Young *et al.* 1960a) and heritability (Beattie 1962; Brown and Turner 1968; Gregory 1982; Groenewald *et al.* 1999; Mortimer and Atkins 1993; Mortimer *et al.* 2009; Young *et al.* 1960b) estimates agree with those estimated for medium to broad wools. Wrinkle scores are therefore under a similar degree of genetic control in fine wool flocks as fibre diameter and clean fleece weight (Hatcher and Atkins 2000) and will respond to single trait selection in much the same manner as medium to broader wools (Turner *et al.* 1970).

Table 1. Variance components, heritability and repeatability (± se) of wrinkle scores

Wrinkle	Variance component					Repeatability	Heritability
	σ^2_w	σ^2_b	σ^2_s	σ^2_p	σ^2_f	t^2	h^2
Neck	0.24±0.01	0.32±0.02	0.06±0.01	0.62±0.02	0.07±0.03	0.61±0.02	0.40±0.07
Body	0.29±0.01	0.24±0.02	0.05±0.01	0.57±0.02	0.04±0.02	0.50±0.02	0.35±0.07
Total	0.77±0.03	1.04±0.05	0.22±0.04	2.04±0.06	0.21±0.11	0.62±0.02	0.44±0.08

The phenotypic and genetic correlations (0.71 and 0.99 respectively) between neck and body wrinkle were strong, positive and close to unity indicating that either score will provide a reliable visual description of the degree and quantity of wrinkles on an individual animal. This finding agrees with previous work (Beattie 1962; Jackson and James 1970; Lewer *et al.* 1995; Mortimer and Atkins 1993; Mortimer *et al.* 2009). For the purpose of brevity total wrinkle will be used when discussing the correlations with other traits.

Phenotypic correlations between wrinkle and the other assessed traits tended to be negligible except for density (-0.3) (Table 2). So plain-bodied fine wool sheep would tend to have denser fleeces that would be marginally softer with slightly better defined crimp and enhanced style. Phenotypic correlations between wrinkle and the measured traits were generally of a higher magnitude, but still ranged from negligible to low. Most measured traits

had weak complimentary phenotypic associations with wrinkle, plain-bodied animals tended to have lower fibre diameter, lower variability in fibre diameter (both standard deviation and coefficient of variation), increased yield and lower resistance to compression. Plain-bodied individuals would also phenotypically tend to maintain liveweight. However antagonistic correlations were identified between wrinkle and wool production, with plain-bodied animals cutting lighter greasy and clean fleeces with shorter staples. There was no phenotypic relationship between wrinkle and staple strength.

Table 2. Phenotypic (r_p) and genetic correlations (r_g) (±se) between wrinkle scores and wool production and quality traits

Trait	Neck		Body		Total	
	r_p	r_g	r_p	r_g	r_p	r_g
Assessed traits						
Handle	0.12±0.02	0.30±0.13	0.11±0.02	0.28±0.14	0.12±0.02	0.30±0.13
Dust penetration	0.02±0.02	0.12±0.19	0.01±0.02	0.04±0.19	0.02±0.02	0.08±0.18
Crimp definition	0.09±0.02	0.14±0.14	0.05±0.02	0.11±0.14	0.08±0.02	0.12±0.14
Staple structure	0.02±0.02	0.08±0.17	0.04±0.02	0.07±0.18	0.03±0.02	0.07±0.17
Density	-0.28±0.02	-0.65±0.10	-0.27±0.02	-0.63±0.11	-0.30±0.02	-0.63±0.10
Colour	-0.01±0.02	-0.08±0.15	-0.03±0.02	-0.06±0.15	-0.02±0.02	-0.07±0.14
Style	0.08±0.02	0.02±0.20	0.10±0.02	0.02±0.21	0.09±0.02	0.02±0.20
Fleece rot	0.02±0.01	0.12±0.17	no estimates available			
Bacterial stain	no estimates available		-0.00±0.02	-0.21±0.14	0.00±0.02	-0.13±0.14
Flystrike	0.02±0.02	-0.06±0.21	0.01±0.02	-0.18±0.22	0.01±0.02	-0.11±0.21
Body wrinkle	0.71±0.01	0.99±0.00	no estimates available			
Measured traits						
GFW (kg)	0.36±0.02	0.61±0.09	0.30±0.02	0.52±0.10	0.35±0.02	0.56±0.09
CSY (%)	-0.17±0.02	-0.40±0.12	-0.15±0.02	-0.38±0.12	-0.17±0.02	-0.39±0.11
CFW (%)	0.23±0.02	0.39±0.11	0.18±0.02	0.29±0.12	0.22±0.02	0.34±0.11
LWT (kg)	-0.03±0.02	-0.05±0.15	-0.07±0.02	-0.12±0.15	-0.06±0.02	-0.07±0.15
FD (µm)	0.19±0.02	0.34±0.12	0.16±0.02	0.31±0.13	0.19±0.02	0.31±0.12
FDSD (µm)	0.26±0.02	0.40±0.11	0.23±0.02	0.33±0.12	0.26±0.02	0.36±0.11
FDCV (%)	0.20±0.02	0.42±0.12	0.18±0.02	0.33±0.13	0.20±0.02	0.37±0.12
FC (°/mm)	0.01±0.02	-0.03±0.15	0.03±0.02	-0.05±0.15	0.02±0.02	-0.04±0.14
SL (mm)	-0.23±0.02	-0.54±0.11	-0.24±0.02	-0.51±0.11	-0.26±0.02	-0.53±0.10
SS (N/ktex)	-0.00±0.02	0.06±0.14	0.03±0.02	0.09±0.14	0.02±0.02	0.08±0.13
RTOC (kpa)	0.19±0.02	0.33±0.12	0.17±0.02	0.29±0.13	0.19±0.02	0.31±0.12
Colour (Y-Z)	0.07±0.02	0.36±0.16	0.06±0.02	0.29±0.17	0.07±0.02	0.33±0.16

The genetic correlations between wrinkle and the assessed and measured traits were of the same sign, except for the incidence of flystrike, and stronger than the phenotypic correlations (Table 2). The genetic correlations between the assessed traits and wrinkle ranged from high and complimentary (density) to negligible and antagonistic (staple structure). However most were negligible. Selection for reduced wrinkle will lead to denser fleeces with a softer handle. No change would be expected in dust penetration, staple structure and assessed colour. The incidence of fleece rot would decrease and while there was evidence of an increased incidence of flystrike, the standard errors for these 2 estimates were both large. Greasy and clean fleece weights both had economically antagonistic relationships with wrinkle but the correlation with greasy fleece weight was higher (0.6 and 0.3 respectively). Staple length had a medium negative, economically favourable relationship (-0.5). Genetic correlations with the other

measured traits were low and favourable so selection for plain-bodied fine wool sheep would produce finer, higher yielding fleeces with longer staple length, lower variation in fibre diameter, improved resistance to compression and measured colour. Liveweight, fibre curvature and staple strength were only weakly correlated with wrinkle.

CONCLUSIONS

Inheritance of wrinkle in fine wool sheep is similar to broader bloodlines. The antagonistic correlation between wrinkle and clean fleece weight is of a similar magnitude to that reported between fibre diameter and clean fleece weight (Safari *et al.* 2005), and it has been demonstrated that simultaneous improvement in these two traits is achievable (Mortimer *et al.* 2006). The genetic relationships between wrinkle and liveweight were positive, albeit small indicating that selection for plain-bodied fine wools will not negatively impact carcase weight.

REFERENCES

Austin, H.B. (1943) "The Merino Past, Present and Probable" Grahame Book Company: Sydney.

Beattie, A.W. (1961) *Q. J. Agric. Sci.* **18**:437.

Beattie, A.W. (1962) *Q. J. Agric. Sci.* **19**:17.

Brown, G.H. and Turner, H.N. (1968) *Aust. J. Agric. Res.* **19**:303.

Crook, B.J. and James, J.W. (1991) *Proc. Aust. Assoc. Anim. Breed. Genet.* **9**:372.

Dun, R.B. and Eastoe, R.D. (1970) "Science and the Merino breeder" Victor C.N. Blight Government Printer: Ultimo.

Gilmour, A.R., Gogel, B.J., Cullis, B.R. and Thompson R. (2006) "ASReml User Guide Release 2.0." VSN International Ltd, Hemel Hempstead.

Gregory, I.P. (1982) *Aust. J. Agric. Res.***33**:355.

Groenewald, P.G., Olivier, J.J. and Olivier, W.J. (1999) *S.A. J. Anim. Sci.***29**:174.

Hatcher, S. and Atkins, K.D. (2000) *Asian-Aust. J. Anim. Sci.* **13** Sup. A:293.

Hatcher, S., Atkins, K.D. and Thornberry, K.J. (2005) *Proc. Assoc. Advmt. Anim. Breed. Genet*. **16**:314.

Jackson, N. and James, J.W. (1970) *Aust. J. Agric. Res.* **21**:837.

Lewer, R.P., Woolaston, R.R. and Howe, R.R. (1995) *Aust. J. Agric. Res.* **46**:379.

Mortimer, S.I. and Atkins, K.D. (1993) *Aust. J. Agric. Res.* **44**:1523.

Mortimer, S.I., Robinson, D.L., Atkins, K.D., Brien, F.D., Swan, A.A., Taylor, P.J. and Fogarty, N.M. (2009) *Anim. Prod. Sci.* **49**:32.

Mortimer, S.I., Taylor, P.J. and Atkins, K.D. (2006) In "Trangie QPLU$ Merinos - Open Day 2006" p. 7editor CE Pope, NSW Department of Primary Industries, Orange.

Safari, E., Fogarty, N.M., Gilmour, A.R. (2005) *Livest. Prod.Sci.* **92**:271.

Sutton, J.M., Williams. A.J., Nicol, H.I. and Thornberry, K.J. (1995) *Wool Tech. Sheep Breed.* **43**:183.

Swan, A.A., Purvis, I.W. and Piper, L.R. (2008) *Aust. J. Exp. Agric.* **48**:1168.

Swan, A.A., Purvis, I.W., Piper, L.R., Lamb, P.R. and Robinson, G.A. (2000) In "Finewool 2000 Breeding for Customer Needs:' p 65 editors A.A. Swan and L.R. Piper CSIRO Livestock Industries and The Woolmark Company.

Turner, H.N., Brooker, M.G. and Dolling, C.H.S. (1970) *Aust. J. Agric. Res.* **21**:955.

Turner, H.N., Hayman, R.H., Riches, J.H., Roberts, N.F. and Wilson, L.T. (1953) "Physical definition of sheep and their fleece for breeding and husbandry studies with particular reference to Merino sheep" CSIRO Div. Anim. Health and Prod. Report No.4 Melbourne.

Young, S.S.Y., Turner, H.N. and Dolling, C.H.S. (1960a) *Aust. J. Agric. Res.* **11**:257.

Young, S.S.Y, Turner, H.N. and Dolling, C.H.S. (1960b) *Aust. J. Agric. Res.* **11**:604.

HERITABILITY AND PHENOTYPIC CORRELATIONS FOR BREECH STRIKE AND BREECH STRIKE RESISTANCE INDICATORS IN MERINOS

J.L. Smith, H.G. Brewer and T. Dyall

FD McMaster Laboratory, CSIRO Livestock Industries, Armidale NSW 2350

SUMMARY

Mulesing is a means of blowfly strike control in Australian Merinos. Among potential mulesing alternatives, selective breeding for resistance is widely viewed to be the best long-term solution. Using data from the CSIRO Breech Strike Resistance Flock at Armidale (n=1656), heritability and phenotypic correlations between breech strike and indicator traits (including breech and crutch cover, wrinkles, dags, urine stain and fleece characteristics) are reported. Weaner breech strike heritability was estimated at 0.32 (0.11). With the exception of dags (0.09) all of the breech strike indicators were at least moderately heritable (>0.20). Breech wrinkle and dags were the indicator traits most closely correlated phenotypically with breech strike (both 0.22).

INTRODUCTION

Selection of Australian Merinos for breech strike resistance is a potential alternative to mulesing as a means of reducing breech strike (James 2006). It is undesirable on both ethical and economic grounds to artificially challenge animals with flystrike for selection purposes, so knowledge of potential indirect selection criteria is necessary to aid incorporation of breech strike resistance into Merino breeding programs. Using the CSIRO Armidale Breech Strike Resistance Flock, heritability and phenotypic correlations for breech strike and a suite of potential indicator traits were estimated with the aim of determining which of those indicators are the most promising candidates as selection criteria for breech strike resistance in Merino breeding programs.

METHOD

Animals. The Breeding for Breech Strike Resistance Project is an evaluation of the impact of not mulesing and the effectiveness of breeding for breech strike resistance using indicator traits. The program is being carried out in two different environments, Armidale in NSW representing the summer rainfall zone and Mt Barker in WA representing the winter rainfall zone. The flock at each location comprises 3 lines – an unselected control (UC), a commercial improvement line (CI, utilising a base of unselected ewes mated to plain breech sires), and a plain breech line (PB, both ewes and sires selected for plain breech characteristics). The Armidale flock was established in 2005 from a combination of purchased 2005 drop ewe weaners (n=644, from 11, mostly finewool, eastern states flocks) and 600 CSIRO Chiswick Station ewes, screened into the 3 selection lines based on phenotypic adult body wrinkle. Replacement ewes for the PB line are selected using breech strike resistance indicator traits (primarily breech cover, crutch cover and breech wrinkle), and for the UC and CI lines are chosen at random and the number balanced across lines. Link sires are used across years and sites. Most of the sires represented are from industry flocks (of a range of wool types) evaluated in Sheep Genetics (Brown *et al.* 2007). Half of the animals within each selection line are mulesed and the other half are unmulesed. No preventative chemical treatments for flystrike are applied. The 2005 drop animals, sourced from industry had unknown pedigree. In total, 34 sires from 17 Merino studs are represented with 14-78 progeny per sire group.

Measurements. Incidence of breech and body strike was recorded during the first flystrike season for weaner animals born between 2005 and 2008 inclusive (n=1656). Flystrike usually occurs in

the Armidale environment in the period Oct-Apr inclusive and the animals were approximately 2 mths of age at the start of the flystrike season. The sheep were inspected for flystrike at least 3 times weekly throughout the flystrike season. Struck animals had details of the strike recorded and were treated with short-acting insecticide (Extinosad^{TM}). The breech strike trait reported here is a count (natural logarithm transformed) of breech strikes during the flystrike challenge period.

Breech strike indicator traits were recorded at post-weaning stage (6mths, late February). The breech traits were breech cover, crutch cover, body wrinkle, breech wrinkle and dags (AWI 2007)). Urine stain was scored 1-5 on females. Yearling (pre-shearing) greasy wool colour and fleece rot scores were also recorded. All of these traits were approximately normally distributed.

Statistical analyses. Univariate analysis of covariance was conducted using ASReml to determine fixed effects (Gilmour *et al.* 2002). The factors included were selection line, sire and dam wool type, year/property-of-origin, mulesed/unmulesed, sex, damage, birth-rearing type, lambing management flock, scorer, and cannonbone length (as an indicator of body size) and weaning bodyweight were fitted as covariates. The sire and dam wool types, based on Sheep Genetics (2006) groupings, were effectively genetic group effects fitted to adjust for sheep type differences among selection lines. Bivariate sire models were fitted to estimate heritabilities and phenotypic correlations among the breech traits. At this stage the dataset is not sufficiently robust to justify reporting of the genetic correlations, the standard errors of which are large. Initially, repeatability analysis of flystrike (by month) was attempted, but the flystrike data were too sparse for that statistical model to operate effectively, which is why the trait reported upon here is the total count of flystrikes for the season.

RESULTS

Flystrike rates. 2006 and 2007 were the higher flystrike challenge years (Table 1). Mulesed animals consistently had lower breech strike rates than unmulesed animals and the extent of selection line differences varied with year. Breech strike rates were not significantly different among selection lines but were affected by sire wool type ($P<0.001$), where progeny from ultrafine/superfine sires had higher breech strike rates than other wool types. Year/property-of-origin ($P<0.01$), mulesing ($P<0.001$), sex ($P<0.05$, males lower than females), birth-rearing type ($P<0.01$, singletons higher than multiples) and weaning bodyweight ($P<0.001$, where smaller animals were more likely to get breech strike) were also significant effects on breech strike.

Main effects on indicator traits. Animals in the selected lines had significantly lower body and breech wrinkle and breech and crutch cover than those in the control line ($P<0.001$). Mulesed animals had significantly less breech wrinkle and cover than unmulesed animals ($P<0.001$). Progeny of ultrafine/superfine sires had higher wrinkle, breech and crutch cover than progeny of either fine/fine medium or medium/strong wool sires ($P<0.001$). Males had less body wrinkle than females ($P<0.001$), and animals born as singletons had lower body and breech wrinkle than those born and reared as multiples ($P<0.05$).

Heritability and phenotypic correlations. Body and breech wrinkle were correlated (0.29), and more wrinkly animals tended to have more urine stain (Table 2). Breech and crutch cover were correlated (0.29), but both were poorly correlated with wrinkles. Among the indicator traits, breech wrinkle and dags were those most closely correlated with breech strike (both 0.22). Breech strike and breech cover were not correlated. Yearling greasy wool colour and fleece rot were correlated (0.20), but neither of those traits were correlated with any of the other breech strike indicators, nor breech strike itself (-0.05 – 0.09), except that wool colour was correlated with urine stain (0.20).

Body and breech strike were not correlated (0.08). Body strike was correlated with fleece rot (0.23), but not wool colour (0.07).

Table 1. Percentage flystrikes by year, selection line (UC=unselected control, CI = commercial improvement, PB = plain breech), and mulesing group (M=mulesed, UM=unmulesed)

Drop	Line	n		Body strikes[#] (%)		Breech Strikes (%)		Total strikes (%)	
		M	UM	M	UM	M	UM	M	UM
2005	UC	105	111	6	0	0	7	6	7
	CI	109	109	0	1	2	14	2	15
	PB	105	105	0	0	0	8	0	8
2006	UC	70	72	3	6	17	90	20	96
	CI	67	66	6	9	0	24	6	66
	PB	70	75	1	1	1	22	3	24
2007	UC	38	38	8	13	5	71	13	84
	CI	43	44	16	20	2	36	19	57
	PB	43	43	14	7	2	33	16	40
2008	UC	52	55	4	0	4	25	8	25
	CI	60	59	3	2	3	14	7	15
	PB	60	60	0	2	5	8	5	10

[#] includes poll strikes

Table 2. Heritability (bold), phenotypic variance (V_p) and phenotypic correlations among breech strike and indicators.

Trait	V_p	Body wrinkle	Breech wrinkle	Breech cover	Crutch cover	Dag	Urine stain	Breech strike
Body wrinkle	0.38	**0.25 (0.10)**	0.29	0.03	0.11	-0.00	0.22	0.04
Breech wrinkle	0.57		**0.36 (0.12)**	0.02	0.09	0.06	0.21	0.22
Breech cover	0.45			**0.23 (0.09)**	0.29	0.01	0.05	0.01
Crutch cover	0.35				**0.47 (0.14)**	0.04	-0.05	0.09
Dag	0.43					**0.09 (0.06)**	0.00	0.22
Urine stain	0.40						**0.30 (0.20)**	0.04
Breech strike	0.06							**0.32 (0.11)**

s.e. on all correlations 0.03-0.04 except for those with urine stain which were 0.06-0.07

DISCUSSION

Low breech wrinkle and dag were the characteristics with greatest effect on breech strike rate. There was no evidence that lower breech cover reduced breech strike. Breech wrinkle and dag have also been shown to be associated with breech strike in the Mt Barker, WA flock, but there is evidence in that flock that reduced breech strike rate is also associated with lower breech cover (Greeff and Karlsson 2009). The mean breech cover score in the WA flock is lower than that in the Armidale flock which may explain the different outcomes observed.

Breech strike and, with the exception of dags, all of the breech traits were moderately heritable. However, only breech wrinkles and dags were sufficiently correlated with breech strike to be useful selection criteria. Published heritability estimates for dags vary widely with environment and age (Greeff and Karlsson 1998, 1999; Woolaston and Ward 1999). Compared to winter rainfall areas, dags in the summer rainfall environment is a somewhat 'transient' trait. This may make dags of limited use as a selection criterion for breech strike in summer rainfall areas.

Breech wrinkle however, is moderately correlated with breech strike, and is also heritable, repeatable across ages, and exhibits approximately normal distribution in the Merino population (Raadsma and Rogan 1987; Lewer *et al.* 1995; JL Smith, *unpublished data*). Breech wrinkle is also highly correlated genetically with body wrinkle (reviewed by James 2006), for which there exists considerable genetic data, including that body wrinkle measured at different ages is highly repeatable (Robinson *et al.* 2007; Hatcher *et al.* 2009). All of these features indicate breech wrinkle to be a good candidate as an indirect selection criterion for breech strike in Merinos.

The phenotypic correlation between body and breech wrinkle reported here (0.29) is lower than reports in the literature of wrinkle score correlations across the body (James 2006), and also lower than at other ages in this experiment (not reported here). This may be due to the wool length (6mths) at wrinkle scoring employed here (i.e. there is lower variance in body wrinkle with more wool length, J.L. Smith, *unpublished data*).

Preliminary evidence from the project indicates the phenotypic correlation between weaner and yearling age breech strike is moderate (not detailed here). A review by Raadsma and Rogan (1987) concluded repeatability across ages of fleece rot and body strike to be age and environment (incidence) dependent. There are insufficient data at this point to determine the correlation between weaner and adult breech strike. However, given that wrinkle score is repeatable across ages and breech wrinkle is correlated with breech strike, it could be expected that selection for low wrinkle in young animals (weaner or yearling) will be a useful selection criterion for lifetime breech strike resistance.

Breech strike was affected by sire wool type (rather than selection line), where progeny of ultrafine/superfine sires were more likely to be affected by breech strike than those with broader wool type sires. There were however, statistically significant differences among both the sire wool types and the selection lines in all of the breech strike indicator traits except dags. This might suggest that with further selection, line differences in breech strike will emerge.

ACKNOWLEDGEMENTS

AWI provided financial support to this Project. Ray Honnery, Grant Uphill and Brian Dennison contributed invaluable technical support in checking, treating and recording flystrikes. Andrew Swan was instrumental in designing and setting up this Project and Laurie Piper provided statistical advice.

REFERENCES

AWI (2007) 'Visual Sheep Scores'. AWI and MLA.

Brown, D.J., Huisman, A.E., Swan, A.A., Graser, H-U., Woolaston, R.R., Ball, A.J., Atkins, K.D. and Banks, R.G. (2007) *Assoc. Advmt. Anim. Breed. Genet.* **17:**187.

Gilmour, A.R., Gogel, B.J., Cullis, B.R., Welham, S.J. and Thompson, R. (2002) ASReml User Guide Release 1.0 VSN International Ltd. Hemel Hempstead, HP1 1ES, UK.

Greeff, J.C. and Karlsson, L.J.E. (1998) *6th World Congr. Genet. Appl. Anim. Prod.* **24:**63.

Greeff, J.C. and Karlsson, L.J.E. (1999) *Assoc. Advmt. Anim. Breed. Genet.* **13:**508.

Greeff, J.C. and Karlsson, L.J.E (2009) These proceedings.

Hatcher, S., Atkins, K.D. and Thornberry, K.J. (2009) *Assoc. Advmt. Anim. Breed. Genet.* **18:** 330.

James, P.J. (2006) *Aust. J. Agric. Res.* **46:**1.

Lewer, R.P., Woolaston, R.R. and Howe, R.R. (1995) *Aust. J. Agric. Res.* **46:**379.

Sheep Genetics (2006) MERINOSELECT wool types, accessed 08/04/2009. http://www.sheepgenetics.org.au/ViewImage.aspx?ITEM=59&NAME=MERINOSELECT%20Types.pdf

Raadsma, H.W. and Rogan, I.M. (1987) In 'Merino Improvement Programs in Australia', p. 321, editor. B.J. McGuirk, Australian Wool Corporation, Melbourne.

Robinson, D.L., Mortimer, S.I., Swan, A.A. and Purvis, I.W. (2007) *Assoc. Advmt. Anim. Breed. Genet.* **17:** 336.

Woolaston, R.R. and Ward, J.L. (1999) *Assoc. Advmt. Anim. Breed. Genet.* **13:**512.

THE RELATIONSHIPS BETWEEN CRUTCH COVER SCORE AND PRODUCTION AND EASY CARE TRAITS IN MERINO SHEEP

D.H. Smith[1], E. Safari[2], F.D. Brien[1], K.S. Jaensch[1] and R.J. Grimson[1]

[1] South Australian Research and Development Institute, Roseworthy, SA, 5371
[2] School of Biological Science, Flinders University, Bedford Park, SA, 5001

SUMMARY

A preliminary study of the relationship between crutch cover score, measured at hogget and adult age, and production and easy care traits found moderately negative phenotypic correlations between hogget score and hogget fibre diameter and body weight, and a moderately positive phenotypic correlation with hogget face cover. Genetic correlations between hogget crutch cover score and hogget fibre diameter, staple strength, staple length and body weight were moderately to strongly negative and strongly positive for coefficient of variation of fibre diameter and face cover. Correlations with hogget fleece weight, yield, neck and body wrinkle were not significantly different to zero. Heritability of crutch cover score at both ages was high. Strategies to eliminate the need for mulesing by placing selection emphasis on decreased hogget crutch cover should result in genetic gains in hogget performance for body weight, staple strength, staple length and to a lesser degree plainness but attention should be paid to fibre diameter to prevent it from increasing. Better estimates of the genetic correlations between hogget crutch cover score, other hogget traits and adult traits are needed to predict future hogget and adult flock performance, however our preliminary results suggest that selection to reduce hogget crutch cover should give minor gains in adult body weight and reproductive performance, with negligible impact on adult fleece weight in the current flock.

INTRODUCTION

James (2006) argued that breeding sheep resistant to flystrike provides the best long term alternative to surgical mulesing as long as those sheep are then managed within a best practice integrated pest management program**.** Murray *et al*. (2007) demonstrated that selection of Merino ewe lambs for low wrinkle, dag and urine score and high breech bareness can be as effective as mulesing in preventing breech blowfly strike. Edwards *et al.* (2009) concluded that crutch cover score was moderately to highly heritable, irrespective of the age at measurement. They also concluded that the genetic correlation with economically important traits, other than weight of belly wool and skirtings, was low however they did not attempt to define the relationship between hogget score and adult performance.

Breeding sheep that are less susceptible to breech strike, when un-mulesed, will involve placing increased selection emphasis on easy care traits such as reduced breech wrinkle, and decreased breech and crutch cover. The effect of this emphasis on current and later age production is of considerable interest to producers. This preliminary study examines the phenotypic and genetic correlations of crutch cover score with production and easy care traits in Merino sheep.

MATERIALS AND METHODS

Data. Mulesed ewe hoggets from SARDI's five South Australian Merino Selection Demonstration Flocks were scored for crutch cover at hogget shearing (16-18 months of age). The ewe hoggets represented 87 sires. The scoring system for crutch cover is described in Edwards *et al*. (2009*)* where 1 = bare, 5 = woolly, and conforms to the scoring system described in the AWI Visual Sheep Scores guide (2007). Unskirted greasy fleece weight including belly wool was recorded at

shearing. Yield, fibre diameter, coefficient of variation of fibre diameter, staple length and staple strength were measured on mid-samples taken prior to shearing. Body weight, face cover score (1 = open, 5 = muffled), neck wrinkle score (1 = plain, 5 = wrinkled), and body wrinkle score (1 = plain, 5 = wrinkled) were recorded 10-14 days after the hogget shearing. Adult crutch cover score, weight of belly wool, skirted and unskirted greasy fleece weight, fibre diameter, coefficient of variation of fibre diameter, and body weight were also recorded at shearing in 2009 for ewe hoggets retained as breeding ewes. Records were also available on number of lambs weaned during 2008 and lifetime number of lambs weaned per lambing opportunity. Summary statistics are given in Table 1.

Table 1. Summary statistics for production and easy care traits

Trait	N	Mean	S.D.	Range
Hogget				
Clean fleece weight	875	4.1	1.1	1.6 - 7.1
Greasy fleece weight (unskirted, including belly)	875	5.7	1.5	2.4 - 10.0
Fibre diameter	875	19.5	2.0	14.9 - 27.1
Yield	875	70.7	4.6	57.3 - 82.7
Coefficient of variation for fibre diameter	875	21.5	3.0	13.6 - 31.5
Staple strength	875	27.7	12.3	5.1 - 79.0
Staple length	875	102.2	19.7	63.4 - 156.5
Body weight	875	43.3	6.2	19.5 - 63.5
Face cover score	873	1.8	0.6	1 - 4
Neck wrinkle score	876	2.5	0.8	1 - 5
Body wrinkle score	876	1.5	0.6	1 - 4
Crutch cover score	865	3.3	0.8	1 - 5
Adult				
Weight of belly wool	401	0.4	0.1	0.0 - 0.7
Skirted greasy fleece weight (not including belly)	403	5.7	0.9	3.9 - 8.5
Unskirted greasy fleece weight (not including belly)	402	6.5	0.9	4.7 - 9.4
Fibre diameter	403	20.9	1.8	16.5 - 26.9
Coefficient of variation for fibre diameter	403	16.7	1.8	12.6 - 22.2
Body weight	403	70.5	7.8	49.7 - 93.2
No. of lambs weaned 2008	430	1.3	0.7	0 - 3
Av. no. of lambs weaned per lambing opportunity	769	1.0	0.5	0 - 2.5
Crutch cover score	400	2.8	0.9	1 - 5

Statistical analysis. Variance components were estimated with ASREML software (Gilmour *et al*. 2006). The model included birth type (single or multiple), rearing type (raised as a single or multiple), age of dam (2 to 7 years), birth year (2004 or 2005), flock (measured performance, classer assessment, elite wool, fibre meat plus, and control) as fixed effects and a direct animal genetic component as a random effect. For the analysis of adult ewe crutch score reproduction status (number of lambs weaned in 2008) was fitted as another fixed effect in the model. Phenotypic and genetic correlations were estimated from bivariate analysis of hogget crutch cover score with hogget and adult traits.

RESULTS AND DISCUSSION

The distribution of hogget crutch score was 0.5%, 17.1%, 38.3% 37.9% and 6.2% for scores 1 to 5 respectively. For adult crutch score the values were 10.5%, 23.8%, 44.0%, 20.8% and 1.0%.

Only flock and birth year were significant ($P<0.05$) for both hogget and adult crutch scores. However for adult crutch score, flock and birth year became non-significant after the inclusion of the number of lambs weaned in 2008 in the analytical model. Type of rearing and year of birth were the only significant fixed effects for face cover score, while type of birth, flock and year of birth were significant for both wrinkle scores. A significant ($P<0.05$) flock by year of birth effect was observed only for body wrinkle score.

Estimated heritabilities were 0.54 ± 0.11 and 0.48 ± 0.18 for hogget and adult crutch scores, respectively. The high heritability estimates for crutch score in hogget and adult are similar to those found by Edwards *et al.* (2009), but slightly higher than those reported by Scobie (2007). Estimated heritabilities for face cover, neck and body wrinkle scores were 0.38 ± 0.10, 0.33 ± 0.10 and 0.34 ± 0.11, respectively which agree with those reported by Mortimer *et al.* (2009).

Phenotypic and genetic correlations between hogget crutch cover score and hogget traits are shown in Table 2. Moderately negative phenotypic correlations were found between hogget crutch cover score and hogget fibre diameter and body weight, and a moderately positive phenotypic correlation with hogget face cover. The genetic correlations had high standard errors but suggest a moderate to strong negative relationship with hogget fibre diameter, staple strength, staple length and body weight and a strong positive relationship with coefficient of variation of fibre diameter and face cover. Correlations with hogget fleece weight, yield and neck and body wrinkle were not significantly different to zero. Edwards *et al* (2009) also showed negative, albeit weaker, correlations with fibre diameter but slightly positive correlations with greasy fleece weight.

Table 2. Correlations (± s.e.) between hogget crutch cover score and other hogget traits

Traits	Phenotypic	Genetic
Clean fleece weight	0.01 ± 0.04	-0.12 ± 0.19
Greasy fleece weight	0.02 ± 0.04	-0.03 ± 0.18
Fibre diameter	-0.20 ± 0.04	-0.37 ± 0.15
Yield	-0.00 ± 0.04	-0.11 ± 0.17
Coefficient of variation of fibre diameter	0.17 ± 0.04	0.58 ± 0.21
Staple strength	-0.16 ± 0.04	-0.27 ± 0.16
Staple length	-0.16 ± 0.04	-0.27 ± 0.15
Body weight	-0.32 ±0.03	-0.51 ± 0.12
Face cover score	0.29 ± 0.03	0.61 ± 0.14
Neck wrinkle score	0.13 ± 0.04	0.23 ± 0.19
Body wrinkle score	0.07 ± 0.04	0.18 ± 0.19

The phenotypic correlations between hogget crutch cover score and adult traits are shown in Table 3. There were insufficient records on adult animals to provide meaningful estimates of genetic correlations between hogget crutch score and adult traits. At the phenotypic level, hogget crutch cover score appears to be moderately and positively correlated with adult belly wool weight, only weakly correlated with adult unskirted greasy fleece weight but uncorrelated with skirted fleece weight. Slightly unfavourable phenotypic correlations between hogget crutch cover and adult fibre diameter are evident in the data but the relationship with body weight and lifetime reproductive performance is slightly favourable.

Table 3. Phenotypic correlations (± s.e.) between hogget crutch cover score and adult traits

Traits	Phenotypic
Weight of belly wool	0.39 ± 0.05
Skirted greasy fleece weight (not including belly)	0.01 ± 0.05
Unskirted greasy fleece weight (not including belly)	0.10 ± 0.05
Fibre diameter	-0.13 ± 0.06
Coefficient of variation for fibre diameter	0.25 ± 0.05
Body weight	-0.17 ± 0.05
No. of lambs weaned 2008	-0.04 ± 0.05
No. of lambs weaned per lambing opportunity	-0.10 ± 0.05
Adult crutch cover score	0.57 ± 0.04

In conclusion, these preliminary results indicate that selection for decreased hogget crutch cover may result in genetic gains in hogget body weight, staple strength, staple length and plainness but attention needs to be paid to fibre diameter to prevent it from increasing. Such selection on hogget performance should also lead to minor current flock gains in adult body weight and reproductive performance, with only slightly unfavourable impacts on adult fleece weight. Notwithstanding, precise estimates of genetic correlations are needed between hogget crutch cover and wrinkle scores with later-age productivity before robust selection indices incorporating visual traits can be offered to industry. More exact estimates should be possible by combining our data with that available in other datasets ie the Falkiner Sheep Genome flock, the Adelaide University Calkookera flock and the Breeding for Breech Strike Resistance project. In the interim any independent culling level method used to select sheep with more resistance to flystrike, as described by Greeff *et al.* (2008), should be structured to allow sufficient selection pressure to be placed on production traits, especially fibre diameter, to ensure present production levels are maintained.

ACKNOWLEDGEMENTS

The SA Selection Demonstration Flocks (SDFs) received major funding support from Australian Wool Innovation and SARDI. We acknowledge Dr. Raul Ponzoni as the founder of the SDFs. Following the conclusion of the project in July 2006, funding for several projects through the Sheep CRC has enabled a percentage of the SDF ewes to be retained.

REFERENCES

Edwards, N.M., Hebart, M.L. and Hynd, P.I. (2009). *Anim. Prod. Sci.*,**49:**56

Gilmour, A. R., Cullis, S. J., Welham, S. J., and Thompson, R. (2006) ASReml User Guide. New South Wales Agriculture, Orange, Australia.

Greeff, J. C., Karlsson, J. E., Slocombe, S., Jones, K. and Underwood, N. (2008) *Agribusiness Livestock Updates, Perth, W.A. 1-2nd July, 2008.*

James, P.J. (2006) *Aust. J. Exp. Ag.* **46**:18

Murray, B., Karlsson, J. and Greeff, J. (2007) *Agribusiness Livestock Updates, Perth, W.A. 24-25th July, 2007.*

Mortimer, S.I., Robinson, D.L., Atkins, K.D., Brien, F.D., Swan, A.A., Taylor, P.J. and Fogarty, N.M. (2009). *Anim. Prod. Sci.* **49**:32

Scobie, D.R., O'Connell, D.O., Morris, C.A. and Hickey, S.M. (2007). *Aust. J. Agric. Res.* **58:**161

Visual Sheep Scores Guide. Australian Wool Innovation (2007). www.wool.com.au/publications.

GENETIC IMPROVEMENT PROGRAMS FOR AQUACULTURE SPECIES IN DEVELOPING COUNTRIES: PROSPECTS AND CHALLENGES

R.W. Ponzoni, N. H. Nguyen and H.L. Khaw

The WorldFish Center, Jalan Batu Maung, 11960 Bayan Lepas, Penang, Malaysia

SUMMARY

Aquaculture in developing countries is largely based on unimproved fish strains. There is ample evidence indicating the potential of genetic improvement programs and a range of selection methods may be used. Examples of the application of mass, cohort, within family, and combined between-within family are given. The methods are discussed in terms of their effectiveness and suitability. It is concluded that in principle all the methods can work well, provided the selection program is started with a population with a broad genetic base and that during its conduct a balance is struck between selection intensity and containment of inbreeding. Limitations to the implementation of genetic improvement programs and the challenges they face are discussed.

INTRODUCTION

Aquaculture is predicted to play a major and ever increasing role in meeting human needs for protein. In terrestrial animal and plant species genetic improvement programs have made a substantial contribution to productivity and viability. By contrast, most aquaculture stocks in current use in developing countries are genetically similar or inferior to wild, undomesticated stocks. A range of methods of varying complexity is available for selection purposes, but their suitability for different circumstances is not always clear. In this paper we briefly present the main selection methods that have been used or advocated, and discuss their virtues and shortcomings (a more detailed description of the methods is given by Ponzoni et al. 2006). When possible, we make reference to practical examples of their application. We also present evidence about the economic worth of genetic improvement programs and discuss some of the challenges faced when implementing such programs in aquatic animals.

APPROACHES TO GENETIC IMPROVEMENT

Aquatic animals allow the implementation of several approaches to genetic improvement. These include hybridization and cross breeding, chromosome manipulation, sex control, transgenesis, and selective breeding. These are almost always mentioned in aquaculture genetics reports, papers and meetings without making a judgement about their relative practical value. For instance, it is seldom, if ever, stated that of all the genetic approaches only selective breeding offers the opportunity of continued genetic gain, that the gains made can be permanent, that it is the only approach in which the gain can be transmitted from generation to generation, and that gains in a nucleus can be multiplied and expressed in thousand or millions of individuals in the production sector (Ponzoni et al. 2007, 2008). In the cases they are useful the other approaches result in 'once off' expressions of the benefit. They may be applied at the multiplication (hatchery) level, but not at the nucleus level.

SELECTION APPROACHES

General. We present the different selection approaches in increasing order of complexity, beginning with the simplest one. In each case, we refer to specific requirements that may constitute a limitation for their implementation in developing countries. Note that we assume that there is genetic variation for the trait(s) of interest in the population undergoing selection and that

it does not suffer from problems (e.g. bottlenecks, inbreeding) created by earlier genetic mismanagement. Such problems could undermine the effectiveness of any selection program (e.g. Teichert-Coddington and Smitherman, 1988; Huang and Liao, 1990). The presentation could be considered repetitive and unnecessary in a livestock or crops context, but not in aquaculture where the application of quantitative genetics lags decades behind the two former fields.

Individual or mass selection. The terms 'individual selection' and 'mass selection' are often used interchangeably, and they refer to selection solely based on the individual's phenotype. It has been a common strategy with fish because of its simplicity. It does not require individual identification or the maintenance of pedigree records, hence it may be considered the least costly method. In principle, it can produce rapid improvement if the heritability of the trait(s) under selection is high. Under those circumstances, however, there is risk of inbreeding due to inadvertent selection of progeny from few parents producing the best offspring, especially if progeny groups are large. For growth rate and morphological traits (easily assessed, expressed in both sexes) it can be quite suitable. By contrast, individual selection is not suitable for situations in which the estimation of breeding values requires slaughter of the animals (e.g. carcase and flesh quality traits) or challenge of some sort (e.g. selection for salinity tolerance or for disease resistance).

Hulata *et al.* (1986) carried out two generations of mass selection for growth rate with Nile Tilapia (*Oreochromis niloticus*) and observed no improvement over the original base population. They attributed the lack of response to selection to a number of possible factors, including inbreeding and genetic drift. They concluded that mass selection was not a promising method unless measures could be taken to control inbreeding. WorldFish (unpublished) records indicate that the experience with Silver Barb (*Barbonymus gonionotus*) in Bangladesh and Thailand and Common Carp (*Cyprinus carpio*) in Vietnam has been of satisfactory response to selection in early generations up to the fourth or fifth, declining sharply thereafter.

Overall, the evidence suggests that simple, unstructured, mass selection will result in problems unless the number of parents is large (Gjerde *et al.* 1996; Villanueva *et al.* 1996), and even so, chance could have a negative effect. Some form of structuring to control the parental contribution to the next generation appears necessary. If controlled pair matings can be carried out, the results of Bentsen and Olesen (2002) can be used to formulate the design of the breeding program. These authors investigated the effect of number of parents selected and of number of progeny tested per pair for a range of population sizes and heritability values. They show that inbreeding rates can be kept as low as one per cent per generation if a minimum of 50 pairs is mated and the number of progeny tested from each pair is standardized to 30 to 50 progeny. Note that although not requiring individual identification of the fish, the schemes suggested by these authors entail the conduct of pair matings, initial maintenance of the progeny of such pair matings in separate enclosures, and controlled contribution of each full sib family to the next generation at the time the fish are assigned to communal rearing. We have found that in some developing countries implementation of these practices was not possible and we had to change the strategy to that described in the following section.

Selection within cohorts and exchange of breeders. Eknath (1991) reports the genetic deterioration taking place in hatcheries in India due to poor brood stock management. To remedy the situation he suggested that brood stock could be arbitrarily divided into several groups. Mating could then be performed between individuals from the different groups on a rotational basis to avoid inbreeding. In this section we develop that notion further, based on the mating design used by McPhee *et al.* (2004) for weight selection in redclaw crayfish (*Cherax quadricarinatus*). These authors divided the population into cohorts, namely, groups sampled from a previously established

foundation population. A selection line was created, consisting of 20 cohorts, where each cohort had 15 female and 10 male foundation parents. A control line of eight cohorts of the same size was also established. One hundred individuals were measured per cohort. Offspring of cohorts were hatched and grown in separate pens within a pond. At harvest time individuals of the heaviest weight in each cohort were chosen as parents of the next generation in the selection line, whereas individuals of average weight were chosen in the control line. In either case, selection was based on the difference between the harvest weight of an individual and its cohort mean. This within cohort selection aims to eliminate the environmental effect of cohorts on growth differences among individuals. The same number of individuals was selected from each cohort. Animals selected in one cohort were mated with those selected in another one to avoid mating related animals. After four generations of selection harvest weight in the selection line was 1.25 times greater than in the control line.

Note that although with this mating scheme the exact number of parents contributing to the next generation is not known and the rate of inbreeding can be calculated for the worst case, that is, only one pair per cohort left offspring. By designing the selection program in such a way that even in the case that only one pair from each cohort produced progeny the inbreeding rate was not excessive, then we would be able to ensure that we would not run into problems due to inbreeding. With regards to the exchange of breeders between cohorts, this could be achieved by shifting the males born in one cohort to another one in a pattern as described by Nomura and Yonezawa (1996), following for instance Cockerham's cyclical mating system (Cockerham, 1970). In practice, we have found that, in contrast to single pair matings, selection within cohorts with exchange of breeders between cohorts following a prescribed pattern is a feasible design even with limited resources. Field personnel feel comfortable with it, and will thus rigorously adhere to the instructions provided.

Within family selection. The method requires identification of the families. This may be achieved by maintaining them in separate tanks, cages, hapas or any other means of containment, without necessarily tagging the fish. The criterion of selection is the deviation of each individual from the mean of the family to which it belongs. Within family selection is especially advantageous when there is a large component of environmental variance common to members of the same family. Full sib groups reared in unreplicated hapas or any other form of containment fall into this category (e.g. see estimates in Ponzoni *et al.*, 2005). Under such circumstances selection between families would be misleading from a genetic viewpoint because of the confounding between genetic merit and common environmental effects. The method can make very effective use of facilities. If replacements are chosen so that every family contributes the same number of individuals to the next generation (e.g. choose one female and one male from each family) the effective population size is twice the actual (Falconer and Mackay, 1996). However, not all the additive genetic variance is available for selection, but only a fraction equal to the coefficient of relationship among the family relatives in question (i.e. 0.5 and 0.25 for full and half sibs, respectively) will be available. If for a particular trait the heritability in the population is 0.2 and the families are full sib groups, then $h^2_w = 0.11$. The lower within family heritability can be compensated for by the high within family selection intensity that can be applied without increasing the rate of inbreeding. The selection intensity within families will be limited only by the number of individuals tested per family. The number of families involved in the program will determine the lower limit of inbreeding, which can easily be controlled by applying a rotational mating system such as that earlier suggested for selection within cohorts.

The use of within family selection was recommended for SE Asian countries by Uraiwan and Doyle (1986). It was successfully applied in the selection program that resulted in an improved

Tilapia strain developed in the Philippines by the Freshwater Aquaculture Center (FAC) of Central Luzon State University. The strain is known by a variety of names, FAC-selected, FaST, and IDRC strain (in recognition of the support received from the International Development Research Centre of Canada). The selection program and the strain's performance have been described by Abella *et al.* (1990), Camacho *et al.* (2001) and Bolivar and Newkirk (2002). The selection line started from a base population combining four strains of Tilapia, namely, Israel, Singapore, Taiwan and a 'FAC' strain available at the time. Nineteen full sib groups were established, and the basis of selection was body weight at 16 weeks of age. The heaviest male from a given family was mated to the two heaviest females of another family to avoid inbreeding. After 12 generations of selection the genetic gain in body weight has been estimated at 12.4 per cent per generation. Camacho *et al.* (2001) comment that within family selection was easy to manage, and that taking care of inbreeding by means of a rotational mating posed no difficulties. The method reduces the need for tagging large numbers of individuals. Note that Bolivar and Newkirk (2000) and Ridha (2004, 2006) compared FAC-selected with GIFT (Genetically Improved Farmed Tilapia), the latter strain resulting from a program combining individual and family selection (dealt with in a later section of this paper), and found that there were no differences in growth rate between both strains, but in some of the experiments (Ridha, 2004) GIFT had greater (23 per cent) survival rate. However, both strains were significantly more productive than other, unimproved, strains.

Combined selection. We use the term 'combined selection' in a broad sense, meaning selection that is based on individual information as well as on information coming from relatives (e.g. full and half sibs, progeny). In this case all of the additive genetic variance is available for selection and the use of information from relatives increases the accuracy of the estimation of breeding values. Furthermore, relatives' records can be used to estimate breeding values for traits that require slaughter of the animals (i.e. carcase and flesh quality traits) or that entail a risky challenge (i.e. disease resistance, tolerance to some environmental component). This is not possible with the other methods (e.g. mass selection or within family selection).

Three documented examples of the successful application of combined selection to the improvement of fish in developing countries will be cited here (in all cases growth rate was the main focus of selection): (i) The GIFT project in Philippines, which reported genetic gains of 12 to 17 per cent per generation in Nile Tilapia, over five generations (Eknath *et al.*, 1998); (ii) The Jayanti Rohu (*Labeo rohita*) selective breeding project in India, which reported a genetic gain of 17 per cent per generation over five generations (Reddy *et al.*, 1999; Mahapatra, 2005, personal communication); and (iii) The selection project of a Malawian indigenous Tilapia, *Oreochromis shiranus*, where the accumulated gain over two generations was 13 per cent (Maluwa, 2005). GIFT and Jayanti Rohu have been tested extensively on farm and proven to outperform other strains used by farmers. We earlier mentioned that the GIFT and FaST strains have very similar growth performance, but GIFT has shown greater survival rate, possibly due to the broader genetic basis in the population originally assembled and to the greater effective population size relative to FaST. Although the program with *O. shiranus* is at an earlier stage than the other two, the strain has now been tested extensively on farm with very positive results.

These three programs (GIFT, Jayanti Rohu, *O. shiranus*) have a number of features in common: (i) They all started with the assembly of a base population drawn from different sources in order to capture genetic variation; (ii) Controlled matings of identified females to identified males were conducted and complete pedigrees were maintained; (iii) Full sib groups were kept together until tagging; (iv) Approximately, 50 to 200 fish per full sib group were tagged and destined to communal rearing in a range of production environments in order to estimate genotype by environment interactions. In the case of GIFT and Jayanti Rohu a selection index combining individual, full sib and half sib information was used to rank individuals on genetic merit, whereas

BLUP breeding values were estimated in *O. shiranus*. BLUP procedures are also used in the selection of GIFT in the population that was transferred to Malaysia (Ponzoni *et al.*, 2005). The sound design coupled with rigorous conduct and analysis accounts for the gains achieved in these programs. Furthermore, data sets of this nature, developed over a number of generations, provide great research opportunities in the area of estimation of phenotypic and genetic parameters, as well as of environmental effects and genotype by environment interactions. As a by-product of the genetic improvement program, opportunities for local staff capacity building are created around it. If captured, these opportunities can result in the training of staff to a level that enables them to independently plan and conduct genetic improvement programs. Note that the amount of information that can be extracted from a pedigreed population is much greater than from a non-pedigreed one.

PROSPECTS FOR GENETIC IMPROVEMENT PROGRAMS

In this paper we focused on selection approaches that in our perception can be managed with the resources that are available in developing countries. A rigorous comparison of different methods based on published evidence is not possible. In practice the outcome of a program will be affected by many factors other than the selection method itself. When planning a new program in a developing country, a way of approaching the problem could be to begin thinking about and outlining the most complete one (i.e. full pedigrees and BLUP estimates), and to simplify it gradually until it becomes feasible with the available resources, working backwards through the methods we presented. The final decision before implementation will be a matter of judgment. In any case, starting with a population with ample genetic variation is a trademark of successful fish genetic improvement programs. Although this in itself is not a sufficient condition for success, it is indeed a necessary condition. Sophisticated designs and genetic evaluation procedures are no substitute or remedy for a genetically deteriorated base population. The failure of some attempts to achieve genetic improvement with aquatic animals may have been due more to weaknesses in the base population than to the selection method utilized. Irrespective of the method of choice, continued genetic improvement will hinge upon the adequate balance between high selection intensity and the maintenance of low inbreeding rate.

From an economic viewpoint, investment appraisal studies indicate very favourable benefit cost ratios for genetic improvement programs for both Nile tilapia (Ponzoni *et al.* 2007) and for common carp (Ponzoni *et al.* 2008). This was shown to be so even for situations in which there was genotype by environment interaction (Ponzoni *et al.* 2008) and a single program had to service more than one environment. It is reasonable to think that these results can be generalised to other, similar, aquatic animal species.

It is sometimes suggested that, given the long term implications of genetic improvement programs, the stock to be improved should be part of an already viable aquaculture sector. This argument deposits little faith in the power of genetic improvement which can turn a non viable aquaculture sector into a prosperous one (e.g. due to 100% superiority of improved stock over farmers' strains). What viability and prosperity can one expect in an aquaculture sector based on stock that grows about 40 per cent less that their wild counterparts (Brummett *et al.* 2004)? Commercially viable farms cultivating crops or running livestock are not based on plants or animals that are significantly less productive than their wild relatives.

LIMITATIONS AND CHALLENGES TO AQUACULTURE DEVELOPMENT

Limitations. Lack of resources is at the forefront in this area. The limitations are often both financial and human. Technical staff involved in aquaculture most often come from a fish biology background and have limited (if any) training in quantitative genetics. The limitations in capacity

are more important than the financial limitations because well trained personnel can make important achievements with limited material resources, but there is no amount of resources that will compensate for lack of capacity. Hence, training of local staff in quantitative genetics and animal breeding should be a priority.

When a genetically improved strain is available multiplication and dissemination face difficulties due to lack of capacity at the hatchery level. The notion of paying a greater price for stock of greater genetic merit is not frequently part of the culture. In turn, producers most often lack both the understanding and the financial means to acquire fingerlings from hatcheries breeding an improved strain. The sad reality is that we have been more successful at developing improved strains than at achieving impact with them at the producer level.

Conservationists and biodiversity groups show great concern about the risks involved in using genetically improved strains. Aquatic animals are more difficult to contain than their terrestrial counterparts and escapees may interbreed with wild stock thus eroding or changing genetic variation in those populations. The sentiments guiding conservationists and biodiversity groups are legitimate and noble, but there is sometimes an incomplete understanding of the issues and the genetic risks are magnified relative to others. For instance, the threats to wild populations from pollution and human encroaching may be greater than those from possible escapes of fish from an improved strain, especially if dealing with a non carnivore such as Nile tilapia.

Challenges. The principles used in the genetic improvement of crops and livestock are applicable to aquatic animals but in practice there are many issues that are peculiar to this latter group.

In order to use combined selection, identifiable families have to be produced. The progeny of the different families must be marked so that they can be communally stocked and tested for genetic evaluation purposes. Aquatic animals are generally very small at spawning. They are kept in their family (usually full sib) groups until they are large enough to be tagged. This often results in an appreciable common environmental effect in traits such as growth rate. The most commonly used tags with fish are Floy Tags ® and PIT (passive integrated transponder) tags, the latter being about five times more expensive than the former, but far superior in terms of retention rate. In the hands of unskilled staff the combined effects of poor reproductive rates, large common environmental effects, and high tag losses can negate the theoretical virtues of elaborate selection methods. DNA technology is available that enables ascertaining parentage after mass spawning and communal rearing of fry prior to tagging. Communal rearing soon after spawning virtually eliminates the common environmental effect (Ninh 2009). It also enables greater growth rate than rearing in hapas or tanks in full sib groups, it may help reduce the generation interval and increase selection response. The challenge consists of reducing the cost of the DNA technology to the point that it can be applied in practical genetic improvement programs.

There are issues related to animal behaviour, genotype by environment interaction, variability at harvest and animal welfare that can be dealt with newly developed (and developing) quantitative genetic theory. These constitute a fertile ground for research and development with aquatic animals. With the rapid expansion of the aquaculture industry (FAO 2009), we can foresee that fish welfare will become one of the major challenges in fish breeding programs. Welfare can be linked to product quality and quantity and to production efficiency (Ashley 2007). There is no universal definition of fish welfare, but it is commonly defined as representing the physical and mental state of well-being of the animal in relation to its internal and external environment (Ellis *et al.* 2002). Farmed fish are often subject to stress during crowding, handling, transport and controlled reproduction (Kubilay and Uluköy 2002). In developing countries, fish selection programs have been mainly focused on growth rate. Such a breeding objective favours the selection of more aggressive animals (Lopez 1996). Social interactions among individuals can have profound influences on the expression of performance and welfare traits (Muir, 1996;

Brichette *et al.* 2001; Denison *et al.* 2003; Muir, 2005; Bijma *et al.* 2007). They can reduce growth due to competition for limited resources and result in mortality due to cannibalism. The stress effects from high stocking density under intensive culture systems (Ellis *et al.* 2002) coupled with such behavioural change may act against the welfare of genetically improved farmed fish. A strong case can be made for selective breeding programs to include behavioural traits that may lead to reduced aggression, greater uniformity in harvest weight and that are related to fish health and welfare.

In developing countries, virtually always, at best, a single genetically improved line is developed for each species of interest. This poses two obvious problems: (i) Sustainability of the line in relation to effective population size and risk of loss in case of a disaster, and (ii) Requirement to service a range of production environments. The first problem can be easily handled with adequate resources. The second one calls for a breeding strategy that goes beyond common animal breeding practice. The term phenotypic 'plasticity' is used more often in aquatic animal literature than in livestock. Phenotypic plasticity is the property of organisms of a genotype to develop systematically different phenotypes in different environments. Genotypes showing highly variable phenotypes across environments are 'plastic', whereas genotypes that show little variability are 'robust' (de Jong and Bijma, 2002). The concept of robustness has been advocated as a breeding goal (Ellen *et al.* 2008) because it would have a positive effect on welfare and health without loss of integrity. Robust animals are organisms that have the ability to maintain their performance across different environmental conditions (Waddington 1960). Selection for robustness in fish breeding programs would result in animals that have a greater ability to maintain their homeostatic state when exposed to stressors, thus improving fish welfare (Ellen *et al.* 2008). James (2009) discusses this issue in the context of genotype by environment interactions. When a genetic improvement program has to service several environments selection index theory could be applied, using the relative importance of the environments as weighting factors. Because the breeding goal should target the future circumstances the weighting factors should reflect the anticipated future relative importance, rather than the present one. Given the high reproductive rate of aquatic animals and the relatively modest resources (compared to livestock) that would be required for testing in multiple environments it appears that research in this area would be worthwhile. To our knowledge there have been no genetic improvement programs with aquatic animals formally targeting multiple environments.

There are other challenges such as feed utilization efficiency, flesh quality, fitness related issues, and disease outbreaks due to intensive aquaculture that are not dealt in the present review.

CONCLUDING REMARKS

Genetically improved strains are essential to aquaculture development. The application of proven quantitative genetic theory should continue for relevant species. There is ample proof of the success such programs can have. However, greater emphasis is required in the areas of dissemination of the improved stock to farmers in order to ensure impact at the production level. New developments in quantitative genetic theory will help prepare the aquaculture industry for the future. Broadening of the breeding objective to include behavioural traits and selection for multiple environments are areas where much fruitful research and development could be conducted.

REFERENCES

Ashley, P.J. (2007) *Applied Animal Behavior Science* **104**: 199.
Abella, T.A., Palada, M.S. and Newkirk, G.F. (1990) *Proc. 2nd Asian Fisheries Forum*: 515.
Bentsen, H.B. and Olesen, I. (2002) *Aquaculture* **204**: 349.

Bjima, P., Muir, W.M., Ellen, E.D., Wolf, J.D. and van Arendonk, J.A.M. (2007). *Genetics.* **175:** 289.
Bolivar, R.B. and Newkirk, G.F. (2002) *Aquaculture* **204**: 371.
Brichette, I., Reyero, M. I. and Garcı´a, C. (2001). *Aquaculture* **192**: 155.
Camacho, A.S., Abella, T. and Tayamen, M.M. (2001) *Fish Genetics, ICLARM Conf. Proc.* **64**: 71.
Brummett, R.E., Angoni, D.E. and Pouomogue, V. (2004). *Aquaculture* **242**: 157.
Cockerham, C.C. (1970) In "Mathematical Topics in Population Genetics", p. 104, editor K. Kojima, Springer Verlag, New York.
De Jong, G and Bijma, P. (2002). *Livest. Prod. Sci.* **78**: 195.
Denison, R. F., Kiers, E. T. and West. S. A. (2003). *Q. Rev. Biol.* **78**: 145.
Eknath, A.E. (1991) *NAGA, The ICLARM Quarterly* **738**: 13.
Eknath, A.E., Dey, M.M., Rye, M., Gjerde, B., Abella, T.A., Sevilleja, R., Tayamen, M.M., Reyes, R.A. and Bentsen, H.B. (1998) *Proc. 6th WCGALP* **27**: 89-96.
Ellen, E.D., Star, L., Uitdehaag, K.A. and Brom, F.W.A. (2008) *J. of Agricultural and Environmental Ethics* **21**: 109
Ellis, T., North, B., Scott, A.P., Bromage, N.R., Porter, M. and Gadd, D. (2002) *J. of Fish Biology* **61**: 493
Falconer, D.S. and Mackay, T.F.C. (1996) "Quantitative Genetics" Longman Group Ltd., Harlow, UK.
FAO (2009). "The State of World Fisheries and Aquaculture 2008". Food and Agriculture Organisation, Rome.
Gjerde, B., Gjoen, H.M. and Villanueva, B. (1996) *Livest. Prod. Sci.* **47**: 59.
Huang, C.M. and Liao, I.C. (1990) *Aquaculture* **85** : 199.
Hulata, G., Wohlfarth, G.W. and Halevy, A. (1986) *Aquaculture* **57**: 177.
James, J. (2009). In "Adaption and fitness in animal populations", p. 151, editors van der Werf, Graser, H.-U., Frankham, R., and Gondro, C., Springer Netherlands.
Kubilay A. and Uluköy G. (2002) *Turk. J. Zool* **26**: 249.
Lopez, J.M.A. (1996). *Informacion Tecnica Eco. Agraria*. **92A**: 177.
Mahapatra, K.D. (2005) Personal Communication.
Maluwa, A.O.H. (2005) PhD Thesis, Norwegian University of Life Sciences, Norway.
McPhee, C.P., Jones, C.M. and Shanks, S.A. (2004) *Aquaculture* **237**: 131.
Muir, W.M. (1996). *Poult. Sci*. **75**: 447.
Muir, W.M. (2005). *Genetics*. **170**: 1247.
Ninh, N.H. (2009). PhD thesis. University of Stirling, UK.
Nomura, T. and Yonezawa, K. (1996) *Genet. Sel. Evol*. **28**: 141.
Ponzoni, R.W., Hamzah, A., Tan, S., Kamaruzzaman, N. (2005) *Aquaculture* **247**: 203.
Ponzoni, R.W., Nguyen, H.N. and Khaw, H.L. (2006) *Proc. 7th WCGALP*: 09-02 [CD-ROM]
Ponzoni, R.W., Nguyen, H.N. and Khaw, H.L. (2007). *Aquaculture* **269** : 187.
Ponzoni, R.W., Nguyen, H.N., Khaw, H.L. and Ninh, N.H. (2008). *Aquaculture* **285** : 47.
Reddy, P.V.G.K., Gjerde, B., Mahapatra, K.D., Jana, R.K., Saha, J.N., Rye, M. and Meher, P.K. (1999) "Selective breeding Procedures for Asian Carps" CIFA, Bhubaneswar, India.
Ridha, M.T. (2004) *Proc. 6th ISTA* **I**: 60.
Rhida, M.T. (2006) *Aquaculture Research* **37**: 172.
Teichert-Coddington, D.R. and Smitherman, R.O. (1988) *Transactions American Fisheries Society* **117** : 297.
Uraiwan, S. and Doyle, R.W. (1986) *Aquaculture* **57**: 93.
Waddington, C. H. (1960) *Genet. Res*. **1**: 140.
Villanueva, B., Woolliams, J.A. and Gjerde, B. (1996) *Anim. Sci.* **63**: 563.

DEVELOPMENT OF A BREEDING STRATEGY FOR HYBRID ABALONE

M. G. Hamilton[1], P. D. Kube[1], N. G. Elliott[1], L. J. McPherson[2] and A. Krsinich[2]

1. CSIRO Food Futures Flagship, Castray Esplanade, Hobart, TAS 7001
2. Great Southern Waters Pty Ltd, 366 The Esplanade, Indented Head, VIC 3223

SUMMARY

The Australian temperate abalone fishery is based on two main species, the greenlip and the blacklip abalone. Sectors of the emerging abalone aquaculture industry have a commercial preference for the hybrid between these two species and wish to genetically improve the hybrid. To decide on an appropriate hybrid breeding strategy, Great Southern Waters and CSIRO, have initiated a trial over four year classes designed to determine if pure-species breeding values are correlated with hybrid breeding values. Discouragingly, early results from the first year class show a relatively weak and non-significant inter-species family correlation at 21 months of age; suggesting that recurrent selection for weight at 21 months in a blacklip breeding population may not result in the improvement of F1 hybrid growth. However, this weak and non-significant correlation may have been due to a number of extraneous factors and further studies are planned to investigate these.

INTRODUCTION

Abalone aquaculture is a relatively new but expanding industry in Australia. Blacklip (*Haliotis rubra* (Leach)) and greenlip (*H. laevigata* (Donovan)) abalone, and their interspecific hybrid are the principal species farmed. These species and their hybrid are in the early stages of domestication and, accordingly, there is only a limited understanding of their genetic architecture. Sectors of the industry have a commercial preference for the hybrid as it exhibits favourable growth and behavioural characteristics in culture, as well as desirable product qualities.

Great Southern Waters Pty Ltd (GSW). GSW produces farmed blacklip × greenlip hybrid abalone. The company initiated a family-based selective breeding programme in 2003 with the objective of increasing growth and meat yield while maintaining the distinctive aesthetic characteristics and meat quality of its 'Jade Tiger™' product. To date, the company has focused its breeding research and development efforts on (i) establishing a closed, diverse and healthy breeding population, (ii) establishing trials to verify the advantages of hybrid abalone under GSW's culture system, (iii) developing appropriate data capture and management systems and (iv) establishing progeny trials to estimate genetic parameters required to design, refine and implement a hybrid abalone breeding strategy.

Benefits of hybrids. Hybridisation has been used as a tool for improving the growth, behavioural, flavour, reproductive and processing characteristics of many commercially important plant and animal species (Kerr *et al.* 2004). Possible explanations for hybrid superiority over pure species, include: (i) heterosis; (ii) complementarity (e.g. where the intermediate expression of traits in the hybrid is superior, from an economic or management perspective, to the extreme expression of these traits in the ancestral species); and (iii) exploitation of greater allelic diversity through recombination and selection within backcrossed or advanced-generation (i.e. composite) hybrids (Falconer and Mackay 1996, Potts and Dungey 2004).

Abalone reproductive biology. Although unilateral cross-incompatibility is not complete, the use of blacklip dams and greenlip sires is more successful than the reciprocal cross. Female blacklip

abalone take approximately three years to reach sexual maturity and males can be mature at two years. Most animals live approximately five years in culture. Abalone are highly fecund (0.1 to 8.0 x 10^6 eggs per female depending on body size), but one of the principal difficulties experienced by abalone breeders is the synchronised induction of spawning (Elliott 2000). However, these problems have to some extent been overcome through the use of domesticated broodstock and the adoption of appropriate spawning methods and infrastructure.

For breeding purposes animals are maintained in individual family tanks until they reach a size suitable for physical tagging (about 15 mm). Difficulties with spawning, limited infrastructure, and limited skilled personnel restrict the number of individual families that can practically be produced in any one spawning season to a maximum of 40 to 50.

Hybrid breeding strategies under consideration. Three hybrid breeding strategies are being considered: (i) pure species selection (ii) reciprocal recurrent selection with forward selection and (iii) composite hybrid breeding (Kerr *et al.* 2004). To decide on a strategy, GSW and CSIRO have initiated a trial over four year classes, designed to determine if pure-species breeding values are correlated with hybrid breeding values (i.e. to establish to what extent recurrent selection in pure species could be used to genetically improve deployed F1 hybrids). This paper outlines the preliminary findings from the first year class which was spawned in the 2006/07 season.

MATERIALS AND METHODS

In this study, the 'half-sib' hybrid and pure-species progeny of 21 wild-caught blacklip dams were compared. Each of these dams was crossed with one of seven groups of blacklip sires and one of seven groups of greenlip sires. These crosses were made over 3 spawning runs (Table 1).

Pair-wise fertilisations were undertaken separately and, after fertilisation, eggs from a single dam were mixed to create a pure species and a hybrid maternal family from each dam. Animals were then maintained in single-species (blacklip and hybrid) family tanks for approximately 10 months before 540 individuals from each family were tagged. Total live weight from a random sample of 130 animals per family were taken at tagging and, at approximately 21 months of age, an independent random sample of 60 animals per family were weighed in the same manner. Not all animals weighed at 21 months were weighed at 10 months. Preliminary univariate restricted maximum likelihood (REML) mixed-model analyses, for each species and measurement age, were undertaken by fitting the following model in ASReml (Gilmour *et al.* 2006):

$$\text{WEIGHT} = \text{MEAN} + \text{SPAWN\_RUN} + \textit{SIRE\_GROUP} + \textit{FAMILY} + \textit{RESIDUAL}$$

where WEIGHT is a vector of log-transformed weight observations, MEAN is the mean weight, SPAWN_RUN are the spawn run (i.e. spawning day confounded with the ratio of wild and farm males in greenlip sire groups; Table 1) effects fitted as a fixed factor, *SIRE_GROUP* are the sire group effects fitted as a random factor, *FAMILY* are the maternal family within sire group effects fitted as a random factor and *RESIDUAL* is a vector of residuals. A mulitivariate (two species by two measurement ages) model, which extended the univariate model and allowed for covariation between family effects and within-species residuals, was then fitted to estimate variance components and correlations. One- and two-tailed likelihood ratio tests were used to test the significance from zero of family variances and correlations, respectively (Gilmour *et al.* 2006).

For each species and assessment age, the narrow-sense heritability (h^2) and coefficient of additive genetic variation (CV_a) were estimated assuming family variance was explained entirely by additive genetic variation among dams (Falconer and Mackay 1996):

$$h^2 = \frac{4 \times \sigma^2_{\text{fam}}}{(\sigma^2_{\text{fam}} + \sigma^2_{\text{e}})} \qquad CV_{\text{a}} = \frac{\sqrt{4 \times \sigma^2_{\text{fam}}}}{\bar{x}}$$

where σ^2_{fam} is the family within sire group variance; σ^2_{e} is the residual variance and $\bar{x}$ is the log-transformed trait mean.

Table 1. Crossing design for the 2006-07 year class (grey = farm sire, black = wild sire).

Blacklip dam	Spawning day	Blacklip sire group ID	Greenlip sire group ID	Blacklip sire groups (four to nine individuals per group)	Greenlip sires groups (four individuals per group)
A	0	1	8		
B	0	1	8		
C	0	1	8		
D	0	1	8		
E	0	2	9		
F	0	2	9		
J	0	3	10		
K	0	3	10		
L	12	4	11		
M	12	4	11		
N	12	4	11		
O	12	4	11		
P	12	5	12		
Q	12	5	12		
R	12	5	12		
S	22	6	13		
T	22	6	13		
U	22	6	13		
V	22	7	14		
W	22	7	14		
X	22	7	14		

RESULTS AND DISCUSSION

Hybrid progeny grew more rapidly than blacklip progeny (Table 2). First-generation hybrids are typically found to give superior performance to pure species in the farmed situation. However, in this study the different genetic background of blacklip and greenlip sires may also have been a factor, since all blacklip sires were wild-caught and most greenlip sires were domesticated farm stock selected according to their size (Table 1).

Table 2. Back-transformed least squares (LS) means, narrow sense heritabilities (h^2) and coefficients of additive genetic variation (CV_a) for weight by species and assessment age.

Species	Time of assessment	LS mean (95% CI) (g)	h^2 (SE)	CV_a (%)
Blacklip	10 months (tagging)	1.55 (1.44 - 1.67)	0.97 (0.25)***	22.5
Blacklip	21 months	12.64 (11.89 - 13.44)	0.20 (0.09)***	7.2
Hybrid	10 months (tagging)	1.75 (1.57 - 1.93)	1.29 (0.30)***	29.7
Hybrid	21 months	17.08 (15.03 - 19.41)	0.28 (0.12)***	7.7

*** P<0.001

The extreme heritabilities and high coefficients of additive genetic variation reported for 10

months of age indicated that family variation observed at this early measure was not entirely explained by additive genetic variation but was, in part at least, explained by nursery effects (e.g. maternal, fertilisation, early-husbandry and tank effects).

The inter-age hybrid family correlation was strong (Table 3), suggesting that relative family performance did not change substantially between 10 and 21 months of age, possibly due to the persistence of nursery effects. However, the inter-age blacklip family correlation was weak (albeit positive) and not significantly different from zero. This weak and non-significant inter-age correlation could have been due to genotype-by-age interaction in growth, but was also consistent with a reduction in the influence of nursery effects over time (Kube *et al.* 2007).

Table 3. Family correlations for weight between assessment ages within species and between species within assessment ages.

Comparison	Category	Additive genetic correlation (SE)
Inter-age	Blacklip	0.37 (0.25) ^{ns}
Inter-age	Hybrid	0.72 (0.15)**
Inter-species	10 months (tagging)	0.74 (0.11)***
Inter-species	21 months	0.38 (0.29) ^{ns}

^{ns} not significant, * P<0.05, ** P<0.01, *** P<0.001

The strong and significant inter-species family correlation observed at 10 months, probably reflected nursery effects (excluding tank effects, as they were independent across species) more than additive genetic effects. Discouragingly, the 21-month inter-species family correlation, which is likely to be a more accurate estimate of the additive genetic correlation, was only weakly positive and non-significant. This suggests that recurrent selection for weight within the blacklip population may not result in the genetic improvement of F1 hybrid growth. However, this weak and non-significant correlation may have been due to nursery effects persisting beyond 10 months of age (particularly in the hybrid); biases among families due to variation in the survival of paired crosses after egg mixing; or a lack of statistical power (i.e. Type II error), given the relatively low number of families studied. Further studies are planned to investigate these possible explanations using data from this year class and others but, until these studies are completed, pure-species recurrent selection cannot be pursued with confidence.

Future studies. A sample of animals from each half-sib family will be genotyped to enable pedigree reconstruction, the precise estimation of genetic parameters and the examination of non-additive genetic effects. Assessment of later-age weight and other economically important traits will also be undertaken. Additionally, studies on advanced-generation hybrids and backcrosses will be carried out to determine the suitability and optimal composition of a composite breed.

REFERENECES

Elliott, N.G. (2000) *Aquac Res* **31**:51-59.

Falconer, D.S. and Mackay, T.F.C. (1996) "Introduction to quantitative genetics" Longman, Harlow.

Gilmour, A.R., Gogel, B.J., Cullis, B.R. and Thompson, R. (2006) "ASReml User Guide Release 2.0" VSN International Ltd, Hemel Hempstead, UK.

Kerr, R.J., Dieters, M.J. and Tier, B. (2004) *Can J For Res* **34**:209.

Kube, P.D., Appleyard, S.A. and Elliott, N.G. (2007) *J Shellfish Res* **26**:821.

Potts, B.M. and Dungey, H.S. (2004) *New For* **27**:115.

Aquaculture

ASSESSMENT OF THE LEVEL OF HETEROZYGOSITY IN THE TASMANIAN ATLANTIC SALMON (*SALMO SALAR*) POPULATION USING SINGLE NUCLEOTIDE POLYMORPHISM MARKERS

S.Dominik[1, 2], J.M. Henshall[1,2], P.D. Kube[1,3], H. King[4], S. Lien[5], M. Kent[5] and N.G. Elliott[1,3]

[1]CSIRO Food Futures Flagship
[2]CSIRO Livestock Industries, Armidale, NSW 2350
[3]CSIRO Marine and Atmospheric Research, Hobart, TAS 7000
[4]Salmon Enterprises of Tasmania Pty. Ltd., Wayatinah, TAS 7140
[5]CIGENE, Norwegian University of Life Sciences, Aas, Norway

SUMMARY

The level of heterozygosity in the Tasmanian Atlantic salmon population was investigated using SNP data from 93 fish from unselected broodstock and 2991 single nucleotide polymorphisms (SNPs). Three approaches were used as measures of the likelihood of two alleles being identical by descent: two homozygosity indices, which have not been used on SNP data before and an inbreeding coefficient for the SNP markers based on observed and expected heterozygosity. As expected, the two homozygosity indices yielded highly correlated results, because biallelic SNP markers are all equally informative. The observed and expected heterozygosity at the SNP loci were very similar and that was reflected in intermediate inbreeding coefficients. All three analysis methods indicated a moderate likelihood that alleles are identical by descent, and offspring tested came from unrelated parents. Most markers were in Hardy-Weinberg-equilibrium. The SNP data support previous microsatellite studies concluding that the Tasmanian Atlantic salmon population has a moderate level of genetic diversity, as indicated by homozygosity indices and the inbreeding coefficient, and that hatchery protocols over the 40 years since the population was established in Australia have maintained a genetically healthy population.

INTRODUCTION

Salmon Enterprises of Tasmania Pty. Ltd. (SALTAS) embarked on a selective breeding program (SBP) in 2004 (Elliott and Kube 2009). One of the main factors determining potential genetic gain is the level of relatedness present in the population. Genetic diversity in the Tasmanian Atlantic salmon population was previously investigated based on sparse microsatellite markers (Innes and Elliott 2006). That study concluded that despite loss of alleles, heterozygosity had been maintained in the Tasmanian population when compared to its progenitor population. It was further concluded that sufficient genetic diversity was present to support a selective breeding program, but regular monitoring was recommended to detect any decline genetic diversity. Stochastic simulation studies by Henshall and Dominik (unpublished) modelled the history of the population and predicted the average rate of inbreeding to be below 1% per generation.

The aim of this study was to assess the level of heterozygosity in offspring of unselected broodstock of the founder generations of the SALTAS selective breeding program based on single nucleotide polymorphisms (SNP). Three approaches including two measures of homozygosity, which have not previously been applied to SNP data, are used.

MATERIALS AND METHODS

The Tasmanian Atlantic salmon population. The history of the Australian Atlantic salmon population is described in detail in Reilly *et al.* (1999). Two key events led to the establishment of the Tasmanian Atlantic salmon population. The first event was the importation of around 400,000

ova from Nova Scotia, Canada into Gaden, New South Wales (NSW), Australia between 1963 and 1968. The second event occurred between 1984 and 1986 when 570,000 ova were brought from NSW into Tasmania leading to the establishment of the SALTAS hatchery. These imports formed the foundation of Australia's leading aquaculture industry based in Tasmania. Following those events the population has been closed to any further importations and SALTAS has been producing broodstock and smolt for the industry over the last 20 years.

Sample preparation and genotyping. DNA was extracted from fin-clippings of 95 offspring across three year classes of founder families (2004 – 2006) from the SALTAS selective breeding program. Fish were chosen to be unrelated with the exception of the inclusion of the parents of one of the individuals. DNA was diluted in TE (10mM Tris pH 7.5, 1mM EDTA). A minimum of 20ul per sample at 50ng/ul (total = 1ug DNA) was prepared to facilitate pipetting.
The 95 samples plus one control sample were genotyped using a 15,225 SNP iSelect Atlantic salmon chip developed by Illumina (www.illumina.com) for the Centre for Integrative Genetics (CIGENE; www.cigene.no) as described by Kent *et al.* (2009). BeadStudio software (Illumina, 2007) was used to generate the allelic data, based on identifiable clusters of the normalised intensity data. SNPs were classed into polymorphic and monomorphic SNPs in unique and duplicated genomic regions. Only SNPs from unique regions were used for further analysis.

Measures of identity by descent. Two approaches that provide a measure of the likelihood of alleles to be identical by descent at the individual level were used.

Internal relatedness Index. The measure of internal relatedness (IR) was developed by Amos *et al.* (2001). It accounts for differences in allelic frequencies. IR compares pairs of loci based on their number of homozygotes and frequencies of alleles. IR ranges from IR = -1 for complete heterozygosity to IR = 1 for complete homozygosity of an individual across all loci.

$$IR = \frac{(2H - \sum f_i)}{(2N - \sum f_i)}$$

with H = number of homozygote loci
N = the total number of loci
f_i = frequency of the ith allele contained in the genotype

Homozygosity by loci Index. Aparicio *et al.* (2006) used the homozygosity by loci (HL) index to improve on the Amos *et al* (2001) approach and to reflect differences in the level of information of loci; the IR was suggested to underestimate heterozygous individuals carrying rare alleles (a possible issue with SNP data). HL = 0 indicates heterozygosity across all loci and HL = 1 reflects complete homozygosity.

$$HL = \frac{\sum E_h}{\sum E_h + \sum E_j}$$

with E_h = expected heterozygosity of loci that an individual carries in homozgyosity
E_j = expected heterozygosity of loci that an individual carries in heterozygosity

Coefficient of inbreeding. The simplest parameter of Wright's F-statistics (Wright 1951) was used. The inbreeding coefficient (F) describes the probability that two alleles at a locus taken at

random in a population are identical by descent. F can range between -1, when there is no inbreeding, to 1 if all alleles are identical by descent.

$$F = \frac{(H_e - H_o)}{H_e}$$

with H_e = expected heterozygosity based on allele frequencies
H_o = observed heterozygosity

RESULTS AND DISCUSSION

Of the 15,225 SNPs contained on the Atlantic salmon SNP chip, 2991 markers were found to be both polymorphic and present in unique genomic regions in our samples. Since the SNP chip was developed using data from the Norwegian breeding population, it was expected that only a proportion of the markers would be polymorphic in the Tasmanian Atlantic salmon population. The genotyping results showed very few missing values with a mean of 92.7 samples per SNP and on average 2982.5 SNPs per sample genotyped.

Table 1. Summary statistics of the homozygosity measures (IR, HL), for the observed and expected heterozygosity (H_o and H_e), the allele frequency for one of the alleles of each SNP (Allelefreq) and the inbreeding coefficient of the loci (F) for 93 samples of Atlantic salmon.

	Mean	Maximum	Minimum	Standard deviation
IR	-0.01	0.15	-0.20	0.05
HL	0.64	0.69	0.54	0.02
H_o	0.25	1.00	0.00	0.18
H_e	0.25	0.50	0.01	0.17
Allelefreq	0.46	0.99	0.01	0.35
F	-0.01	1.00	-1.00	0.15

Alleles cover the possible range of frequencies with some SNPs having rare alleles, which are reflected in the maximum and minimum values of the Allelefreq (Table 1). 87.5% of the polymorphic SNPs were in Hardy-Weinberg-Equilibrium (HWE) as assessed by a Chi-square test ($P< 0.05$), which is confirmed in Figure 1. It shows that the relationships between genotype and gene frequencies are close to what is expected when markers are in HWE.

The mean inbreeding coefficient is close to zero, which is a result of no major discrepancies in the observed (H_o) and expected (H_e) heterozygosities. Four loci had F = -1, which reflects no inbreeding and allele frequencies of both alleles of 0.5. Thirteen SNPs resulted in F = 1. In these loci, no heterozygote genotypes were observed.

IR and HL are distributed in the intermediate range of possible values, indicating a moderate likelihood of alleles being identical by descent, which indicates moderate genetic diversity in the original broodstock. As expected both measures were highly correlated with $r^2 = 0.97$, due to the biallelic nature of the markers.

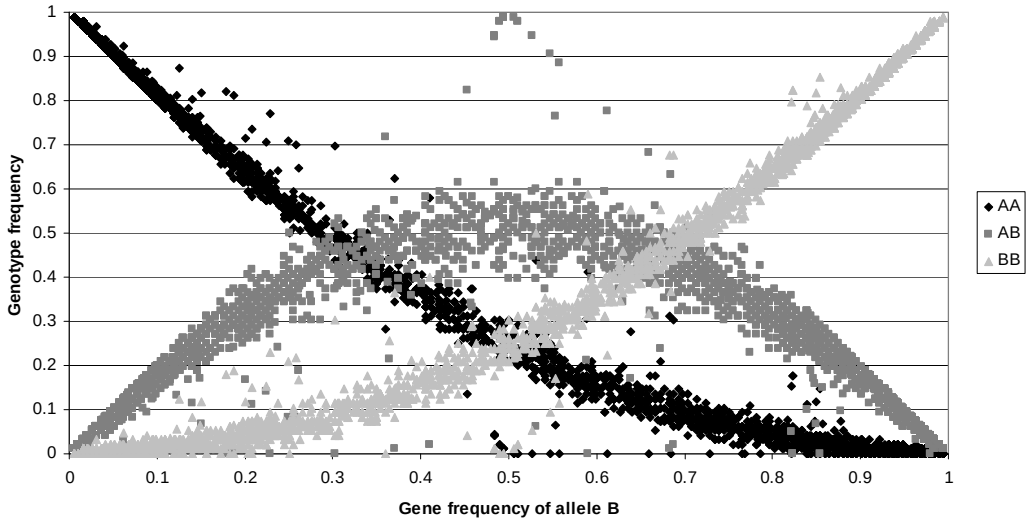


Figure 1. Relationship between genotype frequencies and gene frequencies for the two SNP alleles (A and B) in 93 Tasmanian Atlantic salmon samples.

CONCLUSION

The results of this study using high density biallelic SNPs support those obtained with sparser highly polymorphic microsatellites. Despite anecdotal evidence of small effective population size in the early years of the domestication of Atlantic salmon in Australia, both demonstrate maintenance of heterozygosity and indicate a healthy level of genetic diversity in Tasmanian Atlantic salmon population at the start of its selective breeding program. The genetic diversity found across 2991 SNP markers suggests sufficient diversity for genetic progress in the SALTAS selective breeding program. SNP technology can provide a useful tool to monitor the level of heterozygosity, as suggested by Reilly *et al.* (1999), and should be used alongside other diversity measures to avoid a decline in genetic diversity in the population.

ACKNOWLEDGEMENT

This research was supported by the Fisheries Research and Development Cooperation.

REFERENCES

Amos, W., Worthington Wilmer, J., Fullard, K., Burg, T. M., Croxall, J. P., Bloch, D. & Coulson, T. (2001). *Proc. Roy. Soc. Lond. B* **268**, 2021-2027.

Aparicio, J. M., Ortego, J. & Cordero, P. J. (2006) *Mol. Ecol.* **15**, 4659-4665.

Elliott, N.G. and Kube, P.D (2009) *Proc. Assoc. Advmt. Anim. Breed. Genet.* **18:**362.

Innes, B.H. and Elliott N.G. (2006) *Aquaculture Research* **37**, 563 – 569.

Illumina (2007) Technical Note: Illumina DNA analysis. Publication Number 970-2007-005 06 August 2007.

Kent, M.P., Hayes, B., Xiang, Q., Berg, P.R., Gibbs, R.A., Lien, S. (2009) *Proc. PAG XVII*, San Diego, CA.

Reilly, A., Elliott, N.G., Grewe, P.M, Clabby, C., Powell, R., Ward, R.D. (1999) *Aquaculture* **173**: 459.

Wright, S. (1951). *Ann. Eugen. Lond.* **15**: 323.

Aquaculture

HETEROSIS, DIRECT AND MATERNAL GENETIC EFFECTS ON BODY TRAITS IN A COMPLETE DIALLEL CROSS INVOLVING FOUR STRAINS OF RED TILAPIA *OREOCHROMIS SPP*

N.H. Nguyen[1], N. Pongthana[2] and R.W. Ponzoni[1]

[1]The WorldFish Center, Penang, Malaysia
[2]Pathumthani Fisheries Test and Research Center, Aquatic Animal Genetics Research and Development Institute, 39 Klongha, Klongluang, Pathumthani 12120 Thailand

SUMMARY

Heterosis, direct additive genetic and general reciprocal effects were estimated from a complete diallel cross involving four strains of red tilapia *Oreochromis spp* from Malaysia, Stirling, Taiwan and Thailand. The mating involved 16 parental female and male breeders per strain, producing 64 full sib families in total, with four full-sib families per cross. Statistical analyses were carried out on 1280 performance records collected in both fresh water (0 ppt) and saline water (30 ppt) environments. There was a large additive genetic component for body traits in the four strains of red tilapia. The Malaysian strain was the best (7.4% above the overall mean of the pure strains), whereas the strain from Stirling had the poorest additive performance (13.4% below the overall mean of the pure strains). The average heterosis for body weight across the testing environments was low (4.2%) and the average of all crossbreds was not statistically different from the mean of pure strains. Ranking of strains based on estimates of reciprocal effects was generally similar to that of additive genetic effects. Strategies for the future breeding program in red tilapia are discussed.

INTRODUCTION

Information on phenotypic and genetic parameters for red tilapia (*Oreochromis spp*.) is extremely limited. Despite the importance of the species in many Asian countries and the existence of a number of distinct populations (referred to as 'strains' here) there are no published results of strain comparisons or of purebreeding or crossing parameters. In order to gain some understanding of the genetic characteristics of red tilapia, we sampled strains from Malaysia, Stirling, Taiwan and Thailand, and evaluated them in a diallel cross design in both fresh and saline water environments. The main objective of this work was to evaluate relative performance of strains in order to form a synthetic base population for future genetic selection.

MATERIALS AND METHODS

Origin of stocks. This study included four strains of red tilapia: 1) Malaysian red tilapia obtained from a private hatchery in Malaysia (M), 2) Thai red tilapia *O. niloticus* × *O. mossambicus* from the National Inland Fisheries Institute (N), Department of Fisheries, Thailand, 3) red *O. niloticus* originating in Lake Manzala, Egypt (McAndrew *et al.* 1988), and obtained from Institute of Aquaculture, University of Stirling (S) in United Kingdom, and 4) *O. niloticus* × *O. mossambicus* red tilapia from a private hatchery in Taiwan (T). The exotic strains of red tilapia from Malaysia, Stirling and Taiwan were imported into Thailand in mid-2005. They were kept and reared at the Pathumthani Fisheries Test and Research Center, Aquatic Animal Genetics Research and Development Institute, Department of Fisheries, Pathumthani province, Thailand (latitude 14°N, longitude 100°E, 20 km north of Bangkok). A detailed description of the stocks origin is given in Pongthana *et al.* (2009).

Family production and rearing. The pair matings following a complete diallel cross design (Table 1) were conducted in separate 1×1×1 m breeding hapas in February 2006. Sixty four breeding hapas were installed in a pond. In each hapa, one male was mated to one female. A total of 64 full-sib families (four families per cross combination) were successfully produced. After approximately one or two weeks of mating, swim-up fry were collected separately from each hapa and transferred to 1×1×1 m rearing hapas at a stocking density of 500 fry per hapa, one hapa for each full sib group. The date of collection of swim-up fry was recorded. After 3-4 weeks in the rearing hapas, the fry were transferred to B-net hapas (~ 6 mm mesh size) at a stocking density of 300 fry per hapa (1×1×1 m) for further rearing until an average body weight of 5 grams. When the fingerlings reached this weight they were individually tagged using Passive Integrated Transponders (PIT). A total of 20 fingerlings were tagged per family, amounting to 1280 tagged individuals.

Table 1. Diallel cross mating design involving four strains of red tilapia

Female/ Male	M	N	S	T
M	MM	MN	MS	MT
N	NM	NN	NS	NT
S	SM	SN	SS	ST
T	TM	TN	TS	TT

Testing environments. After tagging, the fish were tested in both freshwater floating cages and saline water concrete tanks. The floating cages were located in a 60,000 m^3 water reservoir inside the Pathumthani Fisheries Test and Research Center. Six cages (4 × 4 × 1.5m) adjacent to each other were assembled, and 2 fish from each family were assigned at random to these cages. The initial stocking density was 5 pieces per m^2 of surface water. The fish were fed twice daily a commercial pellet feed with a dietary protein level of 32%, at the rate of 3-5% of their body weight. Contemporaneously with freshwater cage culture, siblings from the 64 families were tested in a 30 ppt of saltwater environment in four concrete tanks (4 × 10 × 1m). The stocking density was 5 fish per m^2. The same feeding, culturing and management practices were applied as used for the cages in freshwater. In both environments, water quality parameters (temperature, pH, dissolved oxygen, alkalinity, total ammonia and saline level) were closely monitored once a week.

Harvest and measurements. Following a grow-out period of 126 days (range from 118 to 147 days), all fish were harvested and immediately transferred to large hapas for one to two days of conditioning without feeding before the individual identification, body measurements and sex were recorded. Over the growth period in cages and tanks 98% of the tested individuals retained their tags. Survival of the fish in both environments was high, averaging 84.3 per cent, calculated from the difference between the number of fish at stocking and at harvest. This trait was not included in the present analyses. Basic statistics of body weight at harvest is given in Table 2

Table 2. Descriptive statistics for harvest weight (g)

Environment	N	Mean	SD	CV
Fresh water	520	579.3	220.8	38.1
Saline water	529	429.2	160.6	37.4
Overall	1049	503.6	206.8	41.1

Statistical analyses. Model 1 was used to estimate additive and non-additive genetic components. The fixed effects included in the model were test environments (fresh and saline water), sex (female and male) and their two-way interactions. Age at harvest was fitted as a linear covariate. The genotype effects were partitioned in terms of direct additive, non-additive and reciprocal effects. The model also fitted full-sib group as the random term. The mathematical expression of the model is as follows:

$$y_{ij} = \mu + F + \sum \alpha_i a_i + \sum \alpha_{ij} h_{ij} + \sum \beta_i r_i + s + e_{ijk} \qquad [1]$$

where F are the fixed effects and covariate as described above, α_i is the proportion of genes contributed by the n^{th} individual originating from the i^{th} strain (α_i =0.0, 0.5 or 1.0 and $\Sigma\alpha_i$ =1.0); a_i is the additive genetic effect of genes originating from the ith strain; α_{ij} is the coefficient of the total heterosis effect for the cross between the i^{th} and j^{th} strains (α_{ij} =0.0 or 1.0; i≠j and ij≠ji and $\Sigma\alpha_{ij}$ =1.0); h_{ij} is the total heterosis effect for the cross between the ith and jth strains (i≠j and ij≠ji); βi is the coefficient of the general reciprocal effect for the ith strain (β_i =0 for purebreds and −0.5 for male strain and 0.5 for female strain, for the crossbreds and $\Sigma\beta_i$ =1.0); r_i is the general reciprocal effect of the ith strain; s is the random effect of full-sib group; and e_{ijk} is the random residual error for the lth individual.

The relative importance of the additive, heterosis and reciprocal effects was assessed using the likelihood ratio test (LRT) by removing each term from the full model at a time (Table 3).

Table 3. Increase in -2logL for harvest weight when one effect at a time was excluded from the model

Effect	Fresh water	Saline water	Both
Additive genetic (a_i)	31.9***	30.9***	30.1***
Total heterosis (h_{ij})	62.9***	60.3***	59.9***
General reciprocal (r_i)	41.4***	42.2***	29.0***

***P<0.001

The additive, heterosis and reciprocal effects were estimated as regression coefficients with one degree of freedom. The additive genetic effects were restricted to $\sum a_i = 0$. The coefficients of the general reciprocal effects set in the present study assume that the additive genetic effects of a given strain are similar regardless of gender of parental breeders. Total heterosis for a cross between two strains was partitioned as $h_{ij} = \bar{h} + h_i + h_j + s_{ij}$, where $\bar{h}$ is the average heterosis effect for all strains involved in the diallel cross, h_i and h_j are the general heterosis effects for the ith and jth stock, respectively, and s_{ij} is the specific heterosis effect of strains.

RESULTS AND DISCUSSION

Additive genetic effects. Table 4 presents estimates of strain additive genetic effects on body weight. Across the testing environments, the Malaysian strain ranked highest, whereas the additive performance of the strain from Stirling was poorest (7.4 and 13.4% above and below the overall mean of the pure strains, respectively). Overall, ranking of genotypes changed between the environments.

Heterosis. The average heterosis for body weight across the testing environments was low (approximately 4.2%) and not statistically different from the mean of pure strains. The level of

average heterosis was lower in fresh than in saline water, but the difference between the two environments was only about 2% (Table 3). The heterotic effect reported on body traits is generally small in other farmed aquaculture species such as Nile tilapia (Bentsen *et al*. 1998) or Rohu carp (Gjerde *et al.* 2000).

Reciprocal effects. In this study reciprocal effects can be considered to be equivalent to maternal components. Across the environments, the Malaysian strain ranked highest, whereas the strain from Stirling was lowest. Furthermore, the general reciprocal effects also varied with testing environments.

Table 4. Estimates of additive genetic effects and heterosis for body weight at harvest

	Fresh water		Saline water		Across	
	Estimate	%	Estimate	%	Estimate	%
Mean	532.9± 23.8		395.9 ± 23.7		465.5 ± 20.2	
Additive genetic (a_i)						
M	83.6	15.7	-0.7	-0.2	34.5	7.4
N	-13.2	-2.5	27.0	6.8	9.3	2.0
S	-61.7	-11.6	-69.7	-17.7	-62.1	-13.4
T	-8.7	-1.6	43.5	11.0	18.4	3.9
Maternal effect (r_i)						
M	17.3	3.3	34.8	8.8	29.6	6.4
N	43.6	8.3	-7.2	-1.8	16.3	3.5
S	-38.2	-7.3	-52.8	-13.3	-43.6	-9.4
T	-22.7	-4.3	25.2	6.4	-2.3	-0.5
Average heterosis ($\bar{h}$)	15.5	**2.9**	19.8	**5.0**	19.5	**4.2**

CONCLUSIONS

In contrast to the low level of heterosis for harvest weight, the large additive genetic component estimated in this study suggests that performance improvement of red tilapia could be effectively based on the exploitation of additive genetic variation (i.e. through selective breeding rather crossbreeding). This study also points to the relative performance of strains as a guide for the establishment of a synthetic base population for future genetic selection. The changes in both the additive and non-additive genetic performance of genotypes with testing environments in the present study also merits further examination to evaluate alternative breeding strategies for the future breeding program in red tilapia.

REFERENCES

Bentsen, H.B., Eknath, A.E., Palada-de Vera, M.S., Danting, J.C., Bolivar, H.L., Reyes, R.A., Dionisio, E.E., Longalong, F.M., Circa, A.V., Tayamen, M.M. and Gjerde, B., (1998) *Aquaculture* **160**:145.

Gjerde, B., Reddy, P.V.G.K., Mahapatra, K.D., Saha, J.N., Jana, R.K., Meher, P.K., Sahoo, M., Lenka, S., Govindassamy, P. and Rye, M., (2002) *Aquaculture* **20:**103.

McAndrew, B.J., Roubal, F.R., Roberts, R.J., Bullock, A.M. and McEwen, I.M., (1988) *Genetica.* **76**:127.

Pongthana, N., N.H. Nguyen and R.W. Ponzoni. (2009) Comparative performance of four red tilapia strains and their crosses in fresh and saline water environments. *In preparation*

DEVELOPMENT AND EARLY RESULTS OF THE TASMANIAN ATLANTIC SALMON BREEDING PROGRAM

N.G. Elliott and P.D. Kube

CSIRO Food Futures Flagship and CSIRO Marine and Atmospheric Research, Hobart, TAS 7000

SUMMARY

Atlantic salmon (*Salmo salar*) farming commenced in Tasmania in 1984, and after 20 years of a general 'no selection' breeding practice to maintain genetic diversity, a family based selective breeding program for the industry commenced in 2004. The commencement of a selection program was delayed due to concerns that the closed population had insufficient genetic variation to warrant a program. The results of DNA marker analyses and now four years of family production and assessment show the Tasmanian population has sufficient genetic variation to make good commercial gains from selective breeding. The Salmon Enterprises of Tasmania (SALTAS) selective breeding program is unique in that it is fully dependent on DNA genotyping for pedigree assignment, and does not use the traditional practice of maintaining multiple family tanks for the nursery phase. In addition, the program differs from other salmonid programs in having resistance to an external gill parasite as the priority breeding objective. The high impact of treating this marine parasite on production costs is a further distinction of the local industry, as well as a very high dependence on all-female commercial production. The latter influences the rate of transfer of genetic gains into commercial returns. The program reached a milestone in 2008 with the production of the first selected families, and in 2009 the first commercial productions were achieved.

INTRODUCTION

History. The first six generations of the Tasmanian Atlantic salmon population were produced with the aim of maintaining genetic diversity or, in other words, preventing possible inbreeding effects. Some selection on 'appearance' and avoidance of outliers was practised. This cautious approach reflected anecdotal information and the understanding of the genetic history of the population. Tasmanian Atlantic salmon originate from the east coast of Canada; imported first to Gaden, NSW in 1965 to 1967, and from there to Tasmania in 1984 to 1986 to start Australia's leading aquaculture industry.

Improving production traits in the population through selection was first considered in the mid-1990s, however, concerns over lack of sufficient genetic diversity due to purported low effective breeding numbers in the early years of domestication in NSW resulted in a continued cautious approach to establishing a selective breeding program.

Genetic diversity in the population and estimates of the historic effective population sizes was examined in multiple year-classes using microsatellite DNA markers (Innes and Elliott 2006). Variation in the local population was compared to that observed in archived scale samples from the ancestral Canadian population. These analyses suggested that, based on microsatellite analyses, sufficient genetic diversity remained in the local population to warrant the establishment of a family based selective breeding program.

Program design. Breeding objectives were relative clear from the outset. These were improving growth (time to harvest weight), increasing disease resistance, reducing incidence of early maturation, and maintaining carcass quality traits (flesh colour and fat). The program design was less obvious and needed to consider the relative size, structure and maturity of the local industry,

biosecurity restrictions on moving between marine and freshwater, lack of opportunity for introducing new genetic stock, and the resources available for the initiation and on-going management of a selective breeding program (or programs).

Traditionally in aquaculture, family based selective breeding programs have been designed around the use of family tanks for the hatchery and nursery phases. Thus each family is maintained in an individual growout unit until animals reach a suitable size for family or individual tagging, after which families can be mixed in a common growout unit. The capital outlay for a new or refurbished hatchery/nursery, with a minimum capacity of 100 family units each with controlled environment system, plus on-going management was beyond the available budget. Therefore a selective breeding program was designed on the basis of mixed family growout at all stages following hatching, and using DNA genotyping to provide the required pedigree information.

The selective breeding program (SBP) was commenced in 2004 by the leading commercial hatchery SALTAS on behalf of the Tasmanian industry. SALTAS is a cooperative venture, providing smolt to its industry shareholders based on their equity, with a current annual production of ca. 3.7 million smolt, which is ca. 40% of total smolt production for the Tasmanian industry. CSIRO through the Food Futures Flagship is partnering in the initial five years of research and development for the SBP.

The program revolves around a three-year production cycle (Figure 1), with three cycles in operation at any time. The operational plan involves freshwater spawning and nursery, with tagging and DNA fingerprinting at 12 months of age, followed by a cohort split into freshwater for growout and broodstock conditioning, and marine commercial growout and harvesting. Approximately 6,000 individuals are tagged and genotyped in each year class, with equal numbers in the marine and freshwater cohorts.

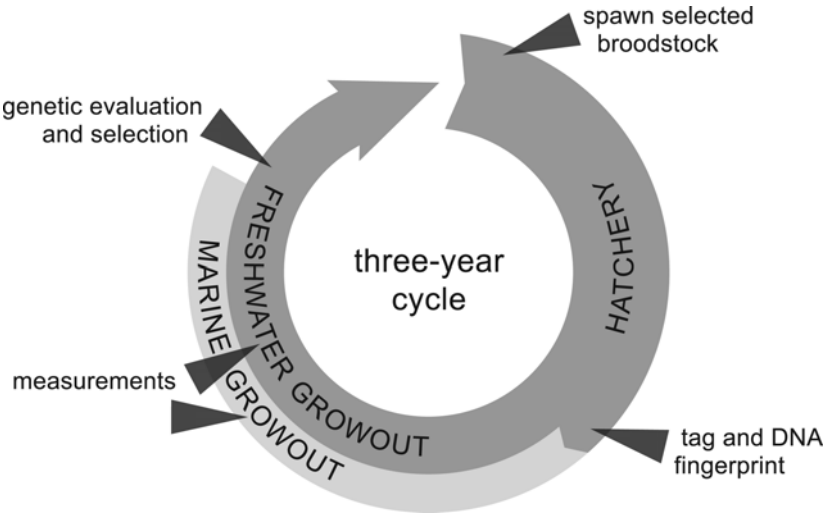


Figure 1. SBP year-class production cycle, showing cohort split at 14 months of age into the marine commercial growout and freshwater potential broodstock cohorts.

RESULTS AND DISCUSSION

Genetic variation. Five SBP year-classes (YC, 2004 to 2008) have been produced. In the first year, 2004, 70 full-sib families were produced and DNA pedigree (through Landcatch Natural

Selection, Scotland) successfully assigned 96% of a random sample of 4350 individuals to a family, with family representation at 22 to 117 individuals. Since then the SBP mating design has involved each broodstock mated twice, and the production of 180 families. Pedigree assignment success has improved, with 99% of 9341 individuals in the 2007YC being successfully assigned.

Moderate genetic variation for multiple traits has been measured across all year classes (Table 1). Consistent patterns of genetic variation have been observed across year classes and a comparison of the performance of repeated spawners (approximately 20 parents per year class) suggests good repeatability of genetic predictions across year classes. The results from the quantitative genetic analysis validated the early microsatellite DNA results that suggested sufficient genetic variation existed within this closed population to justify the development and running of a family selective breeding program.

Table 1. Heritabilities of selection traits in a combined analysis across year classes

Trait	Heritability (± se)
AGD resistance (gill score at first infection)	0.14 ± 0.03
AGD resistance (gill score at infections 2 to 5)	0.30 ± 0.04
Gutted weight (marine)	0.41 ± 0.04
Total weight (freshwater)	0.38 ± 0.04
Incidence early maturation (marine)	0.22 ± 0.03
Incidence early maturation (freshwater)	0.12 ± 0.02
Flesh colour (astaxanthin content)	0.44 ± 0.07
Flesh fat content (total fat content)	0.47 ± 0.05

Marine and freshwater correlation. Biosecurity issues necessitate that individuals that go into marine commercial growout do not return to the freshwater hatchery. Therefore freshwater broodstock selection is based on commercial performance of relatives. Reasonable genetic correlations (r_g = 0.70 ± 0.05) have been observed between same age freshwater and marine growth measures based on total weight.

Disease resistance. A major marine production cost for the Tasmanian industry, and one not experienced to the same degree in other Atlantic salmon producing countries, is treatment and lost productivity due to amoebic gill disease (AGD). The disease can be managed with freshwater bathing but this is costly in terms of both labour and equipment. The estimated cost of AGD to the industry (valued in 2006/07 at $281m) is in excess of $25m a year. A key research question was therefore whether there was genetic variation in resistance to this external gill parasite.

Results from four years' data show that AGD resistance is a quantitative trait, with moderate genetic variation at all infections (an individual fish may have between 6 to 12 infections and subsequent freshwater treatments during a marine growout). However, there is low genetic correlation between the first and subsequent infections (r_g = 0.29 ± 0.11) which suggests first and subsequent infections are different genetic traits. As a consequence, the selection index includes first infection and all subsequent infections as separate traits with a lower weighting on the first infection. Selection is for reduced gill symptoms and it is assumed that this will slow the rate of infection, increase the time between treatments, and therefore reduce the number of treatments required.

Commercial production. The Tasmanian commercial harvest is predominantly (ca. 95%) based on all-female production. This is due to a high incidence of early maturation in males in their first summer at sea, prior to harvest size, which results in poor growth, increased management costs

and lower survival. Although there is genetic potential to reduce the incidence of early maturation through genetic selection, the focus for the breeding program is on growth and disease resistance because all-female production offers a workable solution to this problem.

All-female production requires the making of sex-reversed females or neo-males. Therefore, infusing genetic gains from the selective breeding program into commercial production requires an additional step with a multiplier population to create the neo-males.

CONCLUSIONS

Development of a commercial selective breeding program in an aquaculture industry requires patience as many competing factors influence its progress. These include the structure, business plans and maturity of the industry, as well as the species' biology and commercial production systems. A selective breeding program has been successfully developed for Australia's major aquaculture industry, and moderate genetic gains in commercial production traits for Atlantic salmon are predicted. The program is unique in three major areas - being totally dependent on DNA genotyping for pedigree assignment, the highest selection weighting is on resistance for an external gill parasite, and delivering commercial benefit requires multiplying into an all-female production system. Despite concerns over the importation events of Atlantic salmon into Tasmania, this closed population has been shown to have good genetic variation and predicted genetic gains in key production traits from the SALTAS selective breeding program.

ACKNOWLEDGMENTS

The SALTAS Selective Breeding Program for the Tasmanian Atlantic salmon industry would not exist without the commitment and dedication of many in the industry. In particular the authors acknowledge the enthusiasm of staff at the Saltas hatchery and freshwater facilities (especially Dr Harry King) and at the Tassal marine grow-out and processing facilities. Drs Jay Hetzel and Brian Kinghorn are acknowledged for their value input to an early scoping study for the program.

REFERENCES

Innes, B.H. and Elliott N.G. (2006) *Aquaculture Research* **37:**563.

FIBRE PRODUCTION AND SHEEP BREEDING IN SOUTH AMERICA

R.C. Cardellino[1] and J.P. Mueller[2]

[1]Delta Consultants, Montevideo, Uruguay, [2]National Institute for Agricultural Technology, Bariloche, Argentina

SUMMARY

The production of wool and other fibres, in particular those produced by camelids (special fibres) in South America is reviewed. The two main production systems, namely commercial and smallholding are described, including their geographic locations. Wool is by far the main animal fibre produced in South America with a volume of 143 million kg (greasy) followed by the alpaca fibre with 4 million kg.

Sheep breeding programs in the four South American countries with the largest sheep populations are summarized. Argentina and Uruguay have the most developed sheep genetic improvement programs. Their evolution from the beginning in the 1970's until now is reviewed, including a description of the present programs.

INTRODUCTION

The South American sub-continent covers a range of environments in which sheep, goats and South American domestic camelids (llama and alpaca) produce meat, fibre, milk and skins for a large number of farmers, contributing substantially to their livelihoods and to the national economies. Given the variety of ecological and socio-economical conditions it is difficult to generalize on production areas, production systems and breeding practices. Sheep are most common in the vast temperate rangelands and deserts of the south, while goats are more common towards the north-east of the subcontinent, and camelids are largely found along the Central Andean region. Sheep are for the most part dual purpose (wool-meat) whereas goats are bred for meat and camelids are multipurpose. Dairying with small ruminants is not common in South America. Notable exceptions to this general description are hair-sheep in the northeast of Brazil, bred for meat and skins, and Angora goats and criollo goats in southern Argentina which besides meat produce mohair and cashmere.

Wool is by far the most important animal fibre in South America, but other fibres, usually called "special fibres" like alpaca, llama, and mohair are also produced in large quantities. In addition, vicuña, guanaco, angora (produced by rabbits), silk and cashmere fibres have great potential for development but at present the amount of fibre produced is still low (Table 1).

Table 1. Production of wool and other animal fibres in South America.

Species	Animal Population	Fibre	Fibre Production (kg, greasy)
Sheep	57,500,000	Wool	143,700,000
Alpaca	3,503,774	Alpaca	4,055,595
Llama	4,080,596	Llama	3,342,866
Angora goat	550,000	Mohair	825,000
Vicuña	319,547	Vicuña	5,580
Guanaco	577,697	Guanaco	1,500

Source: Cardellino and Mueller (2008), Quispe *et al*. (2009).

SHEEP PRODUCING AREAS

The main sheep production areas in South America are shown in Map 1. Three main sheep areas can be distinguished. The largest one, indicated as wool producing sheep area, includes the majority of Argentina, southern Chile, Uruguay and southern Brazil. In that area, wool or dual purpose sheep breeds predominate, mainly derived from the Merino. The second area, the criollo sheep region, includes the northern part of Argentina, and the Andean Altiplano regions of Bolivia and Perú. There is a third area, specifically a dry region in northeastern Brazil, where woolless hair sheep are raised basically for meat and skins.

Map 1. Sheep producing areas of South America.

SHEEP PRODUCTION SYSTEMS

Two major sheep production systems can be distinguished: commercial and smallholder. Commercial systems include farmers with a variety of flock sizes depending on the region, but oriented mainly to the production of wool and meat for the market. The main areas of these production systems in Argentina include the regions of Patagonia (dry and cold, Merino and Corriedale), Mesopotamia (mixed cattle-sheep farming, Corriedale and Polwarth) and the Pampas (mixed cropping-sheep farming, Corriedale, Romney Marsh and Lincoln) with about 50,000 growers. Uruguay, with 38,000 growers in mixed farms with beef cattle, run mainly dual-purpose sheep (Corriedale 60% and Merino 25%). Brazil, with 40,000 growers running mixed farms are located mainly in the southern region, with a predominance of dual-purpose sheep. In Chile there are about 60,000 growers. The Patagonian region in Chile contains 60% of the total sheep population in the country, consisting of dual purpose sheep breeds on medium to large farms.

Table 2 shows that commercial production systems comprise some 60% of total sheep numbers and account for 85% of the wool produced. It also can be seen that most wool marketed is fine or medium. It is estimated that about one third of coarse wool and camelid fibres produced by smallholders are not marketed and is instead used in the household or transformed and sold with aggregate value as handcrafts.

Table 2. Sheep population and wool production in South American countries (2008).

Country	N° sheep (mill)	Prod. System	Fine < 24.5 mic	Medium 24.6 – 32.5 mic	Coarse/Criollo > 32.5 mic	Wool Production (mkg, greasy)
Argentina	16.0	commercial/ smallholders	40.3	22.7	2.0	65.0
Uruguay	10.4	commercial	12.0	27.0	2.0	41.0
Chile	3.9	commercial	0.2	10.8	0.2	11.2
Brazil	3.5	commercial	1.0	8.5	1	10.5
Perú	14.7	smallholders	0	5.0	7	12.0
Bolivia	9.0	smallholders	0	0	4	4.0
Total	57.5		53.5	74.0	16.2	143.7

Source: Cardellino, RC. based on FLA, SUL, ODEPA, IICA, IWTO.

The second sheep production system is the smallholder system which corresponds to low input, low productivity small farms with subsistence economies. Flock size is small, 20-40 head and usually mixed with goats or camelids. Sheep are of the criollo type (derived from the original sheep introduced by the Spanish settlers) or non-defined criollo crosses. Main areas where these types of production systems can be found include: the Altiplano of Bolivia, a region at 3000-4500 metres above sea level (masl), involving mostly native communities; the Sierra Central Region of Peru and the area north of the Titicaca Lake, with 43% of very small producers and 32% of peasant communities. The Altiplano sheep production systems extend to the northwest of Argentina (Tempelman and Cardellino 2007).

SHEEP BREEDING PROGRAMS

Argentina. As in other countries of South America, Spanish sheep were introduced into Argentina soon after the discovery. In the early 19th century these populations called “criollos” were upgraded to Saxon, Negrete and Rambouillet Merinos aiming for a better carcass and finer wool. Selection within the resulting genetic pool gave origin to the Argentine Merino. By the end of that century an active European mutton market and the advent of cold-storage plants and ships encouraged the use of the large British Lincoln and Romney Marsh breeds. Both breeds were used for crossing and upgrading. In the 1930s the decreasing demand for sheep meat and increasing wool prices were responsible for the preference of Australian Merinos, Polwarths and Corriedales which were imported from Australia and New Zealand.

Within breed improvement was initiated by pioneer breeders who later associated into Breed Societies and implemented pedigree registration following British Flock Book rules. These Societies are largely responsible for the present genetic structure, production level and genetic improvement rate of most breeds. Today a hierarchical genetic structure characterize Argentine sheep breeds.

Performance recording started in the 1960s, heavily influenced by Australian scientists such as Helen Newton Turner and Brian Jefferies who visited and worked in Argentina as consultants for a FAO funded Patagonian Sheep Production project run by the National Institute for Agricultural Technology (INTA). Following their advice, a fleece testing laboratory was constructed and objective fleece testing was promoted. A performance testing scheme was implemented in 1970 as a selection aid for Merino and Corriedale ram breeders in Patagonia. In 1990 the program

developed into the National Sheep Evaluation Service (PROVINO) based on a joint agreement between INTA and several breed societies. Initially the program provided within contemporary group breeding values and sire-summaries.

In 1991 Central Progeny Testing (CPT) of sires started in the Merino breed followed by Corriedale, Romney Marsh and Polwarth. Particularly successful is the Merino CPT scheme which is still operating. Interestingly the Merino CPT scheme often includes show champions and imported sires on invitation. It has to be remembered that major Argentine Merino studs regularly import rams or their semen from Australia. Amongst the imports are Australian show champions including supreme champions (see for example Top Sire 2005). Several of these studs furnish PROVINO with pedigree information allowing the calculation of within flock BLUP of expected progeny differences (EPDs). Since 2006 the accumulated data from these studs and those from the CPT sites are merged enabling across flock evaluations which are published by the Argentinean Merino Breeders Association (see sire evaluation results at AACM 2009).

Thus, at present PROVINO provides two types of genetic evaluations: within flock EPDs and across flock EPDs. At the end of the 2007/2008 production year, PROVINO evaluated about 16,000 ram hoggets from a total of 176 farms and ranked 236 Merino sires. These figures highlight that selection procedures are still largely based on visual inspection only. An important step towards the use of objective measurements was taken by the Argentine Merino Breed Society with its "Pure Registered Merino" program based on identification of rams performing above average (on breeder chosen PROVINO index) in its contemporary group and visually accepted by an authorized classer.

A further important development has been the enforcement of a "National Sheep Recovery Law" in 2001, which provides state funds for stock recovery, farm infrastructure, feed production, breeding plans, large scale sheep health programs, AI projects, semen imports and central progeny testing.

Brazil. Brazil has around 14 million sheep concentrated in two main regions: South (wool sheep) with 33% and Northeast (hair sheep) with 56% of the total population. In the southern region temperate and subtropical climate predominates. Sheep improvement began with the absorption of mixed criollo flocks of Spanish origin (mainly Churro) by traditional wool and meat sheep breeds (Merino, Polwarth, Corriedale, Romney Marsh and Hampshire Down), followed by 50 years of visual selection by breed standards. In 1942 the Brazilian Sheep Breeders Association (ARCO) was created.

In 1978 the Sheep Genetic Improvement Program (PROMOVI), based on objective measurement of wool traits, was initiated and remained officially active from 1981 to 1995. The importation of meat breeds (Ile-de-France, Suffolk and Texel) increased in the 1990s. From 1995 to 2000 ARCO implemented the Central Performance Test for meat breed lambs and a Performance Test on farms. Since 2000 no objective evaluations that could serve as basis for sheep genetic improvement programs have been carried out in this region (Cardellino 2009).

Semi-arid conditions predominate in the Northeast region of Brazil. Hairy sheep originated from African and Caribbean stocks and resulted in well defined naturalized breeds (Santa Ines, Morada Nova, Rabo Largo and Cariri). Since 2002 a research institute in Paraiba State (EMEPA) and the sheep national centre in Ceara State (EMBRAPA) have conducted feedlot performance testing in Santa Ines males. Other initiatives are the Genetic Improvement Program for Goats and Meat Sheep (GENECOC) by EMBRAPA and the Genetic Improvement Program for Santa Ines breed by the University of Sao Paulo.

All current programs have a very low level of participation by sheep breeders, especially studs. In an expanding market, these breeders obtain good prices for their sheep stock (rams, hoggets and ewes) based only on visual selection and aggressive marketing. In 2008 EMBRAPA proposed the

development of a Research Network for Sheep Genetic Improvement with the participation of its research units, universities and state research organizations. It is being currently developed under a service contract with ARCO and aims at establishing a national sheep genetic improvement scheme with enough versatility to include all breeds of wool and hair sheep in both main sheep producing regions of the country (Benítez *et al*. 2008).

Chile. At present there are no regional or national sheep breeding programs operating. In the Chilean Patagonia, where most of the sheep are concentrated, a cooperative nucleus breeding scheme in the Corriedale breed was set up in the 1980s supported by the National Institute of Agricultural Research (INIA). More recently a number of breeds were introduced and evaluated, including Dohne Merino, Poll Dorset, Suffolk and East Friesian. (R. Lira 2009, pers.comm). Breed improvement relies heavily on conventional procedures, including visual classing of replacement candidates and promotion of show champions. Nevertheless Corriedale rams have ranked high in progeny tests including New Zealand and Argentinean sires (Mueller *et al*. 2007).

Peru. The sheep population in Peru is estimated to be 14 million head, 85% of which is run by peasant communities, and belong to the criollo breed and their non-planned crosses (W. Vivanco 2009 pers.comm.). The remaining 15% correspond to Corriedale and Junín (a local breed based on the Corriedale, Polwarth and Panamá breeds), and are in the hands of communal cooperatives and private farmers. There are no well structured and organized sheep breeding programs in the country, nor performance recording systems and genetic evaluations. Large scale improvement programs have been successfully implemented with support of the University of La Molina (Mueller *et a*l. 2002). Recently, the Dohne Merino and the East Friesian have been introduced.

Uruguay. The predominant breeds in Uruguay are Corriedale, Merino and Polwarth, which represent 60, 20 and 10 per cent of the national sheep flock, respectively. These breeds can be defined as multi- purpose in the sense that they generate income from the sale of wool and sheep meat (surplus offspring and cast for age animals), forcing breeders to consider several traits in their selection programmes. Traditionally, wool has been the main product of the system. However in recent years, the importance of sheep meat (lambs and mutton) has increased significantly. The breeding structure of sheep industry in Uruguay follows in general the common hierarchical pattern with “top”, and “multiplier” studs. Flock size in studs is large enough to allow effective “within flock” programs of genetic improvement (on average 500 breeding ewes). The evolution of sheep breeding programs and genetic evaluation procedures since 1970 to date has followed the same pattern of more developed sheep producing countries, mainly Australia and New Zealand.

A performance recording scheme started in 1969 with the purpose of introducing more effective breeding methods, through the objective measurement of economically important traits. This service was implemented and supervised by the Uruguayan Wool Secretariat (SUL), a private growers´ organization. The Flock Testing service is still operating and involves at present 137 studs and 23,000 animals in 2008. The rate of adoption by the most important studs (top studs) has been very high (95% in Corriedale, 80% in Polwarth and Merino studs).

After many years of operation of the Flock Testing service, it gradually became obvious that a formal definition of breeding objectives and selection criteria was essential. It was recognized that such a formal definition would enable to more precisely determine the relative importance of different traits and also to offer breeders the possibility of combining various selection criteria in an index (Cardellino and Ponzoni 1986). In recent years, new updated selection indices were developed for general use.

In 1994, with the objective of comparing the genetic merit of rams from different studs, several Central Progeny Test (CPT) operations started in all the breeds, with the support of their Breeders’

Associations and the technical support of SUL. Until year 2000, dam identification was not recorded, so genetic evaluation of different traits was performed with a sire–model. Gradually the concept of reference sires was introduced, allowing the genetic comparison of rams used in different CPTs and years and also the feasibility of performing across–flock genetic evaluations became more obvious. Multi–trait BLUP methodology was then introduced as well as the concept of EPD's (Expected Progeny Differences), replacing the traditional use of phenotypic values in the selection of the animals. At the same time, new traits were included in the breeding plans: FEC (faecal egg count), as indicator of genetic resistance to internal parasites, weaning weight, eye muscle area and fat depth.

The experience acquired and the results produced by CPTs, created a suitable environment to start with genetic evaluation across–flocks, through the use of reference sires among the participating studs. At present the number of studs involved in across–flock evaluations is: Corriedale (24), Merino (14), Polwarth (6), Merilín (5), Texel (5) and Romney (3) (see sire evaluation results at SUL 2009). Genetic evaluation analyses are performed by SUL and INIA. Every year special meetings with stud breeders and technical staff are carried out, to define the use of "reference sires", checking that the whole system remains well connected.

The initial system of collecting and storing data was not reliable enough which represented a limitation to the growth of a sire reference scheme with a high standard. SUL developed software (called SULAR) which performs quality control of data when it is incorporated. Special emphasis was placed in the registration of information on the performance of dams, and survival of lambs with the objective of producing EPD's for these traits. The centralized data base stores the identification of the animals, plus information on production and pedigree records (Gimeno and Cardellino 2006).

In 1998, it started a Fine Merino Project with the participation of the Merino Breeders' Association, INIA and SUL. A highly selected screened nucleus was formed with the objective of generating and distributing genetically superior sires for the production of fine and superfine wool less than 19.5 microns (Montossi *et al.* 2005). The initial nucleus included 742 ewe hoggets selected from 37 contributing flocks, screened from a population of 5170 ewe hoggets. Contributing producers annually receive genetically superior hogget rams, for their own use. Frozen semen from highly selected Australian merino rams has been used, particularly from the New England area. Results so far have shown very good progress in reducing diameter, while maintaining fleece and body weights.

WOOL vs MEAT IN SHEEP BREEDING PROGRAMS

Historically the production of sheep meat has been a by-product of wool production, with practically no areas specialized in the production of fat lambs. However, in the last 5-10 years, as a result of low and fluctuating wool prices, meat production has increased its importance in sheep production systems, representing a higher proportion of total sheep income.

Uruguay is the first exporter of sheep meat in the region with 22,000 ton, followed by Argentina (6,100) and Chile (5,100). However, the importance of fat lamb production and consequently prices for lamb meat are much lower than Australia and New Zealand, probably due to a lower exporting capacity and also very reduced home market consumption (6.2, 2,5 and 0.3 kg/head in Uruguay, Argentina and Chile, respectively)

In general terms, in Argentina and Uruguay, the increased importance of meat relative to wool has lead to changes in their genetic improvement programs, in particular: the inclusion of new traits more related to meat production (type of birth, weaning weight, muscle and fat depth) as well as changes in the relative importance of meat traits in the breeding objective (H) and the corresponding Selection Indices (I) for the different breeds. In addition, direct and maternal breeding values for weaning weight have been included.

This new scenario has also led to the introduction of new breeds, some of them for terminal crossbreeding (Poll Dorset, Suffolk), and others like the Dohne Merino, in particular to be used on Corriedale flocks to reduce the diameter of the wool, without lowering lamb meat production.

This general trend to improve meat production has also been followed by a generalized emphasis in producing finer wools, both in Merino and Corriedale flocks.

THE PRODUCTION OF OTHER ANIMAL FIBRES

In South America these fibres are produced almost exclusively by smallholders in low input systems where they are critical for the subsistence of its producers. Most fibre production systems are located in marginal areas with goats and camelids grazing natural rangelands. Alpacas, llamas and vicuñas are typically found in high altitudes of the central Andes while goats producing mohair or cashmere and guanacos are largely found in the Patagonian desert.

Alpaca. The Alpaca (*Vicugna pacos*) is a domesticated South American camelid species whose wild ancestor is the vicuña. Alpacas are raised in the highlands of Peru, Bolivia and Chile. More than 80% of the world's alpaca population can be found in southern Peru, northwest of Lake Titicaca at 3700-5000 masl. The alpaca is a symbol of Peruvian national identity. 85% of the alpacas are run by smallholders with less than 50 animals each, or are kept in farmer communities. Alpacas are particularly prized for their fibre, which is noted for its fineness, softness, exceptional warmth, hygroscopic features, resistance, elasticity, and natural colors. The soft touch is related to the fineness of the fibre but also to the arrangement of the scales along the fibre. "Baby" alpaca fibre diameter averages 22 micron and alpaca "fleece" averages 26 micron. An adult alpaca produces 1.5-2.8 kg of fibre per year. Alpaca is the main special fibre produced in South America.

Llama. The llama (*Lama glama*) is the other domesticated South American camelid species, its wild ancestor being the guanaco. Both, llamas and guanacos are larger animals than alpacas and vicuñas. Bolivia has the largest llama population, about 2.4 million, largely on the high-plateau (Altiplano) at 4000 masl in the west of the country. Peru with 1.2 million llamas is the second largest producer, while Argentina ranks third. It is estimated that in Bolivia there are 54,000 producers, 80% having less than 90 llamas each. Llamas are multipurpose animals; they are raised for their meat, power and fibre. As with alpacas, there is a strong cultural tie between llamas and their producers and communities. The fibre produced by llamas is not as fine as that of the alpacas. In Bolivia adult llamas produce fibre with an average diameter of 33 mic, and may yield up to 93% of its original weight when processed. Due to its multiple breeding objectives, llamas were selected for high body weight and fleece weight (1.5-3.5 kg) but less for fibre traits such as fineness and uniformity of color, which are more difficult to measure.

Vicuña. The vicuña (*Vicugna vicugna*) is the smaller of the two wild South American camelids and its undercoat fibres are extremely valuable and "special", not only for its textile characteristics but also for its rareness and association to exotic environments and culture. After a period of near extinction, the vicuña population recovered substantially in population size in all Andean countries. In Peru the vicuña population is now 140,000 and increasing, as well as in Argentina with a population of 133,000. Vicuñas are captured, shorn and released using different methods. At present Peru is producing most of the vicuña fibre, about 5500 kg/year with a fibre diameter between 10-15 mic. Yarn and fabrics made of vicuña fibre have the highest market price of all special fibres but its production is not easy due to its short staple length and the necessity of separating manually guard and dead fibres from the fine down fibres.

Guanaco. The guanaco (*Lama guanicoe*) is the biggest of the two wild South American camelid species and its population is much larger than that of the vicuña. More than 90% of the world guanaco population is in Argentina and the remainder in Chile and Peru. The population of 550,000 guanacos in Argentina is largely concentrated in the southern part of the country (the Patagonian desert). Guanacos roam freely in this sparsely inhabited region where sheep production is the main agricultural activity. Capture of guanacos is difficult as they can easily jump regular fences to escape at very high speed when mustered. Special techniques have been developed in order to capture, calm, shear and release guanacos avoiding unnecessary fear and injury. Fleece weight is approximately 1 kg for a two-year growth. Guanaco fibre is not as fine as that of the vicuña (16-22 microns) but otherwise quite similar, including its color variations of brown.

Mohair. About 650,000 Angora goats are run in the northwest of Argentina's Patagonia where they produce 825,000 kg mohair of competitive quality. Argentina is among the top world producers of mohair. Notable of Angora goats in this part of the world is the uniformity in color. Almost all Angora goats in Argentina are white, as opposed to central Asian Angora goats where other colors are very common. Angoras are shorn twice a year and produce a total of about 1.5-2.5 kg mohair. Mohair from young animals, (first and second shearing) is much finer (24 mic) than mohair from adult animals (29 micron and more).

ACKNOWLEDGEMENTS

The authors wish to thank Drs. Daniel Benítez (Brazil), Ricardo A. Cardellino (Brazil), Raúl Lira (Chile), Diego Gimeno (Uruguay) and William Vivanco (Perú) for their reviews, comments and valuable information used in this paper.

REFERENCES

AACM (2009) http://www.merino.org.ar/index2.htm (accessed May 2009).

Benítez, D., Cardellino, R.A. and Souza, W.H. (2008). Soc. Bras. Melh. Animal, São Carlos, São Paulo, Brazil. http://www.sbmaonline.org.br/.

Cardellino, R.A. (2009) Revista Ceres, Universidade Federal de Viçosa, Minas Gerais (in press). Requests to rcard@ufpel.edu.br.

Cardellino, R.C. and Mueller, J.P. (2008) Natural Fibres Symposium. FAO, Rome 2008 (in press)

Cardellino, R.C. and Ponzoni, R. (1986) 3rd World Congr. Genet. Appl. Livest. Prod. IX 650-657.

Gimeno, D. and Cardellino, R.C. (2006) CD ROM 8th World Congr. Genet. Appl. Livest. Prod.

Montossi, F., De Barbieri I., Mederos, A., Ciappesoni, G., Furgón, J.,Martinez, H., Luzardo, S.,Nolla, M., Dighiero, A., Levratto, J., Grattarola, M., Fros, A. and Pérez Jones, J. (2005) Serie de Actividades de Difusión Nro. 439, INIA, Uruguay.

Mueller, J.P., Flores, E.R. and Gutierrez, G. 2002. 7th World Congr. Genet. Appl. Livest. Prod. Comm. 25-12.

Mueller, J.P., Paz, A.P. and Suárez, P. 2007. XIII World Corriedale Congress, Lincoln, New Zealand. Comunicacion Técnica INTA EEA Bariloche PA 515.

Quispe, E.C., Rodriguez, T.C., Iñiguez L.R. and Mueller, J.P. (2009) Producción de fibra de alpaca, llama, vicuña y guanaco en Sudamérica. Animal Genetic Resources Information, FAO (in press).

SUL (2009) http://www.sul.org.uy/EvalGene.htm (accessed May 2009)

Tempelman, K.A. and Cardellino, R.A., Eds (2007), FAO, Rome, Italy. Download: *People and Animals* at: http://www.fao.org/ag/againfo/resources/en/pubs_gen.html

Top Sire (2005) http://www.merinonsw.com.au/Topsire_julysept_2005.pdf (accessed May 2009)

EFFECTS ON LIFETIME REPRODUCTIVE PERFORMANCE OF PHENOTYPIC SELECTION FOR FLEECE WEIGHT, FIBRE DIAMETER, BODY WEIGHT AND RELATED SELECTION INDEXES

L.R. Piper[1], A.A. Swan[2], and H.G. Brewer[1]

[1]CSIRO Livestock Industries, F D McMaster Laboratory, Locked Bag 1, Armidale, NSW 2350,
[2]Animal Genetics and Breeding Unit, University of New England, Armidale, NSW 2351

SUMMARY

The effects of simulated selection at hogget age for fleece weight, fibre diameter, body weight and two relevant selection indexes on lifetime fertility (EL/EJ), litter size (LB/EL), lamb survival LW/LB) and reproduction rate (LW/EJ) has been examined for a medium-wool random breeding control flock of Merino sheep grazing in south west Queensland. There were no significant effects on lifetime reproduction rate or any of the three component traits, of selection for fleece weight, fibre diameter or either of the two selection indexes. Selection for body weight had a significant positive effect on lifetime litter size ($p<0.001$) and an almost significant positive effect on lifetime reproduction rate ($p=0.059$). These data do not support the view that phenotypic selection for increased fleece weight will have adverse effects on lifetime reproduction rate.

INTRODUCTION

A recent series of papers (Cloete *et al.* 2002; Greeff 2005; Adams *et al.* 2006) has suggested that reproductive performance may potentially be compromised in animals with increased capacity for wool production especially if feed supplies are reduced. To further examine this proposition, Piper *et al.* (2007) analysed data from long term selection and control flocks of medium-wool Merino sheep grazing at Cunnamulla in the semi-arid zone of south west Queensland. Rainfall in the Cunnamulla region averages about 375 mm *per annum* but there is considerable annual variation and rainfall unreliability is the main factor limiting feed supply from pasture. In the selected lines fleece weight increased at about 2 % per year but as expected did not change in the control line. There was no change in lifetime reproduction rate (LW/EJ) in either the selection or control lines. The authors concluded that "long term breeding programs for Merino sheep, which include increased fleece weight as a component of a multi-trait breeding objective, can be implemented without necessarily reducing reproduction rate."

Hatcher and Atkins (2007) examined the effects of simulated phenotypic selection for hogget fleece weight, fibre diameter and body weight on lifetime reproduction rate and its components in ewes from the medium-wool strain of the Trangie QPLUS flock (Taylor and Atkins 1997). They found that within-flock selection for body weight would lead to significant improvements in reproductive performance, for fibre diameter would have no significant effects on reproductive performance and for fleece weight would lead to fewer progeny surviving to weaning. The data analysed by Hatcher and Atkins (2007) came from four flocks undergoing long term selection for a range of micron premium breeding objectives and from a related control flock. It is not clear whether the reproductive performance results may have been influenced by including data from the four long term selection flocks where the breeding objectives and selection indexes included the traits fleece weight, fibre diameter and body weight. In this paper, to avoid any possible influence on the results of including data from long term selection flocks, we have analysed data from a random mating control flock grazing in the relatively harsh semi-arid environment of the Cunnamulla district in south west Queensland. In the analyses reported below, we have examined the effects on lifetime reproduction rate and its components, of simulated phenotypic selection for fleece weight, fibre diameter, body weight and two related selection indexes.

MATERIALS AND METHODS

Sheep. The reproductive performance of 615 medium-wool Merino ewes, first mated at around 18 months of age between 1950 and 1964, was the focus of these analyses. The foundation ewes for this flock were typical medium-wool Merino ewes of mixed Peppin origin. The mating design for the flock has been described by Turner *et al.* (1968) and the environment and management of the flock at the CSIRO National Field Station, Gilruth Plains, Cunnamulla, Queensland, has been described by (Turner *et al.* 1959).

Observations and data analysis. Ten wool and body characteristics were measured on all animals using the techniques described by Turner *et al.* (1953). For the purpose of these analyses, the data consists of measurements of greasy fleece weight (GFW), fibre diameter (FD), and body weight (BWT) taken from 18 *mo* ewes (previously shorn at 6 *mo*) and the reproduction records (fertility (EL/EJ), litter size (LB/EL), lamb survival (LW/LB), and reproduction rate (LW/EJ) of the same ewes at their first six lambings (aged 2-7 years). All of the wool and body measurements and most of the reproduction records were obtained at Gilruth Plains. The flock was transferred to the CSIRO Field Station, Longford, Armidale, NSW in 1966 and some of the later reproduction records for the 1961 to 1964 drop ewes were obtained at Longford.

Allocation of ewes to high and low groups for the production traits. Linear models adjusting for significant fixed effects were fitted using the statistical software R (R Development Core Team, 2008). For GFW and BWT these effects included contemporary group defined as year of birth by management-flock subclasses, birth type, and rearing type, all fitted as factors. Age of dam (years) and age of measurement (days) were fitted as covariates, including a quadratic term for age of dam. For FD, only contemporary group and birth type were significant.

Residual values from these single trait models were used to allocate animals to High and Low trait groups within each year of birth, thus simulating current flock selection. Animals with residual values superior to the median value for the year were allocated to the High group, and those with values inferior to the median were allocated to the Low group. The mean difference in performance between the High and Low groups for each trait is shown in Table 1.

Table 1. Predicted means for, and differences between the High and Low groups for GFW (kg), FD (micron), BWT (kg), and the Merino 7% and 14% indexes (M7 and M14)

	High (se)	Low (se)	H-L	(H-L)/L*100
GFW	3.89 (0.04)	3.25 (0.04)	0.64	19.6
FD	23.19 (0.15)	20.61 (0.15)	2.58	12.5
BWT	32.80 (0.28)	27.97 (0.28)	4.83	17.3
M7	105.31 (0.46)	94.54 (0.47)	10.77	11.4
M14	106.59 (0.59)	93.21 (0.60)	13.38	14.4

The residual values for fleece weight and fibre diameter were also used to calculate selection indexes for the Merino 7% and 14% breeding objectives used by MERINOSELECT (Swan *et al.* 2007). Selection index weights were derived for these objectives using MERINOSELECT relative economic values and genetic parameters, assuming the measurements available included own performance for greasy fleece weight and fibre diameter. The index weights (dollars per ewe) for greasy fleece weight and fibre diameter were 9.8 and -3.6 for the Merino 7% objective, and 5.9 and -5.1 for the Merino 14% objective. Animals were allocated to High and Low index groups

within year of birth using the procedure described above for individual traits. Differences in performance for the two indexes are shown in Table 1.

Analyses of the reproduction data. Repeated record mixed linear models, adjusting for fixed effects were fitted using ASReml (Gilmour *et al.* 2006). The effects fitted included management group defined as year of birth by management-flock subclass, birth type, age of dam (years), own age (years) and group (high or low) all fitted as factors with ewe fitted as a random effect. Management group was significant (P<0.001) for all combinations of reproduction and production traits. Own age was significant (P<0.001) for all combinations of production traits and the reproduction traits LB/EL and LW/EJ but not for any of the production trait combinations with EL/EJ or LW/LB. Birth type and age of dam were not significant for any combination of the reproduction and production traits.

RESULTS AND DISCUSSION

The number of observations for each of the reproduction trait analyses was 2461 for fertility, 2185 for litter size, 2177 for lamb survival and 2454 for reproduction rate. The predicted mean values for the high and low groups for each production trait by reproduction trait combination are shown in Table 2.

Table 2. **Predicted mean values (se) for the high and low groups for each production trait by reproduction trait combination**

	Fertility (EL/EJ)	Litter Size (LB/EL)	Survival (LW/LB)	Rep Rate (LW/EJ)
GFW - H	0.91 (0.01)	1.14 (0.01)	0.75 (0.01)	0.81 (0.01)
GFW - L	0.90 (0.02)	1.14 (0.01)	0.77 (0.01)	0.82 (0.02)
FD - H	0.91 (0.01)	1.14 (0.01)	0.76 (0.01)	0.82 (0.02)
FD - L	0.89 (0.01)	1.13 (0.01)	0.76 (0.01)	0.81 (0.02)
BWT - H	0.91 (0.01)	1.18 (0.01) ***	0.75 (0.01)	0.83 (0.02) †
BWT - L	0.90 (0.01)	1.10 (0.01) ***	0.77 (0.01)	0.80 (0.02) †
M7 - H	0.91 (0.01)	1.13 (0.01)	0.76 (0.01)	0.81 (0.02)
M7 - L	0.90 (0.01)	1.14 (0.01)	0.76 (0.01)	0.81 (0.02)
M14 - H	0.90 (0.01)	1.13 (0.01)	0.76 (0.01)	0.81 (0.02)
M14 - L	0.91 (0.01)	1.15 (0.01)	0.76 (0.01)	0.81(0.02)

Significance of difference between high and low groups.;. *** P<0.001; † P=0.059; remainder, ns

Inspection of the contrasting high and low group means for each production trait, reproduction trait combination in Table 2 shows that with only two exceptions, there were only negligible effects of simulated selection for production traits on subsequent lifetime reproductive performance. The exceptions were that simulated selection for increased body weight produced a significant increase (P<0.001) in litter size and an almost significant increase (P=0.059) in reproduction rate.

The differences between the high and low groups and the percentage difference between them (H-L/L*100) are shown in Table 3 for each production trait – reproduction trait combination. With the exception of the effects on litter size and reproduction rate of selection for body weight, the majority of the differences are less than one percent and there is no consistent pattern of positive or negative effects. These very small and non-significant effects of phenotypic selection for production traits on reproductive performance have occurred despite the highly significant and sizable effects of that simulated selection on the differences between the high and low groups for the production traits (Table 1).

These results confirm the findings of Hatcher and Atkins (2007) in respect of the positive effect of selection for body weight on reproductive performance and the lack of effect of simulated selection for fibre diameter on reproductive performance. However, in this study there was also no effect on reproductive performance of selection for fleece weight or of selection for optimal indexes (seven and fourteen percent micron premium) combining fleece weight and fibre diameter. If anything, the grazing environment at Cunnamulla would be expected to be somewhat more harsh or marginal than that at Trangie and the proposition that sheep with high potential for fleece production may be compromised in respect of reproduction rate when unfavorable feed conditions are more likely to occur (Adams *et al.* 2006) was not observed in this study.

Table 3. **Percentage differences (H-L/L*100) between the high and low groups for each production and reproduction trait combination.**

	Fertility (EL/EJ)	Litter Size (LB/EL)	Survival (LW/LB)	Rep Rate (LW/EJ)
GFW	0.42	0.13	-1.42	-1.19
FD	2.29	0.85	-0.34	1.03
BWT	0.13	6.62 ***	-0.82	3.85 †
M7	0.40	-1.03	0.62	0.41
M14	-0.50	-1.34	0.49	-0.34

Significance of difference between high and low groups; *** P<0.001; † P=0.059; remainder, ns

CONCLUSIONS

The results from this study, which examined the phenotypic consequences of simulated selection for production traits on reproductive performance and of the previous study (Piper *et al.* 2007) which focused on the genetic consequences of such selection, do not support the view that sheep with increased capacity for wool production may have reduced reproductive performance when variable feed availability challenges animal production from pasture. These current findings are consistent with published estimates of the phenotypic correlations among the traits examined.

REFERENCES

Adams, N.R., Briegel, J.R., Greeff, J.C. and Bermingham, E.N. (2006) *Aust. J. Agric. Res.* **57**:27.

Cloete, S.W.P, Greeff, J.C. and Lewer. R.P. (2002) *Aust. J. agric. Res.* **53:**271.

Gilmour, A.R., Gogel, B.J., Cullis, B.R., and Thompson, R. (2006) ASReml User Guide Release 2.0 VSN International Ltd, Hemel Hempstead, HP1 1ES, UK.

Greeff, J.C. (2005) *Proc. Assoc. Advmt. Anim. Breed. Genet.* **16:**16.

Hatcher, S. and Atkins, K.D. (2007) *Proc. Assoc. Advmt. Anim. Breed. Genet.* **17:**260.

Piper, L.R., Swan, A.A and Brewer, H.G. (2007) *Proc. Assoc. Advmt. Anim. Breed. Genet.* **17:**103.

R Development Core Team (2008). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org.

Swan, A.A., van der Werf, J.H.J. and Atkins, K.D. (2007) *Proc. Assoc. Advmnt. Anim. Breed. Genet.* **17**: 483.

Taylor, P.J. and Atkins, K.D. (1997) *Wool Tech. Sheep Breed.* **45:**92.

Turner, H.N., Dolling, C.H.S. and Kennedy, J.F. (1968) *Aust. J. agric. Res.* **19:**79.

Turner, H.N., Dolling, C.H.S. and Sheaffe, P.H.G. (1959) *Aust. J. agric. Res.* **16:**699.

Turner, H.N., Riches, J.H., Hayman, R.H., Roberts, N.F. and Wilson, L.T. (1953) Physical definition of sheep and their fleece, with special reference to the Merino. CSIRO Div. Anim. Health, Div Rept. No. 4, Ser. SW-4.

VARIATION IN THE LIFETIME REPRODUCTIVE PERFORMANCE OF MERINO EWES

G. J. Lee, K. D. Atkins and M. A. Sladek

Cooperative Research Centre for Sheep Industry Innovation, NSW Department of Primary Industries, Orange Agricultural Institute, Forest Road, Orange, NSW 2800, Australia

SUMMARY

The variation between ewes in lifetime (2-6 years of age) reproductive performance within 3 research flocks maintained at Trangie in central western NSW was analysed. For each of the flocks, there were large differences ($P<0.001$) in lifetime net reproduction rate (NRR), and each of its components (fertility, fecundity and lamb survival), between each of the NRR quartiles. The difference in net reproduction between the ewes in the top quartile compared with those in the bottom quartile was equivalent to an additional lamb per ewe annually. These results identify reproductive levels that could be achieved, establish a basis for selection to improve performance and provide a means to segment the breeding flock for differential management.

INTRODUCTION

The average reproductive performance of Australian specialist sheep enterprises over the period 1977-2007 was 76.6 lambs marked /100 ewes joined but increasing at only 0.04% units annually (ABARE, 2008). The slow rate of progress is despite the availability of both genetic (Purvis *et al.* 1987) and management options (Langford *et al.* 2004) to improve reproductive rates.

There is a genetic influence on each component but heritability estimates are low (Safari *et al.* 2007), in part due to the nature of the annual expression of the traits. Variability, though, is quite high for reproductive traits such that within-flock selection is likely to lead to permanent but relatively slow genetic improvement.

Within-year management options to improve reproductive performance include nutritional inputs to improve the flock average condition/liveweight at joining and parturition, leading to higher fertility, fecundity and lamb survival (Langford *et al.* 2004). Managing whole flocks to increase liveweight/condition at joining to improve reproductive performance can be economically marginal, even after accounting for flow-on benefits to dam and progeny wool production (Young 2008). However, if we could identify segments of the flock that will achieve above average responses to higher management inputs then we should improve the benefit:cost ratio.

Knowing the extent of variation in the reproductive performance within the flock would indicate the current potential, and potentially identify flock segments that might respond to differential management. This paper reports the variation in the lifetime (2-6 years of age) reproductive performance within 3 research flocks run at the Agricultural Research Centre, Trangie in central western NSW.

MATERIALS AND METHODS

Lifetime reproductive data (ages 2-6 years) were available from 3 flocks of Merino ewes (D-Flock, C-Flock and QPLU$) run at the Agricultural Research Centre, Trangie. In each of these flocks, lambing and weaning performance of the ewes was routinely recorded. Lambing and weaning data for at least 3 joinings were available for 2430 D-Flock ewes (born 1975-1983), 1819 C-Flock ewes (born 1984-1993) and for 3037 QPLU$ ewes (born 1993-2002). Descriptions of flock structure and management have been provided elsewhere for D-Flock (Mortimer and Atkins 1989), C-Flock (Mortimer *et al.* 1994), and QPLU$ (Taylor and Atkins 1997).

Statistical analyses. Data on the number of times each ewe was joined, the number of years the ewe lambed, the total number of lambs born and the number of lambs weaned were obtained over the reproductive life (2-6 years of age) of the ewes in the flock. From these values lifetime fertility (no. times ewe lambed/no. joinings), fecundity (no. lambs born/ no. times ewe lambed), lamb survival (no. lambs weaned/no. lamb born) and net reproduction rate (no. lambs weaned/no. joinings) were calculated for each ewe, giving one lifetime record per ewe.

The term genotype in these analyses represents bloodline (D-Flock), animals with the same proportion of genes from each bloodline (C-Flock) and selection line within bloodline (QPLU$).

Adjustments to lifetime net reproduction for genotype and year of birth were made using ASReml (Gilmour *et al.* 2002) and the residual values used to rank individuals. For each quartile of the distribution of adjusted lifetime net reproduction rate (NRR), the adjusted mean NRR and each of its component traits were estimated using ASReml (Gilmour *et al.* 2002), fitting the effects of genotype, year of birth and NRR quartile.

RESULTS

The mean reproductive performances for each of the component traits and NRR of each flock are shown for each of the performance quartiles in Table 1. Across flocks, the differences in NRR between each of the quartiles were of a similar magnitude. Within each flock, the mean NRR was significantly ($P<0.001$) different between each of the quartiles, with the difference between the lowest and highest quartiles being 0.99 (D-Flock), 1.03 (C-Flock), and 1.19 (QPLU$) lambs weaned annually.

Table 1. Lifetime reproductive performance and its components (adjusted for genotype and year of birth effects) for each quartile ranked on net reproduction of Merino ewes from three different flocks

		1st quartile	2nd quartile	3rd quartile	4th quartile	sed
Fertility	D-Flock	0.481 a	0.713 b	0.820 c	0.907 d	0.010
	C-Flock	0.652 a	0.857 b	0.928 c	0.975 d	0.011
	QPLU$	0.542 a	0.752 b	0.853 c	0.949 d	0.011
Fecundity	D-Flock	1.259 a	1.297 a	1.370 b	1.562 c	0.018
	C-Flock	1.262 a	1.301 a	1.413 b	1.647 c	0.019
	QPLU$	1.310 a	1.396 b	1.497 c	1.693 d	0.018
Survival	D-Flock	0.475 a	0.733 b	0.812 c	0.880 d	0.012
	C-Flock	0.530 a	0.773 b	0.849 c	0.914 d	0.013
	QPLU$	0.401 a	0.716 b	0.821 c	0.906 d	0.011
Net reproduction	D-Flock	0.251 a	0.629 b	0.875 c	1.244 d	0.008
	C-Flock	0.407 a	0.810 b	1.064 c	1.431 d	0.010
	QPLU$	0.246 a	0.689 b	1.002 c	1.433 d	0.008

abcd - means within rows with the same subscript are not significantly different ($P>0.05$)

The differences in both fertility and survival between each of the quartiles were significant ($P<0.001$) in each of the flocks. Over 90% of ewes within the highest performing quartile were fertile compared to only 48 to 65% within the lowest performing quartiles of each flock. Of the lambs born to ewes in the highest quartile, 88 to 91% survived compared with only 40 to 53% of lambs born to the poorest performing ewes, although the former had more than twice as many multiple births. In each of the flocks, twice as many ewes in the highest quartile had multiple

births compared with the poorest performing ewes, and represented more than half the ewes lambing within the highest performing ewes.

DISCUSSION

Differences in net reproduction between the bottom and top quartiles of the order of 3.5-6.5 times demonstrate the large degree of variation between ewes over their lifetime in the reproductive performance within Merino flocks and that achievable reproduction rates by Merino ewes are much higher than current expectations based on whole flock means. Ewes from the highest quartile were each producing at least 1 lamb annually more than ewes from the bottom quartile, which on average only lambed every second year and when they did lamb only reared half of their lambs. The top ewes were also able to rear 90% of the lambs born despite having significantly more multiple births (55-69% of ewes lambing).

At least three opportunities are available to take advantage of the large within-flock variation to improve reproductive rates using strategies targeted at different segments of the breeding flock.

Firstly, the influence of the highly productive ewes can be increased by retaining them for an additional year or two beyond the normal cast-for-age. The ewes in the top NRR quartile in these flocks produced 41% of the lambs weaned and there may be production advantages in keeping these ewes to older ages. However, further information on the implications for other production traits of retaining these ewes for longer, and their capacity to continue reproducing (at levels higher than likely replacement ewes) at older ages will be required. This strategy would increase overall performance of the breeding flock through both their higher mean performance and the effect of their retention on flock structure, i.e. reducing the number maiden replacements required.

Secondly, remove ewes with low reproduction from the breeding flock. Only 8% of lambs were produced by the bottom quartile, so removing these ewes from the breeding flock at an early age can substantially improve the average reproductive performance of ewes within the selected age groups. The average reproductive rate of these ewes is so low that over the 5 breeding cycles they would not produce sufficient ewe lambs to replace themselves in the breeding flock. Regardless of their genetic merit in other production traits, removal of these animals from the breeding flock will have little effect on genetic progress in those traits. These ewes might be retained as wool producers, depending on their wool production potential. However, while removing these animals from the breeding flock will increase NRR of the ewes retained, the actual response achieved in the whole breeding flock will depend on (1) the age of culling, (2) the difference in NRR between the culls and replacement ewes and (3) the proportion of replacement maiden ewes.

Thirdly, target management interventions to those flock segments most likely to produce the largest economic responses. While management interventions at the whole flock level to increase average liveweight/condition at joining can have significant economic benefits (Young 2007) in some situations, in others they are economically marginal, even after accounting for flow-on benefits to dam and progeny wool production (Young 2008). Advocates of managing the whole flock to increase liveweight/body condition to improve conception rates and/or survival (Behrendt *et al.* 2006a) appear to assume (in the absence of data to the contrary) that all animals have an equal chance of responding to additional inputs. Little if any attention has been directed to within-flock variability in responses to improvements in, for example, liveweight/body condition. The exception is the management in late pregnancy of ewes based on litter size (Behrendt *et al.* 2006b), acknowledging the relative needs of twin-bearing and single-bearing ewes. Given the extent of variation presented, it appears likely that the requirement for management inputs, and the timing of those inputs, could vary between different segments. Differentially managing flock segments would direct inputs only to those segments with the greatest potential to respond, reducing the total inputs and costs. For example, responses among the poorest performing ewes may not be sufficient

to cover input costs, while the highest performing ewes may already be close to attaining their potential. Additional research is required to test these hypotheses.

The second and third of these strategies require early prediction of ewe lifetime reproductive performance. An earlier report (Lee and Atkins 1996) based on D-Flock ewes suggested that reproductive performance in early life (2- or 3-years old) could predict reproductive performance in later life, although the accuracy of that prediction was increased if the information from the two years was combined. Subsequent analyses of C-Flock and QPLU$ data have confirmed those observations (G.J. Lee, K.D. Atkins and M.A. Sladek, unpublished data) indicating early identification would be feasible.

To obtain lifetime records for individual animals will require permanent individual identification, individual measurement and an appropriate system for recording data. The increasing use of electronic identification of ewes and on-farm scanning services has made the collection of information on individuals less labour intensive and reduced the costs (Pope and Atkins 2008). The availability of lifetime records can increase the accuracy of selection of young animals by including the dam's reproductive performance as a selection criterion in the index (Lee *et al.*, 2009).

ACKNOWLEDGEMENTS

The authors wish to acknowledge the many staff members of the Agricultural Research Centre, Trangie who assisted with data collection and database management.

REFERENCES

ABARE (2008) Farm survey data for the beef, slaughter lambs and sheep industries. *http://www.abare.gov.au/ame/mla/mla.asp*

Behrendt, R., Barber, P., Oldham, C.M., Hocking Edwards, J.E., Hatcher, S. and Thompson, A. (2006a) *Proc. Aust. Soc. Anim. Prod.* **26**: Communication 70.

Behrendt, R., Barber, P., Oldham, C.M., Hocking Edwards, J.E., Hatcher, S. and Thompson, A. (2006b) *Proc. Aust. Soc. Anim. Prod.* **26**: Communication 91.

Gilmour, A.R., Gogel, B.J., Cullis, B.R., Welham, S.J. and Thompson, R. (2002) "ASReml User Guide Release 1.0.", VSN International, Hemel Hempstead, UK.

Langford, C., Alcock, D., Holst, P., Shands, C. and Casburn, G. (2004) "Wean More Lambs: Optimising sheep reproductive performance", Meat and Livestock Australia, North Sydney.

Lee, G.J. and Atkins, K.D. (1996) *Aust. J. Expt. Agric.* **36**:123.

Lee, G.J., Atkins, K.D. and Sladek, M.A. (2009) *Proc. Assoc. Advmt. Anim. Breed. Genet.* **18:** t*hese proceedings.*

Mortimer, S.I. and Atkins, K.D. (1989) *Aust. J. Agric. Res.* **40**:433.

Mortimer, S.I., Atkins, K.D., Eissen, J., Van Heelsum, A., Burns, A.M. and Isaac, B.R. (1994) *Wool Tech. Sheep Breed.* **42**:243.

Pope, C.E. and Atkins, K.D. (2008) *Proc. Aust. Soc. Anim. Prod.* **27:**23.

Purvis, I.W., Atkins, K.D. and Piper, L.R. (1987) In 'Merino improvement programs in Australia', pp. 229-242, editor B.J. McGuirk, Australian Wool Corporation, Melbourne.

Safari, E., Fogarty, N.M., Gilmour, A.R., Atkins, K.D., Mortimer, S.I., Swan, A.A., Brien, F.D., Greeff, J.C. and van der Werf, J.H.J. (2007) *Aust. J. Agric. Res.* **58**:177.

Taylor, P.J. and Atkins, K.D. (1997) *Wool Tech. Sheep Breed.* **45**:92.

Young, J.M. (2007) Implications of Lifetimewool for on-farm management in south west Victoria. *http://www.lifetimewool.com.au/pdf/highrainfallzreport.pdf*

Young, J.M. (2008) Implications of Lifetimewool for on-farm management on the southern slopes (southern NSW & central Vic). *http://www.lifetimewool.com.au/pdf/MIDASsouthernslopes.pdf*

GENETIC PARAMETERS FOR LIFETIME REPRODUCTIVE PERFORMANCE OF MERINO EWES

G.J. Lee, K.D. Atkins and M.A. Sladek

Cooperative Research Centre for Sheep Industry Innovation, NSW Department of Primary Industries, Orange Agricultural Institute, Forest Road, Orange, NSW 2800, Australia

SUMMARY

The genetic variation between ewes in lifetime (2-6 years of age) net reproductive performance, and its component traits, was estimated in each of 3 Merino research flocks. The heritability estimates for each of the lifetime reproductive traits were higher than published estimates based on single year records, and in line with expectation for traits with a repeatability of about 0.15. Therefore, a repeatability model for analysing reproductive performance seems adequate. The lifetime component traits (fertility, fecundity and survival) each had high positive genetic correlations (≥ 0.55) with net lifetime reproductive performance, indicating that all components contribute to genetic improvement in net reproductive performance.

INTRODUCTION

Reproduction rates in Australian sheep flocks are low nationally and have changed little in the last 30 years (ABARE 2008) despite the availability of genetic (Purvis *et al.* 1987) and within-year management options (Langford *et al.* 2004; Kleemann *et al.* 2006) to improve reproductive performance.

Improving reproductive performance increases productivity of the breeding ewe unit, more efficient use is made of available feed (proportionately less of the feed consumed is used for maintenance), more surplus animals are available for sale and greater selection pressure is available, increasing the potential for genetic gains. This can lead to increased profit ($/ha) across the range of sheep breeding enterprises (Langford *et al.* 2004).

While there is a genetic influence on each component, heritability estimates are low (Safari *et al.* 2007b) and this in part is due to the nature of the annual expression of the traits. Using the high variability in annual records (Safari *et al.* 2005) for within-flock selection is likely to lead to permanent but relatively slow genetic improvement. Environmental sources (nutrition, management, health etc) contribute to the large variability in reproductive performance both between-ewes within years and within-ewe between years. Over a lifetime, the between-ewe variation in reproductive performance remains large (Lee *et al.* 2009), so combining a number of annual performances should improve precision in estimating the merit of individuals and potentially improve the rate of genetic improvement.

This study estimates the extent of genetic variation in the lifetime (2-6 years of age) reproductive performance of Merino ewes within 3 research flocks maintained at the Agricultural Research Centre, Trangie in central western NSW.

METHODS

Lifetime reproductive data (ages 2-6) were available from 3 flocks of Merino ewes (D-flock, C-flock and QPLU$) run at the Agricultural Research Centre, Trangie. In each of these flocks, lambing and weaning performance of the ewes were routinely recorded. Lambing and weaning data for at least 3 joinings were available for 2430 D-flock ewes (born 1975-1983), 1819 C-Flock ewes (born 1984-1993) and for 3037 QPLU$ ewes (born 1993-2002). Descriptions of flock structure and management have been provided elsewhere for D-flock (Mortimer and Atkins 1989), C-flock (Mortimer *et al.* 1994), and QPLU$ (Taylor and Atkins 1997).

Statistical analyses. Data on the number of times each ewe was joined, the number of years the ewe lambed, the total number of lambs born, and the number of lambs weaned were obtained over the reproductive life (2-6 years of age) of the ewes in the flock. From these values lifetime fertility (ewes lambing/ewe joined), fecundity (lambs born/ewe lambing), lamb survival (lambs weaned/lamb born) and net reproduction (lambs weaned /ewe joined) were calculated (Lee *et al.* 2009).

Variance and co-variance components were estimated for lifetime net reproduction and the reproduction component traits (fertility, fecundity and lamb survival) using a mixed animal model within ASReml (Gilmour *et al.* 2002), fitting year of birth and genotype effects in univariate and bivariate analyses.

RESULTS AND DISCUSSION

Phenotypic variation. The means of net reproductive performance, and its component traits, over the ewes' lifetime and the phenotypic variation are shown in Table 1. Within-flock phenotypic variation in each of the component traits was less than that in net reproduction. Among the component traits, phenotypic variation was least for fecundity, while survival was the most variable.

As expected, the phenotypic coefficients of variation observed for lifetime records of fertility, fecundity and net reproduction rate were substantially less (33-46%) than that observed by Safari *et al*. (2007a) for the annual records from these same flocks. Assuming equal variances across age expressions and unity genetic correlations between age expressions, the phenotypic variance for a mean (lifetime) trait compared with the phenotypic variance of a single age expression will be:

$$\sigma^2_{\mu} = \{ 1 + (n - 1) t \} / n * \sigma^2_x \quad \ldots\ldots\ldots (1)$$

where σ^2_{μ} = variance of lifetime mean
σ^2_x = variance of single record
t = repeatability
n = number of records

Using n=4 and t=0.15 in equation (1) above shows that the phenotypic variance will be 0.36 of the annual record phenotypic variance or 0.6 (√0.36) for the standard deviation or coefficient of variation. The reduction in coefficient of variation of 33-46% is in broad agreement with the 40% reduction expected for the simple case above.

Table 1. Mean and coefficient of variation of lifetime net reproductive performance and its component traits in 3 flocks of Merino ewes

	D-flock		C-flock		QPLU$	
	mean	CV %	mean	CV %	mean	CV %
Fertility – ewes lambing/ewe joined	0.725	32.8	0.852	24.4	0.771	31.7
Fecundity – lambs born/ ewe lambing	1.374	24.3	1.398	22.5	1.471	24.4
Survival – lambs weaned/lamb born	0.732	34.7	0.770	30.5	0.714	39.6
Net reproduction - lambs weaned/ ewe joined	0.739	52.9	0.929	43.4	0.825	56.2

Heritability. Estimates of the heritability of lifetime net reproduction ranged from 0.108 to 0.193 across the 3 flocks. Both fertility and lamb survival heritability estimates had the largest range

between the flocks (0.012 to 0.196 and 0.036 to 0.195, respectively). Heritability estimates for lifetime fecundity were the most consistent between the flocks and indicated moderate levels (0.19 to 0.26, ± 0.04-0.05) of genetic variation.

Table 2. Heritability (diagonal, bold), genetic (below diagonal) and phenotypic (above diagonal) correlations (±*se*) for lifetime net reproductive performance and each of its components in Merino ewes from 3 flocks

		Fertility	Fecundity	Survival	Net
D-Flock	Fertility	**0.196**	0.047	0.192	0.702
		0.043	*0.023*	*0.021*	*0.011*
	Fecundity	-0.024	**0.191**	-0.132	0.390
		0.169	***0.045***	*0.021*	*0.018*
	Survival	0.791	0.173	**0.036**	0.630
		0.349	*0.381*	***0.038***	*0.013*
	Net	0.811	0.550	0.776	**0.172**
		0.072	*0.136*	*0.171*	***0.043***
C-Flock	Fertility	**0.012**	0.122	0.200	0.644
		0.032	*0.025*	*0.024*	*0.014*
	Fecundity	0.628	**0.260**	-0.074	0.496
		0.804	***0.052***	*0.024*	*0.019*
	Survival	1.032	-0.389	**0.078**	0.666
		0.895	*0.227*	***0.043***	*0.014*
	Net	1.070	0.687	0.614	**0.108**
		0.578	*0.136*	*0.183*	***0.043***
QPLU$	Fertility	**0.135**	0.125	0.162	0.645
		0.036	*0.020*	*0.020*	*0.011*
	Fecundity	0.436	**0.196**	-0.053	0.430
		0.150	***0.037***	*0.019*	*0.016*
	Survival	0.274	0.139	**0.195**	0.690
		0.163	*0.145*	***0.040***	*0.010*
	Net	0.667	0.638	0.782	**0.193**
		0.094	*0.099*	*0.062*	***0.036***

The average estimate of heritability for lifetime net reproduction, 0.16 across the 3 flocks, was higher than the estimate derived from annual records of these flocks (0.05, Safari *et al*. 2007b). Similarly, the average lifetime fecundity heritability of 0.21 was higher than the pooled estimate of annual records for the 3 flocks (0.074, Safari *et al.* 2007b). Given the assumptions of a repeatability model, the expected heritability for the mean (lifetime) trait compared with the heritability of a single age expression will be:

$$h^2_{\mu} = n / \{ 1 + (n - 1) t \} * h^2_x \quad \ldots\ldots\ldots (2)$$

where h^2_{μ} = heritability of lifetime mean
h^2_x = heritability of single record

For n=4 and t=0.15 in equation (2), the expected heritability of the lifetime trait will be 2.8 times that of the single record. Increases in heritability of 3.2 and 2.9 times for net reproduction rate and fecundity respectively, lend further support to the adequacy of the repeatability model for reproductive traits.

Phenotypic and genetic correlations. The phenotypic correlations (Table 2) of net reproductive performance with each of the component traits were moderate to high and positive, but for each flock were lowest for the relationship of fecundity with net reproduction (0.39 to 0.50). Phenotypic correlations of fertility with each of the other component traits were positive but low (<0.2), while survival and fecundity had a low negative correlation (-0.05 to -0.13) within each flock.

Across all flocks, the genetic correlations of each of the component traits with lifetime net reproduction were positive and high (≥ 0.55). However, the genetic correlations among the component traits were imprecise and/or inconsistent between the flocks.

CONCLUSION

Collecting and updating reproductive performance over the lifetime, as opposed to using annual records, will have the benefits in the current generation through improving the accuracy of selection for ewes to remain in the breeding flock, and will substantially improve the accuracy of selection of young animals (particularly rams in ram breeding flocks and ewes in commercial flocks) by including the dam's reproductive performance as a selection criterion in the index. Further, given the size of the genetic correlations of net reproduction with the component traits, using lifetime net reproductive performance will achieve the best overall response in reproduction rate rather than just focussing on fertility (wet-dry data), litter size (pregnancy scanning) or lamb survival alone.

ACKNOWLEDGEMENTS

The authors wish to acknowledge the many staff members of the Agricultural Research Centre, Trangie who assisted with data collection and database management. Support for the QPLU$ project was provided by Australian Wool Innovation, while the study reported was supported by the Cooperative Research Centre for Sheep Industry Innovation.

REFERENCES

ABARE (2008) Farm survey data for the beef, slaughter lambs and sheep industries. *http://www.abare.gov.au/ame/mla/mla.asp*

Gilmour, A.R., Gogel, B.J., Cullis, B.R., Welham, S.J. and Thompson, R. (2002) 'ASReml User Guide Release 1.0.', VSN International, Hemel Hempstead, UK.

Kleemann, D.O., Grosser, T.I. and Walker, S.K. (2006) *Theriogen* **65**: 1649.

Langford, C., Alcock, D., Holst, P., Shands, C. and Casburn, G. (2004) 'Wean More Lambs Optimising sheep reproductive performance.', Meat and Livestock Australia, North Sydney.

Lee, G.J., Atkins, K.D. and Sladek, M.A. (2009) *Anim. Prod. Sci.* **49**: 624.

Mortimer, S.I. and Atkins, K.D. (1989) *Aust. J. Agric. Res.* **40**: 433.

Mortimer, S.I., Atkins, K.D., Eissen, J., Van Heelsum, A., Burns, A.M. and Isaac, B.R. (1994) *Wool Tech. Sheep Breed.* **42**: 243.

Purvis, I.W., Atkins, K.D. and Piper, L.R. (1987) In 'Merino improvement programs in Australia', pp. 229-242, editor B.J. McGuirk, Australian Wool Corporation, Melbourne.

Safari, E., Fogarty, N.M. and Gilmour, A.R. (2005) *Livest. Prod. Sci.* **92**: 271.

Safari, E., Fogarty, N.M., Gilmour, A.R., Atkins, K.D., Mortimer, S.I., Swan, A.A., Brien, F.D., Greeff, J.C. and van der Werf, J.H.J. (2007a) *Aust. J. Agric. Res.* **58** 169.

Safari, E., Fogarty, N.M., Gilmour, A.R., Atkins, K.D., Mortimer, S.I., Swan, A.A., Brien, F.D., Greeff, J.C. and van der Werf, J.H.J. (2007b) *Aust. J. Agric. Res.* **58**: 177.

Taylor, P.J. and Atkins, K.D. (1997) *Wool Tech. Sheep Breed.* **45**: 92.

LEVELS OF POST- WEANING LOSS IN THE TRANGIE D-FLOCK (1975 -1983)

C. E. Pope and K. D. Atkins

NSW Department of Primary Industries, Locked Bag 21,Orange NSW 2800

SUMMARY

Survival of Merinos from weaning to 9 months of age was analysed in the multiple bloodline flock (D-Flock) maintained at Trangie between 1975 and 1983. Average post-weaning loss was 8.6%, but varied between years from 4.2% to 18.0%. Birth type, age of dam and sex had very small influences on the probability of survival post-weaning. Weight at weaning had a curvilinear relationship with weaner survival such that lighter animals had reduced survival while heavier animals showed little survival advantage over average weight animals. However, this relationship varied across years and was not a significant source of variation in several of the years observed.

INTRODUCTION

The average reproductive performance of Merino sheep in Australia is low by national and international standards. The overall reproductive rate (measured as lambs weaned per ewe joined) can be broken into its component traits of fertility, litter size and lamb survival. Lamb survival is generally limited to survival to weaning, but it can be extended to include post-weaning survival.

Sackett *et al.* (2006) estimated weaner mortality to cost about $89M annually while Lloyd, (Lloyd J unpublished) estimated an annual cost of $75M with an assumed 8% loss in sheep cereal and pastoral zones and a 10% loss in the high rainfall zone. Unpublished work from Angus Campbell of the Mackinnon Project estimated that reducing post-weaner mortality from 12% to 4% per annum would improve net farm profit by 80c/DSE across the whole farm.

Post weaning loss is not well enumerated and potentially represents a significant loss in genetic and financial terms. This study aimed to investigate the extent of post weaning losses and the extent to which factors such as weaning weight and age, birth type, sex and dam age influence weaner survival. The data spanned the period 1975-83 in a multiple bloodline Merino flock. Given the lack of attention that post-weaning survival has received in the past, it is highly likely there has been little change in the performance of Merinos since that time.

MATERIALS AND METHODS

Lambing and survival data for 14,187 animals was collected between 1975 and 1983 from the D-flock located at Trangie, NSW. The D-flock was based on fourteen bloodlines that were representative of the numerical importance of these Merino bloodlines within New South Wales at that time. A fifteenth flock, was similarly formed, and was drawn from the Fertility flock selection line which had been maintained at Trangie since 1959 (Atkins and Robards 1976).

The sheep grazed on largely natural grass pastures and some lucerne. Lambing occurred in July and lambs were weaned at an average age of 120 days (November). Lambs were weighed in May (approx 6 months post-weaning) to determine post-weaning growth and survival. A full description of the flocks and their management is provided by Mortimer and Atkins (1987).

Data were analysed in two stages with a generalized linear model using ASREML software (Butler *et al.* 2007). A logit transformation was used to account for the binary nature of survival. The initial model examined the survival variables lambs marked, lambs weaned and lambs surviving to the May weighing and included terms for birth year, sex, dam age, birth type and flock and any significant interactions. A further analysis of post-weaning survival included only year and flock with covariates for weaning age (linear) and weaning weight (linear and quadratic functions).

RESULTS AND DISCUSSION

What are the losses over time? Across all years, 21.6% of all lambs born died prior to lamb marking, an additional 5.9% died between lamb marking and weaning and 6.1% were lost post-weaning up to 9 months of age. Overall the average total loss was 33.6% between birth and 9 months of age (range 25-42% across years). While there was considerable variation between years, the pattern in the timing of losses across years was reasonably consistent.

The average losses post-weaning, expressed as a percentage of lambs weaned was 8.6%. This is considerably lower than the 12-14% post-weaning losses reported by Campbell and Behrendt (unpublished) in small scale research in Victoria (range 4.5-27.1%) and recent anecdotal reports from the Yass District in NSW indicating that producers were losing up to 25% of lambs post-weaning (R.P. Graham, 2005, pers. com.). The greatest loss in these data occurred in the 1975 birth cohort with 18% of animals dying between weaning and 9 months. Much of this loss occurred during a severe period of flystrike during early autumn (Atkins and McGuirk 1979). The range in losses in the other years was between 4% and 11% with an average of 7.5%.

Is there a difference between the survival of twins and singles? On average the pre-weaning survival rate of single born lambs was 79.8% compared to 66.8% for multiple births (Figure 1). Post-weaning survival, while not significantly different among birth types, reversed this trend. The survival of single males post-weaning was 90.2% compared with 92.4% for single females. For twins, the figures were 91.7% and 90.1% respectively, with a standard error of about 0.65% for each subclass. Only in the unusual year of the 1975 cohort, when overall losses were high, did the losses among multiples exceed that of singles. In that specific year, 12% of singles did not survive after weaning compared to 24% of twins.

a) b)

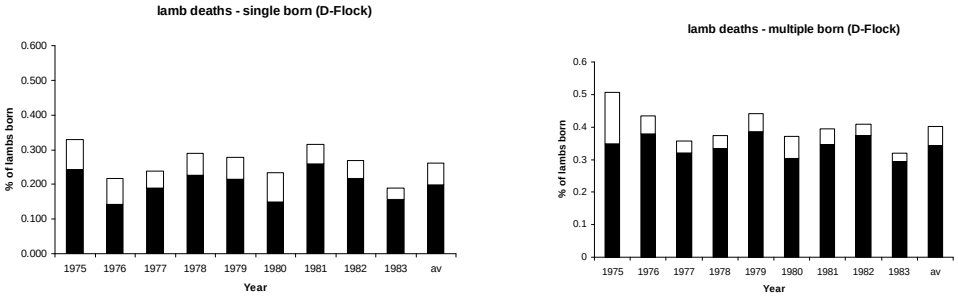


Figure 1. Effect of year on annual lamb and weaner deaths within (a) single and (b) multiple-born lambs. (The solid bars indicate lamb loss pre-weaning, the open bars indicate lambs lost between weaning and May (9 months of age) [predicted means from ASREML analysis?]

Effect of dam age: Pre-weaning survival was lower among progeny of 2-year old dams (69.8%) but little different among progeny of older age groups (73.0% to 74.2%). For post-weaning survival, the pattern with age of dam was somewhat different. The progeny of 2-4 year old dams had an average survival of 92.5% while for 5- and 6-year old dams, the survival rates were only 90.5% and 89.3% respectively. Standard errors for individual dam ages were less than 1%.

Effect of weaning age: Post-weaning survival improved with increasing weaning age. Over the range from 90 to 130 days, the corresponding survival rate increased from 89.4% to 94.5%.

Weaning weight: Figure 2 shows that weight at weaning has a curvilinear relationship with weaner survival such that lighter animals had a reduced survival. However, this relationship varied significantly across years and was not a significant source of variation in several of the years observed. During 1983, for example, the relationship was slightly inverted. The point of inflection for the other years varied over a range of about 12 kgs.

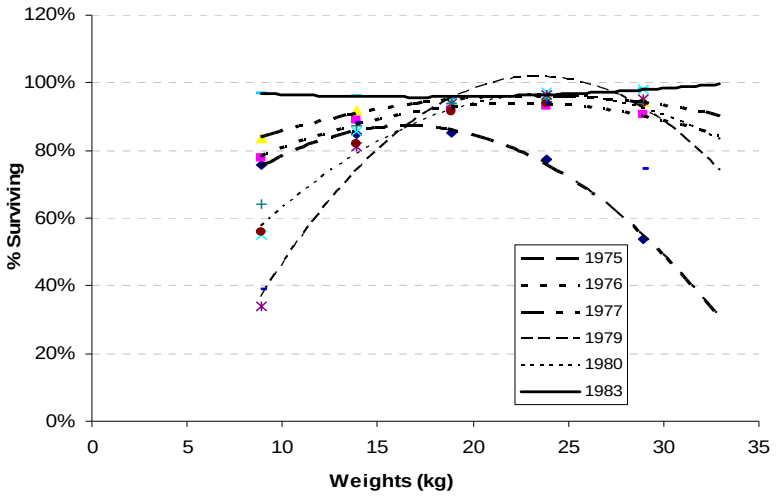


Figure 2. Relationship between weaning weight and postweaning survival, predicted from the weaning weight X year interaction

Weaning weight and flock interactions. Table 1 shows post-weaning survival at different weaning weights for each of the sub-flocks contributing to the D-Flock. The lighter lambs were less likely to survive. However the number sampled at the lower end of the distribution is small (7% vary from the mean by -10kgs or more, 0.2.% vary by +15kgs or more). The mean weight of the bottom 20% of the population was 14.1 kgs with a survival of 77.7% (20% of lambs were less than 16kg). The mean weaning weight of the heaviest 80% was 18.9 $\pm$1.2 kgs with a survival of 92.6%.

One recommendation to improve post-weaner survival is to draft off the lightweight animals at weaning and to provide them with differential feeding and management. The data presented here suggests that such a practice is targeting a small number of animals, not all of which can be expected to respond to nutritional stimuli.

IMPLICATIONS

This analysis of the D-Flock data from Trangie has shown that information relating to post-weaning survival in one region will not necessarily apply elsewhere. This analysis has found:

- **Birth type** had a very small influence on the probability of survival post-weaning.
- **Age of dam** had a very small influence on the probability of survival post-weaning.
- **Sex** had a very small influence on the probability of survival post-weaning.

- **Weight at weaning** had a curvilinear relationship with weaner survival such that lighter animals had a reduced survival. However, this relationship varied significantly across years and was not a significant source of variation in several of the years observed.
- **Economically** focussing on increasing the weight of the lightest 20% of animals is unlikely to be a very cost-effective strategy to increase sheep numbers in the Trangie environment reported here. More analyses from other environments need to be undertaken before global strategies are developed.

Table 1. The effect of weaning weight on weaner survival (%) within each of 15 bloodlines across years

				Weight (kg)				
Flock	Description	mean	n	10	15	20	25	30
36	Fine wool Saxon	16.5	374	84	91	96	99	100
38	Fine wool Saxon	16.1	270	76	87	96	103	107
45	Medium wool Peppin	20.3	1499	86	92	95	96	95
33	Medium wool Peppin	19.1	595	81	89	95	98	100
32	Medium wool Peppin	18.2	672	83	92	98	102	104
34	Medium wool Peppin	18.7	672	85	92	97	100	100
41	Medium wool Peppin	18.1	499	72	81	88	93	95
39	Medium wool Peppin	17.9	498	86	92	96	97	97
43	Medium wool Peppin	19.5	532	81	89	95	99	100
44	Medium wool Peppin	18.8	570	84	90	94	95	94
42	Medium wool Peppin	17.9	583	86	92	96	98	98
37	Medium wool Peppin	18.7	416	79	86	92	95	96
31	Medium wool non-Peppin	19.3	710	86	92	96	97	96
35	Medium wool non-Peppin	18.8	624	80	89	96	101	104
40	Strong wool South Australian	19.8	470	80	88	93	96	97
	Whole Flock	18.8		80	90	95	98	97
	Standard errors			1%	.4%	.4%	.6%	.1.3%

ACKNOWLEDGEMENTS

The authors would like to acknowledge the support of the Sheep CRC and the many DPI staff who have been involved in data collection at Trangie ARC.

REFERENCES

Atkins,K.D. and McGuirk,B.J.(1979). *Wool Technol. Sheep Breed.* **27:**15.

Atkins, K.D. and Robards, G.E.(1976) *Aust. J. Exp.Agr. Anim.Husb*. **16:**315.

Butler,D., Cullis,B., Gilmour,A. and Gogel.B (2007) ASReml-R reference manual, The State of Queensland, Dept of Primary Industries and Fisheries

Lloyd J., Increased lamb and Weaner Survival - Review of the Economic Case for Change (CRC Report)

Mortimer, S.I. and Atkins, K.D. (1989) *Aust J Agric. Res.* **40:**433.

Sackett, D. M., Holmes, P., Abbot, K., Jephcott, S., and Barber, M. (2006) Assessing the economic cost of endemic disease on the profitability of Australian beef cattle and sheep producers. Meat and Livestock Australia Limited, North Sydney

GENETIC VARIATION IN CLEAN WOOL COLOUR IN FINE WOOL MERINOS

J.L. Smith and I.W. Purvis

FD McMaster Laboratory, CSIRO Livestock Industries, Armidale NSW 2350

SUMMARY

Greasy wool colour is a component of wool style and a determinant of price received, but it is clean wool colour that determines potential colour range and dye-ability of wool garments. Selective breeding is a means of improving raw wool clean colour, but its effectiveness is currently unclear due to limited existing information on clean colour variability, heritability and correlations with other fleece traits. Clean colour was measured on yearling (10mths, n=7106) and adult (21mths, n=8078) animals. Other fleece traits, including assessed greasy wool colour, were also recorded and genetic parameters for clean and greasy wool colour estimated. Clean wool colour was moderately to highly heritable at both yearling and adult ages (0.25 and 0.29 respectively), as was greasy colour (0.33 and 0.40). Clean wool colour at yearling and adult age was moderately genetically correlated (0.52), suggesting that selection using yearling measurements would be effective for improving lifetime clean wool colour. Greasy and clean wool colour were poorly correlated phenotypically and genetically at both yearling and adult age. Of the measured fleece traits, fibre diameter and curvature were those with highest genetic correlations with clean wool colour (yearling, 0.52, -0.34). These correlations indicate that superfine wool animals with a classical crimp frequency to fibre diameter ratio (true-to-type) tend to have superior clean wool colour compared to other wool types included here.

INTRODUCTION

Greasy wool colour impacts on price received for wool at auction. With potential development of new markets in trans-seasonal apparel, selective breeding for clean colour is a means of increasing the quantity of white wool supplied to the market. There are many estimates of greasy wool colour genetic parameters for Australian Merinos in the literature (Mortimer 2007) and stud breeders routinely rely on subjective assessment of greasy colour rather than objective measurement of clean colour (Brown 2006). However phenotypically, the relationship between greasy and clean wool colour is poor. There is some evidence though from genetic analysis of other Merino resource flocks (Dowling *et al.* 2007; Hebart and Brien 2009) indicating moderate genetic correlation between greasy and clean wool colour which suggests that selection for clean wool colour based on greasy colour could be effective.

Using data from CSIRO fine wool resource flocks, the aims of this study were to estimate: the variability and heritability of greasy and clean wool colour; phenotypic and genetic correlations between greasy and clean colour and a suite of fleece and production traits; and phenotypic and genetic correlations between clean colour at yearling and adult age. These parameters are necessary for inclusion of clean wool colour in Merino breeding programs.

METHOD

Animals. Measured clean wool colour (cCOL, yellowness Y-Z) and assessed greasy wool colour (gCOL) were recorded on fully pedigreed yearling (10mth) and adult (2yo) sheep in the CSIRO Fine Wool Project (FWP) (Swan *et al.* 2008) and the CSIRO Toward 13 Micron Flock (T13) (Swan and Purvis 2005) (Table 1). Of the 12 bloodlines represented, there were 9 fine/superfine lines and 2 medium lines, plus the T13 ultrafine wool line which was founded on ewes from the fine/superfine lines in the FWP.

Table 1. Sire representation among yearling and adult clean wool colour (cCOL) records

	Animals	Sires	cCOL	Mean progeny group size (range)
Yearling	11856	473	7106	25.1 (1-173)
Adult	8226	416	8078	19.5 (1-106)
Yearling & Adult	8826	296	5598	18.9 (3-87)

Wool colour measurements. cCOL measurements were made by the Australian Wool Testing Authority in CIE tristimulus values as Y-Z (yellowness). gCOL was assessed on a 1-5 scale where 1=whitest. Swan *et al*. (1997) reported a strong association between assessed and measured greasy wool colour, indicating subjective assessment is an accurate method to determine greasy wool colour.

Other measurements. In addition to cCOL and gCOL, traits included in this analysis were greasy and clean fleece weight (GFW and CFW); clean scoured yield (CSY); mean fibre diameter (MFD); standard deviation and coefficient of variation of fibre diameter (SDFD and CVD); staple length and strength (SL and SS); mean fibre curvature (CURV); and bodyweight (BWT). Fleece and bodyweight trait abbreviations are henceforth prefixed with y (yearling) or a (adult) as appropriate.

Statistical analysis. Initially, fixed effects models were fitted using ASReml (Gilmour *et al.* (2002) to determine significance of main effects which included bloodline (12 levels, described above), dam age (maiden and adult), birth-rearing type (born and reared single; born multiple and reared single; born and reared multiple), sex (male and female), flock (combined birth year-management group effect) and age (to first shearing for yearlings and between first and second shearing for adults) which was fitted as a covariate. Scorer of gCOL was also included. Non-significant fixed effects were iteratively removed from the model and no interactions were included in the final models. In analysis of adult records, location (Armidale, Condobolin and Camden) was an additional main effect. For the drops 1991 and 1993-1996 inclusive, approximately ¾ (n=2314, balanced for bloodline and sire) of male progeny from the FWP flock were relocated to the NSW DPI site at Condobolin following hogget shearing for the purpose of determining genotype by environment interactions. In 1992, approximately ¾ (n=396) of the male progeny were re-located to The University of Sydney, Camden, NSW for footrot-related studies. Heritability and correlations among cCOL, gCOL and the suite of fleece and production traits were estimated from bivariate mixed animal models.

RESULTS

Main effects. Bloodlines were significantly different in both gCOL and cCOL and at both ages (Table 2). For ycCOL, the ultrafine bloodline and medium wool bloodlines were significantly different to each other and all the fine/superfine wool bloodlines ($P<0.001$). As adults, the ultrafine line was not significantly different in cCOL to any of the fine wool lines, but the wool from the medium lines remained significantly more yellow than that from the other bloodlines ($P<0.001$).

For ycCOL ($P<0.01$), acCOL ($P<0.001$) and agCOL ($P<0.001$), animals reared as multiples tended to have more yellow wool than those born and reared as singletons. Among adults, males had more yellow wool than females ($P<0.05$). Location was a significant effect on acCOL and agCOL where animals at Camden had more yellow gCOL and less yellow cCOL ($P<0.001$). acCOL and agCOL tended to be more yellow at Condobolin than Armidale ($P<0.001$).

Table 2. Predicted bloodline means for clean (cCOL) and greasy (gCOL) wool colour among yearlings (10mth) and adults (21mth)

		Bloodline											
Age	Trait	1	2	3	4	5	6	7	8	9	10	11	12
Yearling	cCOL	8.2	8.4	8.3	8.3	8.2	8.7	8.1	8.8	8.3	8.1	8.1	7.8
	gCOL	2.3	2.9	2.7	2.4	2.8	3.2	2.1	3.5	2.7	2.2	2.6	2.6
Adult	cCOL	8.3	8.4	8.5	8.4	8.3	9.2	8.4	9.2	8.4	8.3	8.3	8.5
	gCOL	2.1	3.2	2.4	2.3	3.1	3.6	1.8	3.6	2.8	2.1	2.4	2.7

Bloodline 12 = ultrafine wool, 6 and 8 = medium wool, rest = superfine/fine wool
cCOL s.e.'s 0.02 – 0.05; gCOL s.e.'s 0.02 – 0.04

Genetic parameters. The heritability of cCOL was similar, but slightly higher in adults than yearlings (0.29 and 0.25 respectively). For gCOL, the heritability was also higher at adult than yearling age (0.40 and 0.33 respectively), but cCOL and gCOL were poorly correlated both phenotypically and genetically at both ages (Table 3). MFD and CURV were the fleece traits most closely associated with cCOL and those phenotypic and genetic relationships appeared stronger in yearlings than adults (Table 3). Whiter wools tended to have lower MFD and higher CURV.

Table 3. Phenotypic (r_p) and genetic (r_g) correlations among clean wool colour (cCOL, Y-Z) and assessed greasy wool colour (gCOL) and other production traits (s.e. in parentheses) in Yearlings (10mth) and adults (21mth)

	Yearling (10mths)						Adult (21mths)					
			cCOL *x*		gCOL *x*				cCOL *x*		gCOL *x*	
Trait	V_p	h^2	r_p	r_g	r_p	r_g	V_p	h^2	r_p	r_g	r_p	r_g
GFW (kg)	0.11	0.35	0.04	0.12	0.10	0.26	0.20	0.40	-0.01	0.09	0.08	0.25
CFW (kg)	0.07	0.34	0.05	0.14	0.00	0.15	0.13	0.38	0.00	0.15	-0.02	0.07
CSY (%)	15.6	0.55	0.03	0.07	-0.32	-0.45	17.6	0.55	0.00	0.15	-0.26	-0.37
MFD (µm)	1.07	0.68	0.24	0.52	0.10	0.14	1.30	0.69	0.19	0.35	0.06	0.10
SDFD (µm)	0.14	0.50	0.10	0.28	0.14	0.27	0.15	0.52	0.11	0.21	0.10	0.21
CVD (%)	4.7	0.45	-0.03	-0.05	0.10	0.20	3.6	0.44	0.00	0.00	0.08	0.18
SL (mm)	72	0.54	0.04	0.04	-0.04	-0.05	64	0.51	-0.04	-0.03	-0.02	0.05
SS (N/kTex)	89	0.35	0.07	0.04	-0.08	-0.23	83	0.33	0.04	0.12	-0.11	-0.22
CURV (º/mm)	110	0.59	-0.21	-0.34	-0.06	-0.16	109	0.49	-0.13	-0.30	-0.09	-0.24
YBWT (kg)	11.6	0.49	0.02	0.07	-0.02	0.02	20.0	0.50	-0.03	-0.10	-0.03	0.03
cCOL (Y-Z)	0.31	0.25	-	-	0.06	0.13	0.37	0.29	-	-	0.08	0.18
gCOL (1-5)	0.40	0.33	0.06	0.13	-	-	0.52	0.40	0.08	0.18	-	-

s.e on h^2 estimates 0.02-0.04; s.e. on r_p estimates 0.01-0.02; s.e. on r_g estimates 0.04-0.07

The genetic correlation between cCOL at yearling and adult age was moderate (0.52), suggesting that selection using yearling measurements would be reasonably effective for improving clean wool colour of adult animals. gCOL at yearling and adult age was moderately correlated phenotypically (0.40) and highly correlated genetically (0.84). This is of limited practical use for improving clean wool colour due to the poor relationship between the two, but could be a useful selection aid for those wool producers with a focus on wool style, of which greasy colour is a component.

DISCUSSION

The genetic correlation between cCOL at yearling and adult age is moderate, and between gCOL of yearlings and adults is high. That, accompanied by moderate heritabilities and adequate phenotypic variance, bodes well for selection for lifetime greasy or clean colour based on yearling

measurements. However, the genetic correlation between greasy and clean colour at both ages is low at best, which is limiting to the potential for improvement in clean wool colour through indirect selection on greasy wool colour. Clean colour heritability and correlations with key economically important fleece traits are lower here than reported for mainly medium wool types by Hebart and Brien (2009) which may suggest differences among Merino genotypes in clean colour.

At approximately $10 per animal, clean wool colour is an expensive trait to measure and with no clear economic signals to wool producers in improving clean wool colour it is difficult for stud breeders to justify the measurement and inclusion of clean wool colour in breeding programs. Given the genetic correlations between cCOL and MFD and CURV, and the fact that those two traits are more highly heritable than cCOL, a higher rate of response in cCOL may be achievable through indirect selection on MFD and CURV than gCOL.

Benavides *et al.* (1998) argue that the poor association between greasy and clean colour is not surprising as scoured wool colour is dependent upon environmental challenge (Wilkinson 1981). Therefore, selection on cCOL is likely to be ineffective because yellowness is only expressed when environmental conditions are conducive. In effect, this is a genotype by environment interaction. Wilkinson's observation of wool colour deterioration with exposure to heat and humidity led to development of a greasy wool incubation test for prediction of cCOL. Raadsma and Wilkinson (1990) applied that predictive test in genetic studies of wool yellowing. Heritability of the predictive test in Merinos was estimated at 0.31. Further investigation of the relationship between 'incubated' greasy colour and clean colour might be warranted as a means of including a measure of wool colour in Merino breeding programs. To further facilitate inclusion of clean wool colour in breeding programs analysis of auction data to calculate the relative economic value of that trait is needed.

ACKNOWLEDGEMENTS

The Woolmark Company and CRC for Premium Quality Wool supported the CSIRO FWP. This analysis was supported by the CRC for Sheep Industry Innovation. Valuable contributions were made by the CSIRO technical staff, participating sheep breeders in the FWP and T13 Consortium Members.

REFERENCES

Brown D.J. (2006) *Int. J. Sheep Wool Sci.* **54:**1.

Benavides M.V., Mahar A.P., Young M.J., Beatson P.R. and Reid T.C. (1998) *Aust. J. Agric. Res.* **49:**1195.

Dowling, M.E., Schlink, A.C. and Greeff, J.C. (2007) *Proc. Assoc. Advmt. Anim. Breed. Genet.* **17:**328.

Gilmour A.R., Gogel B.J., Cullis B.R., Welham S.J. and Thompson R. (2002) ASReml User Guide Release 1.0 VSN International Ltd. Hemel Hempstead, HP1 1ES, UK.

Mortimer, S.I. (2007) *Int. J. Sheep Wool Sci.* **55:**75.

Raadsma, H.W. and Wilkinson, B.R .(1990) *Aust. J. Agric. Res.* **41:**139.

Swan A.A., Purvis I.W., Hansford K., and Humphries W. (1997) *Proc. Assoc. Advmt. Anim. Breed. Genet.* **12:**153.

Swan A.A., and Purvis I.W. (2005) *Proc. Assoc. Advmt. Anim. Breed. Genet.* **16:**165.

Swan, A.A., Purvis, I.W. and Piper, L.R. (2008) *Aust. J. Exp. Agric.* **48:**1168.

Wilkinson B.R. (1981) *Wool Tech. Sheep Breed.* **29:**169.

MIXED MODELS IN ANIMAL BREEDING: WHERE TO NOW?

A.R. Gilmour

Cargo Vale, CARGO, NSW 2800,
formerly Orange Agricultural Institute, NSW Department of Primary Industries

SUMMARY

Over the past 60 years, mixed models have underpinned huge gains in plant and animal production through genetic improvement. Charles Henderson (1912-1989) established mixed models for estimating breeding values (BLUP) using the popularly called Henderson's Mixed Model and provided early methods (Henderson's Methods I, II and III) for estimating variance parameters. Robin Thompson then published the widely acclaimed REML method for variance component estimation in 1971. These two innovators, along with the development of computing power, have spawned national and international breeding programs in almost all animal species used for human food and fibre.

Our ability to generate data is outstripping our ability to analyse data and this will lead to mixed models playing new roles in genetic estimation. The focus is changing from simply describing the relationship between variables through a correlation, to modelling the relationship based on knowledge of the Genome.

INTRODUCTION

Selective breeding goes back at least to Jacob (1800 BC, Genesis 30) who selected the fitter rams for his own flock. Traditional breeding has largely relied on visual assessment with many such classers having considerable skill in recognising genetic potential with respect to their objective, whether breeding war horses, dogs or pigeons. What characterises modern breeding though is the extensive use of objective measurement and adjustment for environmental effects.


Charles Henderson

The digital age has seen a rapid increase in the number of traits included in a breeding objective or selection criterion, as well as use of data on relatives to improve the separation of genetic from environmental differences. Charles Henderson (1912-1989) *et al.* (1949, 1959) developed and popularised the mixed model equations which underpin the BLUP estimation of breeding values. His development of these equations included use of the additive genetic relationship matrix, showing how it accommodates selection as well as their primary role of adjusting for nuisance environmental effects.


Robin Thompson

However, the mixed model equations used for evaluation assume knowledge of variance parameters. Henderson (1953) defined the main methods used to estimate these until Robin Thompson (Patterson and Thompson 1971) presented the Residual Maximum Likelihood (REML) method. Karin Meyer and Dorothy Robinson produced software to implement REML methods (in animal breeding and

more generally respectively). However analysis was difficult until Robin presented the Average Information method (Johnson and Thompson 1995; Gilmour *et al*. 1995) underpinning ASReml (1997, 2002, 2006, 2009) which become generally available in 1997.

The promise of the genomic revolution is that we may be able to select directly for specific combinations of genes based on reading an individual's genetic code and having good information on the phenotypic and pleiotropic effects of genes/alleles.

MIXED MODEL EQUATIONS AND BLUP

The linear mixed model is written as

$$y = X\tau + Zu + e$$

where $\boldsymbol{X}$ is the design matrix for fixed effects, $\boldsymbol{\tau}$, $\boldsymbol{Z}$ is the design matrix for random effects, $\boldsymbol{u}$, $\boldsymbol{y}$ is the vector of phenotypic measurements and $\boldsymbol{e}$ is the vector of model residuals. The mixed model equations (MME) are conveniently represented in matrix form by

$$\begin{bmatrix} X'R^{-1}X & X'R^{-1}Z \\ Z'R^{-1}X & Z'R^{-1}Z + G^{-1} \end{bmatrix} \begin{bmatrix} \hat{\tau} \\ \tilde{u} \end{bmatrix} = \begin{bmatrix} X'R^{-1}y \\ Z'R^{-1}y \end{bmatrix}$$

where $\text{var}\begin{bmatrix} u \\ e \end{bmatrix} = \begin{bmatrix} G & 0 \\ 0 & R \end{bmatrix}$. Given $\boldsymbol{R}$ and $\boldsymbol{G}$, the solution for the fixed effects given by the mixed model equations is the same as given by solving $X'V^{-1}X\hat{\tau} = X'V^{-1}$ where $V = R + ZGZ'$. The solutions for the random effects are the Best Linear Unbiased Predictors of those effects and as such are ideal for selecting breeding stock.

The power of this system lies in the structure that can be incorporated into $\boldsymbol{X}$, $\boldsymbol{Z}$, $\boldsymbol{R}$ and $\boldsymbol{G}$. It is is not unusual for $\boldsymbol{u}$ to include sub-vectors for various traits and various 'strata' such as direct genetic, maternal genetic, maternal environment, dominance and nuisance blocking effects. This can lead to a fairly complex structure to $\boldsymbol{G}$ involving relationship matrices and variance matrices of various sorts. The main advantage of the mixed model equations is that the left hand side matrix is typically fairly sparse so that large systems of equations can be solved quite efficiently. This arises because matrices $\boldsymbol{X}$ and $\boldsymbol{Z}$, and inverses of $\boldsymbol{R}$ and $\boldsymbol{G}$ are typically sparse.

RESIDUAL MAXIMUM LIKELIHOOD

Without going into the detail, suffice to say that if we assume $\boldsymbol{u}$ and $\boldsymbol{e}$ (and therefore $\boldsymbol{y}$) are normally distributed (given $\boldsymbol{R}$ and $\boldsymbol{G}$), we obtain an expression involving $\boldsymbol{y}$, $\boldsymbol{X}$, $\boldsymbol{Z}$, $\boldsymbol{R}$ and $\boldsymbol{G}$ which is called the likelihood. This expression can be partitioned into two parts; one providing information on $\boldsymbol{\tau}$ conditional on $\boldsymbol{G}$ and $\boldsymbol{R}$ leading to the mixed model equations, the other providing information on $\boldsymbol{R}$ and $\boldsymbol{G}$ conditional on $\boldsymbol{\tau}$. Residual Maximum Likelihood seeks to find the parameter values for $\boldsymbol{R}$ and $\boldsymbol{G}$ that are most likely because they maximise this second part (rather than the whole likelihood). This maximisation exercise though was not trivial when $\boldsymbol{R}$ and $\boldsymbol{G}$ involved more than a few parameters and the problem was large. Consequently, REML estimation was restricted in application to small problems or well structured standard animal breeding models until Thompson presented the Average Information procedure (Gilmour *et al.* 1995, Johnson and Thompson, 1995) which is also centred around the mixed model equations. The implementation in ASReml exploits the sparsity of the mixed model equations though judicious ordering of the equations, avoiding the need to obtain the complete inverse of the left hand side matrix. Now REML can be applied to large problems (with several hundred variance parameters).

WHERE TO NOW?

One thing programming has taught me is that no matter how big you allow, someone will want bigger. While computing technology has helped with the more traits, more records issue of modern animal breeding based on BLUP technology, we are now faced with genome level data of a higher magnitude and methodologies which do not have the statistical and mathematical rigor that supports conventional quantitative genetics. Three problem areas come to mind. The first is the well established variance estimation problem (Hill and Thompson 1978) that when estimating a variance matrix, the probability that the maximum value of the REML likelihood occurs outside the imposed parameter space increases with the matrix size. The second is the application of mixed models to genomic data. The third is how to effectively combine specific genomic data into the BLUP evaluation process.

Structured Variance models. The more traits involved in a REML analysis, the more likely there will be difficulties with the estimation of all the variances and co-variances involved. ASReml will estimate a negative definite matrix if permitted, or attempt to estimate a positive definite matrix which is almost singular. But this raises the issue of whether a reduced parameterization within the parameter space will be preferable. It is not uncommon to find that a matrix can be reduced by use of principal components to a more parsimonious form. That is, the first 1, 2 or 3 principal components will contain the big bulk of the information contained in the matrix. The remaining variation is noise and is often associated with negative eigen values. Therefore it makes sense to estimate the matrix based on some underlying structure. Three structures are common in ASReml. For variates that have no intrinsic ordering, the principal component/factor analytic models allow more parsimonious modelling. For measurements repeated at irregular intervals, the random regression models are often applied but these may produce unreasonable estimates at the ends of the time range. For regular repeated traits, for example weights at successive ages, the expected structure is an autoregressive one for which the Antedependence (Generalised auto regressive) models apply. Jaffrèzic *et al.* (2002) has extended the Antedependence model to a Structured antedependence where a model is imposed on the regression and innovation parameters. Meyer and Kirkpatrick (2009) have investigated a reduced parameterization based on assuming common eigen vectors across strata which is another proposal within this framework. To my mind, this leads to a general area of writing models for the variance parameters, and is the next logical step when it comes to fitting models with hundreds of variance parameters. The question will always be whether a reduced parameterization has adequately captured the real variation without imposing a structure unsupported by the data.

Mixed models for genomic data. There is a huge literature on analysing the huge amount of genomic data that is being presented and little consensus on the best approach. One issue is the diversity of kinds of data available and the other is the sheer volume of data and the knowledge that meaningful/useful variation is present in only a small proportion of it. The issue here is then to separate signal from noise. I believe mixed models could have a bigger role here because signal will represent a covariance (or inflated variance) over the noise (base variance). Mixed models have been successfully used to adjust for spatial variation in genomic slides. They have been used to locate QTL in back-cross/F2 experiments (Gilmour 2007; Verbyla *et al.* 2007) and in association studies where there are often more 'markers' than experimental units. Thomson *et al.* (2009) use mixed models as part of their procedure to combined cattle and sheep genomic data to look for differentially expressed genes. The new outlier method in ASReml 3 may help in this regard.

Incorporating genomic markers in BLUP evaluation. Scientists are an optimistic group when it comes to incorporating genetic markers into BLUP evaluation. I suspect there is a lot of detailed work required before this becomes standard procedure across the industries.

DISCUSSION

Linear Mixed Models have underpinned a revolution in livestock breeding in the last 50 years and despite the huge investment in genomic research and Bayesian methods, there remains a continuing major role for them in the foreseeable future. However, the general model needs adaption for the specifics of each particular species and application. By this I mean, identification of the principle sources of variation, whether they should be accommodated as fixed or random effects, appropriate variance structures and extending the analyses to larger populations and with more traits.

While a bivariate analysis is now readily performed, larger multivariate analyses for the estimation of positive definite variance matrices are often difficult requiring use of structured matrices and raising the issue of whether the structure is adequate. There will undoubtedly be further developments in this area.

The literature on analysis of genomic data reports a wide range of methods as people have hurried to analyse their large amounts of newly acquired data. Some of these analyses have demonstrated the utility of mixed models in this area, but have also shown up limitations due to the amount and structure of the new data. This also will need more attention.

ACKNOWLEDGEMENTS

Upon my retirement, I am grateful to all my colleagues, to NSW DPI, and to the gracious Lord God, creator, for the opportunity given me to contribute to plant and animal breeding through my part in ASReml, in LAMBPLAN (through the program BVEST), through mentoring and assisting other researchers and in the analysis of breeding data, especially for sheep.

REFERENCES

Gilmour, A. R. (2007) *Comp Stats and Data Analysis* **51**:3749

Gilmour, A. R., Thompson, R. and Cullis, B. R. (1995) *Biometrics* **51**:1440

Gilmour, A. R., Gogel, B. G., Cullis, B. R. and Thompson, R. (2009) ASReml 3. www.vsni.co.uk

Henderson, C. R. (1953) *Biometrics* **9**:226

Henderson, C. R., Kempthorne, O., Searle, S. R. And Von Krosigk, C. N. (1959) *Biometrics* **15**:192

Hill, W. G. and Thompson, R. (1978) *Biometrics* **34**: 429.

Jaffrèzic, F., White, I. M. S., Thompson, R., Visscher, P. M. and Hill, W. G. (2002) *7WCALGP*

Johnson, D. L. and Thompson, R. (1995) *J. Dairy Sci.*, **78**:449

Meyer, K., and Kirkpatrick, M. (2009) AAABG

Patterson, and Thompson, R. (1971) *Biometrika* **58**:545

Thomson, P. C., Singh, M., and Raadsma, H. W. (2009) AAABG

Verbyla, A. P., Cullis, B. R. and Thompson, R. (2007) *Theor. And Appl. Genet* **116**:95

LATENT MIXED MODELS

Robin Thompson

Centre for Mathematical and Computational Biology, and Department of Biomathematics and Bioinformatics, Rothamsted Research, Harpenden AL5 2JQ, UK and School of Mathematical Sciences, Queen Mary, University of London, Mile End Road, London E1 4NS, UK

SUMMARY

The linear mixed model has been a major research interest of Dr Arthur Gilmour, motivated by problems arising in research data generated by agricultural scientists. He has developed a variety of computer packages associated with the linear mixed model. He has tirelessly assisted researchers in the analysis and interpretation of their data using these packages. In this quest to help researchers he has made several important innovations. The purpose of this paper is to review some of these innovations, including improved iterative schemes for estimating variance parameters, developing a powerful scheme for specification of linear mixed models and exploiting sparsity to reduce the computational burden. Some of these innovations are only implicitly described in computer user guides and deserve wider recognition.

INTRODUCTION

In this paper we review some of the contributions of Dr Arthur Gilmour to the analysis of correlated data, especially that generated by agricultural scientists. He has developed several packages that help with this analysis including REG (Gilmour 1993a), a generalised linear model program, TwoD (Gilmour 1992), a program for the analysis of spatial data, BVEST (Gilmour 1993b), a program for the prediction of breeding values using best linear unbiased prediction (BLUP) and more recently ASReml (Gilmour *et al.* 2006) a program for the analysis of mixed models. In this paper some of the features involved in the development of this program will be reviewed, including some of the theory behind the program, the specification of the model and taking account of sparsity in various algorithms introduced to reduce the computational burden.

THEORY

Some of the main points arise by considering a linear model of the form

$$\boldsymbol{y} = \boldsymbol{X\tau} + \boldsymbol{Zu} + \boldsymbol{e} \qquad \textbf{(1)}$$

where $\boldsymbol{y}$ is the vector of phenotypic measurements and $\boldsymbol{X}$ is the design matrix for fixed effects, $\boldsymbol{\tau}$, $\boldsymbol{Z}$ is the design matrix for random effects, $\boldsymbol{u}$, and $\boldsymbol{e}$ is the vector of model residuals and where effects vector, $\boldsymbol{u}$, and the residual vector, $\boldsymbol{e}$, are assumed to be multivariate normally distributed with $\text{var}\begin{bmatrix}\mathbf{u}\\ \boldsymbol{e}\end{bmatrix} = \begin{bmatrix}\boldsymbol{G} & \mathbf{0}\\ \mathbf{0} & \boldsymbol{R}\end{bmatrix}$. Because this linear model includes fixed effects, $\boldsymbol{\tau}$, and random effects, $\boldsymbol{u}$, this is called a mixed model. It has many applications. In the analysis of experiments interest is in estimation of treatment effects**,** $\boldsymbol{\tau}$, taking account of the correlated variance structure. In some genetic applications there is interest in estimation of the genetic variances and covariances in **G**, adjusting the data for the fixed effects $\boldsymbol{\tau}$. In other applications, which particularly motivated Henderson, there is interest in predicting the random effects $\boldsymbol{u}$ and Henderson (1973) introduced the system of equations

$$\begin{bmatrix}\boldsymbol{X'R^{-1}X} & \boldsymbol{X'R^{-1}Z}\\ \boldsymbol{Z'R^{-1}X} & \boldsymbol{Z'R^{-1}Z + G^{-1}}\end{bmatrix}\begin{bmatrix}\hat{\boldsymbol{\tau}}\\ \tilde{\boldsymbol{u}}\end{bmatrix} = \begin{bmatrix}\boldsymbol{X'R^{-1}y}\\ \boldsymbol{Z'R^{-1}y}\end{bmatrix} \qquad \textbf{(2)}$$

$$\text{or} \quad [\boldsymbol{C}]\begin{bmatrix}\hat{\boldsymbol{\tau}}\\ \tilde{\boldsymbol{u}}\end{bmatrix} = \begin{bmatrix}\boldsymbol{X'R^{-1}y}\\ \boldsymbol{Z'R^{-1}y}\end{bmatrix} \qquad \textbf{(3)}$$

and showed that the solutions for the random effects are the Best Linear Unbiased Predictors (BLUPs) of those effects. Given $\boldsymbol{R}$ and $\boldsymbol{G}$, the solution for the fixed effects given by these equations is the same as given by solving $\boldsymbol{X'V^{-1}X\hat{\tau} = X'V^{-1}y}$ where $\boldsymbol{V = R + ZGZ'}$. Because of the similarity of these equations to least squares equations, which are the same as **(2)** but without the $\boldsymbol{G^{-1}}$ term, these equations are now called Henderson's mixed model equations.

In other cases there is interest in parameterizing $\boldsymbol{R}$ and $\boldsymbol{G}$, as a function of variance parameters and estimating these variance parameters. Patterson and Thompson (1971) introduced a residual maximum likelihood (REML) method based on maximising the likelihood of error contrasts i.e. contrasts that contributed no information on fixed effects. This REML method agreed with analysis of variance methods in balanced cases and effectively eliminated bias in variance estimation due to not knowing the fixed effects. Convenient forms of the residual log-likelihood (Harville 1974, Smith and Graser 1985) are
L=-(0.5)(D+S), where D=*logdet*$\mathbf{(V)}$+*logdet*$\boldsymbol{(X'V^{-1}X)}$ = *logdet*$\mathbf{(G)}$+*logdet*$\mathbf{(R)}$+*logdet*$\mathbf{(C)}$
and $S = \boldsymbol{y'Py}$, is a residual sum of squares with $\boldsymbol{P = V^{-1} - V^{-1}X(X'V^{-1}X)^{-1}X'V^{-1} = R^{-1} - R^{-1}WC^{-1}W'R^{-1}}$ with $\mathbf{W=[X\ Z]}$.

This residual likelihood is of the same form as the full likelihood with the addition of the last term in D that is sometimes thought of as a penalty for estimating the fixed effects. Smith and Graser (1985) introduced the form that is a function of **C** that naturally leads to sequential formation of the likelihood. The terms D and S can be formed sequentially by using terms associated with eliminating, or absorbing, the fixed and random effects one by one from the MME. An advantage of this is that $\mathbf{C}^{-1}$ is not calculated and, because of the large number of zero elements in **C**, computation can be substantially reduced by using sparse matrix methods. One disadvantage was that the differentials of the likelihood were not easily available. To maximize the likelihood with one parameter Smith and Graser (1985) suggested using a quadratic approximation. With more than one parameter, methods that avoid calculating derivatives become a popular flexible alternative, and the DFREML program written by Meyer (1991a) was used extensively. For example, these methods were used for Animal and Reduced Animal Models, both for univariate and multivariate data (Meyer, 1989, 1991b). The basic framework was extended to include more biologically appropriate models with genetic components, including maternal models with both Wilham and Falconer terms (Koerhuis and Thompson, 1997), and models with mutation terms (Wray, 1990).

A major disadvantage of derivative-free methods is that the computational effort increases dramatically as the number of variance parameters increases. An important advance in derivative-based methods was the rediscovery (Misztal and Perez-Enriso, 1993) of an algorithm (Takahashi *et al*. 1973) allowing calculation of the terms in $\mathbf{C}^{-1}$ required for forming the first differentials, what might be called the sparse inverse of **C**, without calculating all the elements of the inverse. Meyer and Smith (1996) introduced an alternative way of calculating these first differentials by performing the 'automatic' differentiation of the Cholesky decomposition of **C.** These techniques for forming first derivatives both require twice the computational effort of forming the likelihood. An alternative derivation in terms of sequential formation of the sparse inverse of $\mathbf{C}^{-1}$ parallels the sequential formation of the likelihood (Thompson *et al*. 1994). This result allowed the implementation of EM algorithms to estimate variance parameters (Misztal 1994). These were an improvement on derivative free methods but could still be slow to converge. It is possible to calculate second differentials using automatic differentiation (Smith 1995) but the computation of each second differential requires six times as many multiplications as those involved in a single likelihood calculation (Smith 1995), and this becomes more costly as the number of parameters increases. A convenient suggestion (Johnson and Thompson 1995; Gilmour *et al*. 1995) is based on manipulation of the alternative information matrices. They show that the average of the observed and expected information (AI) is relatively easy to calculate. The AI matrix is based on

the differential of the sum of squares S, and can be written in the similar way to S itself, using working variables based on estimates of the fixed effects and predictors of the random effects.

A synthesis of comparisons of different algorithms used to compute REML estimates was carried out by Hofer (1998) and is updated in Table 1. These comparisons show the expected improvement of EM methods over derivative free methods. The comparisons also show that most second differential methods converge in relatively few iterations.

Table 1. Results of empirical comparison of REML algorithms with regards to rounds of iteration (function evaluations for DF) and total time to convergence [a]

Ref[a]	MME[b]	Par[c]	DF[d]		EM		NR/AI[e]	
			F.Eval	Time (h)	Rounds	Time (h)	Rounds	Time (h)
1	4895	3	26	0.01	24	0.05		
	9790	9	238	0.31	33	0.26		
	14685	18	583	1.77	45	1.02		
2	6192	9	699	1.27			6	0.45
	10230	12	1236	2.33			8	0.90
	14274	18	4751	11.10			18	3.33
3	5731	5	169	0.34			6	0.07
4	8765	6	927	70.6	109	1.14	7	1.86
5	5073	2	39	0.02	23	4.97	5	0.02
6[f]	233796	55	37021	2083			185	40.10
7	46581	12	1435	15.2	1006	88.6	6	0.58
	55410	19	5813	30.6			6	1.00

[a] References 1 Misztal 1999; 2 Meyer and Smith 1996; 3 Johnson and Thompson 1995 ; 4 Gilmour *et al.* 1995; 5 Madsen *et al.* 1994 ; 6 Neumaier and Groeneveld 1998; 7 Jensen *et al.* 1997

[b] Dimension of mixed model equations (MME).

[c] Number of (co)variance components.

[d] 'DF' = derivative free

[e] 'NR' = quasi-Newton using computed analytic differences.

[f] quasi -Newton using finite differences.

APPLICATION

With the availability of an efficient algorithm for the calculation of the sparse inverse of **C** and for the calculation of an AI matrix there was interest in developing a general program for estimating variance parameters. This program ,now called ASReml (Gilmour *et al. 2006)*, was designed taking into account features in existing programs. These included REG (Gilmour 1993)a program for estimation of generalized linear models, TwoD (Gilmour 1992), a program for spatial analysis, REML (Thompson 1977,Robinson *et al.* 1982), a program developed for the analysis of variety trials that had been ported to Genstat, (Welham and Thompson 1990) and DFEML (Meyer 1991a).

Linear model specification. The specification of the design matrices **X** and **Z** in **(1)** initially followed the Wilkinson and Rogers (1973) syntax that allowed interaction between factorial terms and the specification of polynomial and regression sub models associated with factorial terms. This specification of models was very general but several modifications were found to be of value to ease the setting up of the linear model terms.

Conditional factors. Sometimes effects are only required to be fitted for subsets of the data. The at() function allows this. For example at(HERD.1).SEASON allows SEASON effects just for data with HERD =1. If no levels of the conditioning factor (at(HERD).SEASON in this case) are specified in the at() function, a complete set of conditioning terms is generate and this allows estimation of variance models for SEASON effects for each level of HERD.

Combining design matrices. There is sometimes the need to combine or overlay design matrices. The and() function allows this. For example SIREBREED and(DAMBREED) allows models with equal SIREBREED and DAMBREED effects to be fitted and this might be useful in the analysis of diallel crosses .Sometimes there is the need to associate competition effects from animals in the same group(Bijma *et al.*,2006). For example if there are 4 animals in each group and DIRECT is a factor indicating each animal and INDIRECT1, INDIRECT2, INDIRECT3 are factors indicating the 3 animals in the same group this can again be modelled using DIRECT and(INDIRECT1) and(INDIRECT2) and(INDIRECT3).

Functions of covariates and factorial effects. ASReml originally allowed model terms that were polynomials of covariates but there is sometimes interest in forming other functions ,for example when fitting splines. To add generality the facility to form user generated functions was added .For example, if the values of a covariate **c** are $\boldsymbol{cv_1}, \ldots, \boldsymbol{cv_L}$ then reading in a file with i-th row $\boldsymbol{cv_i}$ $\boldsymbol{f_j(cv_i)}$ with i=1,…,*L* and j=1,…,*L** and allows the functions $\boldsymbol{f_j(c)}$ with j=1,…,*L** to be constructed and used in a model. Similarly if the effects, $\boldsymbol{u}_i$, for a factor with *L* levels are required to be replaced by $\boldsymbol{u}_i = \boldsymbol{F u_i^*}$, for example if some elements of $u_i$are constrained, then again the ,possibly non-zero, elements of the $L \times L^*$ matrix $\boldsymbol{F}$ can be read in and the L^* effects $\boldsymbol{u_i^*}$ incorporated into the linear model.

Variance specification. In the simplest case when the random effects were all uncorrelated allowed easy specification of the variance structure of the model. However the programs REML and TwoD and the need of users for various models including seperable spatial processes, random regression models and multivariate animal models inspired a wider class of models. Firstly both **R** and **G** were allowed to be expressed as a direct sum of respectively *s* and *b* parts i.e.

$$\boldsymbol{R} = \oplus_{j=1}^{s} \boldsymbol{R_j} = \begin{bmatrix} \boldsymbol{R_1} & \boldsymbol{0} & \cdots & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{R_2} & \cdots & \boldsymbol{0} & \boldsymbol{0} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \boldsymbol{0} & \boldsymbol{0} & \cdots & \boldsymbol{R_{s-1}} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & \cdots & \boldsymbol{0} & \boldsymbol{R_s} \end{bmatrix} \quad \text{and} \quad \boldsymbol{G} = \oplus_{i=1}^{b} \boldsymbol{G_i} = \begin{bmatrix} \boldsymbol{G_1} & \boldsymbol{0} & \cdots & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{G_2} & \cdots & \boldsymbol{0} & \boldsymbol{0} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \boldsymbol{0} & \boldsymbol{0} & \cdots & \boldsymbol{G_{b-1}} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & \cdots & \boldsymbol{0} & \boldsymbol{G_b} \end{bmatrix}$$

The effects associated with each part are uncorrelated with effects in the other parts. This generality of R was partly to allow the analysis of a series of trials with possibly different numbers of plots and different variance parameters in each trial. The parameter b relates to the number of sets of random effects embedded in the random effects $\boldsymbol{u}$. Models for the component matrices $\boldsymbol{R_j}$ and $\boldsymbol{G_i}$ are allowed to be formed as the direct product of models for up to three correlated random factors of the form $\boldsymbol{R_j} = \boldsymbol{R_{j1}} \otimes \boldsymbol{R_{j2}} \otimes \boldsymbol{R_{j3}}$ and $\boldsymbol{G_i} = \boldsymbol{G_{i1}} \otimes \boldsymbol{G_{i2}} \otimes \boldsymbol{G_{i3}}$.

A range of models are available for the components of both **R** and **G**, They include correlation models and covariance models. Correlation models generate variance matrices, **C** , with diagonals one, include uniform, banded and general correlation and models motivated by time-series analysis including autoregressive, moving average and autoregressive-moving average models. The covariance models, generate a variance matrix V, include diagonal, ante dependence, unstructured and factor analytic models. There is also the facility to form homogeneous variance matrices $\boldsymbol{V} = \boldsymbol{\sigma^2 C}$ and heterogeneous variance matrices $\boldsymbol{V} = \boldsymbol{D^{1/2}\, C D^{1/2}}$ and estimate the parameters in the

correlation matrix **C,** the scalar $\boldsymbol{\sigma^2}$ and the parameters in the diagonal matrix $\boldsymbol{D^{1/2}}$**(**Gilmour *et al.* 1998**).** There is also the facility to form an additive relationship matrix from a pedigree, and fit cubic splines by incorporating an equivalent mixed model (Verbyla *et al.*, 1999).Typically a variance structure applies to individual terms in the linear model but the facility is available to impose structure on a combination of terms, for example if there are effects ANIMAL and ANIMAL.TIME with ANIMAL representing animal effects and TIME a covariate and wish to impose random animal terms and animal regression terms and a covariance between these two sets of terms.

Despite the generality of the variance models one cannot always predict the ingenuity of users (for example Jaffervic et al. 2004)and so there is the facility to read in a user generated relationship matrix or its inverse and there is a facility to allow users to define variance matrices (or their inverses) as functions of parameters to be estimated. One problem with user generated relationship matrices is that they may, perhaps due to sampling, be singular. For example a relationship matrix **A** may have rank r and have s singularities, then if we reorder A so that **A** and the associated random effects $\boldsymbol{u_a}$ can be partitioned as $\boldsymbol{A} = \begin{bmatrix} \boldsymbol{A_{11}} & \boldsymbol{A_{12}} \\ \boldsymbol{A_{21}} & \boldsymbol{A_{22}} \end{bmatrix}$ and $\boldsymbol{u'_a} = [\boldsymbol{u'_{a1}}\ \boldsymbol{u'_{a2}}]$ with $\boldsymbol{A_{11}}$ of rank r then if $\boldsymbol{A_{11}^{-1}} = \boldsymbol{A^{11}}$ then because of the s singularities $\boldsymbol{A_{22} - A_{21}\,A^{11}A_{12} = A_{22.1} = 0}$ and the random effects are given by $\boldsymbol{u'_{a2} = A_{21}\,A^{11}u'_{a1}}$**.** One might extend $\boldsymbol{u_a}$ using $\boldsymbol{u_a^{*\prime} = [u'_a\ u^{*\prime}_{a2}]}$ with $\boldsymbol{u^*_{a2}}$ the same size as $\boldsymbol{u_{a2}}$ and setup MME using $\boldsymbol{u^*_a}$ and $\boldsymbol{A^-} = \begin{bmatrix} \boldsymbol{A^{11}} & \boldsymbol{0} & \boldsymbol{A^{11}A_{12}} \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{-I} \\ \boldsymbol{A_{21}\,A^{11}} & \boldsymbol{-I} & \boldsymbol{A_{22.1}} \end{bmatrix}$ replacing **u** and $\boldsymbol{G^{-1}}$ in **(2).**Note that if $\boldsymbol{A_{22.1}}$ is positive definite then with absorption of $\boldsymbol{u^*_{a2}}$ the mixed model equations reduce to the usual form. If $\boldsymbol{A_{22.1} = 0,\ u^*_{a2}}$ can be thought of as introducing Lagrange multipliers to take account of the constraints on $\boldsymbol{u_a}$.

Exploiting sparsity. The MME and the inverse variance matrices involved in the MME are typically sparse with a large number of zero elements. Some of the computational work can be reduced if due account of this sparsity is taken. This reduction includes calculation of the likelihood and its differentials, calculation of variances of linear functions of the effects, and calculations involved with factor analysis models.

Calculation of the likelihood and its differentials. This involves the sequential calculation of the weighted sum of squares, *S*, and the calculation of the sparse inverse of **C** depend on the order that the order of terms in **C.** For example in a simple example with model with a grand mean and *h* uncorrelated herd effects, then **C** has *h*+1 rows and columns and 1+3*h* non-zero elements in **C**.If the sequential formation starts by eliminating the grand mean the resulting updated **C** matrix fills in the non-zero elements of **C** and the sparse inverse requires calculation of all elements of $\mathbf{C}^{-1}$.By contrast, if in the sequential formation the grand mean is the last term then there is no in-fill of **C** and only the 1+3*h* elements, corresponding to the non-zero elements in the initial **C**, need calculating in the sparse inverse of **C**. The computation time with the initial ordering is proportional to $h^3$and proportional to *h* in the later scheme.

In order to reduce the computational time and effort several different orderings were investigated. A simple algorithm would be to (i) count the non-zero elements in each row, (ii) absorb the row with the least number of non-zero elements and update **C**, then repeat the process of the updated **C** matrix. Potentially this requires a large number of ordering operations and so a nested search (NS) algorithm (Gilmour and Thompson, 2007) was suggested. This involves step (i) and step(ii) is extended by finding the row with the least number of non-zero elements and finding the decile of the number of non-zero elements in a row. This decile is used as a cut-off and in step

(iii) absorb rows shorter than the cut-off at the time they are considered for absorption, starting with the shortest row. Once all rows with non-zero elements less than the cut-off have been absorbed then the program goes back to step (i) and the process is repeated until all the rows have been absorbed.

After experimentation an extra grouping step was introduced giving a grouped nested search algorithm(GNS).This grouping step identifies rows which have the same columns present. Each such group of rows is represented by just one member in the list of equations yet to be absorbed; the whole group is absorbed when that representative is absorbed. This has little impact on the order obtained but does reduce the work required to determine the order.

Timings for two examples are given in Table 2 for these two ordering algorithms (NS and GNS) and another suggested ordering (MeTiS). Ducrocq and Druet (2003) identified the MeTiS suite of ordering algorithms (Karypis and Kumar, 1998) as being very suitable for the solution of

Table 2. Fill-in(size of coefficient matrix after absorption) and timing for ordering and one iteration for two examples and three algorithms run on a 32bit Linux.

Example	Algorithm	Fill-in	Order(secs)	Iteration (secs)
Multivariate	MeTiS	7,477,499	1	582
	NS	6,726,733	195	415
	GNS	6,703,056	124	412
Random Regression	MeTiS	39,724,222	33	473
	NS	38,228,850	1984	487
	GNS	38,202,783	45	432

mixed model equations. These algorithms are based on a multilevel k-way partitioning of graphs. MeTiS uses initial coarsening, partitioning and then uncoarsening resulting in multilevel, nested dissection involving top-down and bottom-up ordering. Meyer (2005) evaluated MeTiS in her software, comparing it to a range of other published algorithms including the multiple minimum degree routine GENMMD (George and Liu, 1981, Lui, 1985) and various refinements extracted from the MUMPS package (Amestoy *et al.*, 1998, 2001). She concluded MeTiS produced orderings which required about half the execution time of the minimum degree algorithms although the times vary a little with the particular characteristics of the example.

The first example is a six trait multivariate sire model analysis of data from 26875 animals fitting 31745 equations and 39 variance parameters. The second example is a multivariate random regression of 437,632 data records fitting 536,288 equations and 116 variance parameters. The MeTiS ordering routines quickly find a good order for solving mixed model equations. However, in the context of ASReml, the MeTiS order is on balance not the optimum order when compared with the GNS ordering algorithm. This is often 10-20% faster per iteration.

Calculation of variances of linear functions of the effects. It sometimes required to estimate a linear combination of the fixed and random effects $\boldsymbol{x_p\hat{\tau} + z_p\tilde{u}}$ and with $\boldsymbol{w_p = [x_p\ z_p]}$ the variance of this combination is $\boldsymbol{w_p C^{-1} w_p'}$. By replacing the right hand side of equation (3) by $\boldsymbol{w_p'}$ it can be shown (Gilmour *et al.* 2004) that the variance of the linear combination could be found in a sparse way in exactly same way as the fitted sum of squares from the linear model (1).

Factor analysis models. These models have been suggested as a way of giving a parsimonious way of modelling variance structures. They can be motivated by writing a vector of p random effects $\boldsymbol{u_i}$ as $\boldsymbol{u_{i1} = L_1 f_i + \delta_{i1}, u_{i2} = L_2 f_i}$ *with* $\boldsymbol{u_{i1}, u_{i2}}$ and $\boldsymbol{f_i}$ of size r, s and k respectively.The vector

$\boldsymbol{u_i}$ may represent the performance of the i-th genotype in p environments or p different traits.This is motivated by writing the elements of $\boldsymbol{u_i}$ as linear combinations of the k elements of $\boldsymbol{f_i}$ and the first *r* of the elements of $\boldsymbol{u_i}$ having extra components based on $\boldsymbol{\delta_{i1}}$ .The variance matrix of $\boldsymbol{u_i}$ can be then be written as $\boldsymbol{FA} = \begin{bmatrix} \boldsymbol{L_1 L_1'} + \boldsymbol{D_1} & \boldsymbol{L_1 L_2'} \\ \boldsymbol{L_2 L_1'} & \boldsymbol{L_2 L_2'} \end{bmatrix}$ with var$(\boldsymbol{\delta_{i1}}) = \boldsymbol{D_1}$.If *r*=0 then we have a reduced rank model (Meyer 2005). If $\boldsymbol{u_{i1}}$and$\boldsymbol{f_i}$ are used instead of using $\boldsymbol{u_i}$,in the mixed model equations then the sparsity increases especially if *k* is much less than *p* because the inverse matrix associated with these terms is sparser than the inverse variance matrix associated with $\boldsymbol{u_i}$. Thompson *et al.* (2003) give examples where the alternative parameterization had savings of 50% of computational time when *p*=3 and *k*=3 and savings of 90% when *p*=62 and *k*=3.

Also if the effects $\boldsymbol{u_i}$ are associated with genotype *i*, and are combined into a vector **u ,** and the genotypes have an additive relationship matrix **A** then var**(u**)= $\boldsymbol{FA} \otimes \boldsymbol{A}$ **.** The calculation of the average information matrix requires the calculation of terms such as **s=Ar** where **r** is a vector of residuals **(**Kelly *et al.* 2009). As $\boldsymbol{A^{-1}}$ is much sparser than **A,** Kelly *et al* . (2009) found it useful to think of s as the solution to $\boldsymbol{A^{-1}}$**s=r.** This solution can be found in a recursive way first adjusting the right hand sides of ancestors for direct descendants and then adjusting direct descendents for ancestors (Kelly *et al.* 2009).This computation avoids the formation of **A** and the computational effort is linear in the number of genotypes.

DISCUSSION

The initial motivation in developing the computer program was to generate a kernel to be ported into Genstat. Such was the enthusiasm for the methods that a standalone program was developed, partly using code from REG to allow a wide range of data transformations. This has needed much more support but has the advantage that the algorithm has been made available to a wider user community. The user guide has been cited in over 1,000 publications. These publications are an appropriate testimony to the insight and innovations of Arthur Gilmour.

ACKNOWLEDGEMENTS

This research was supported by Lawes Agricultural Trust. I acknowledge the unstinting help Arthur Gilmour has given me over my career.

REFERENCES

Amestoy, P.R., Duff, I.S. and Excellent, J.-Y. (1998) *Comp. Methods Appl. Mech. Eng.***184** :501

Amestoy, P.R., Duff, I.S., Koster, J. and Excellent, J.-Y. (2001) *SIAM J. Matrix Anal. Appl.***23** :15.

Bijma,P., Muir, W.M., Ellen, E.D., Wolf, J.B. and Van Arendonk, J.A.M. (2007) *Genetics **175***: 289.

Ducrocq, V. and Druet, T. (2003) *54th Annual Meeting, Europ. Ass. Anim. Prod. :1.*

Henderson, C.R. (1973) In "*Proc. Animal Breeding and Genetics. Symposium in Honor of Dr. Jay L. Lush" Amer. Soc. of Anim. Sci. and Amer. Dairy Sci. Assoc., Champaign, Illinois*, p. 10.

George, A. and Liu, J. W.-H. (1981) *Computer Solution of Large Positive Definite Systems.*

Prentice-Hall, Inc., Englewood Cliffs, New Jersey.

Gilmour, A.R. (1992) "TwoD. A program to fit a mixed linear model with two dimensional spatial adjustments for local trend" NSW Agriculture, Tamworth, Australia.

Gilmour, A.R. (1993a) "REG - A generalised linear models program" NSW Agriculture, Orange, Australia.

Gilmour, A.R. (1993b) "BVEST : A multiple trait animal model best linear unbiased predictor breeding value estimation program " NSW Agriculture, Orange, Australia.

Gilmour, A.R., Cullis, B.R., Frensham, A.B. and Thompson, R. (1998). *Compstat'98 Proceedings* 53.
Gilmour, A.R., Cullis, B.C., Welham, S.J., Gogel, B.J. and Thompson, R (2004) *Computational Statistics and Data Analysis* **44**:571.
Gilmour, A.R., Gogel, B., Cullis, B.R. and Thompson, R. (2006) "ASReml 2 User Guide" VSN International, Hemel Hempstead, UK.
Gilmour, A.R. and Thompson, R. (2006) *8th World Congr. Genet. Appl. Livest. Prod.* **27**:12.
Gilmour, A.R., Thompson, R. and Cullis, B.R. (1995) *Biometrics* **51**:1440.
Harville, D.A. (1974) *Biometrika* **61**:383.
Hill, W.G. and Thompson, R. (1978) *Biometrics* **34**: 429.
Hofer, A. (1998). *J. Anim. Breed. Genet.* **115**:247.
Jaffrézic, F. , White, I. M. S. , Thompson, R. , Hill, W. G. and Visscher, P. (2002) *J. of Dairy Science* **85**:968.
Johnson, D.L. and Thompson, R (1995) *J. Dairy Sci.*, **78**:449.
Jensen, J., Mantysaari, E.A., Madsen, P. and Thompson, R. (1997) *Indian Journal of Agric. Stat.* **49**: 215.
Karypis, G. and Kumar, V. (1998) *SIAM Journal of Scientific Computing*, **20:** 359.
Kelly, A.M., Cullis, B.R., Gilmour, A.R., Eccleston, J.A. and Thompson, R. (2009) *Genetics, Selection and Evolution* **41**:33.
Koerhuis, A.N.M. and Thompson, R. (1997). *Genetics, Selection and Evolution* **29**:225.
Liu, J. W.-H. (1985) *ACM Trans. Mathh. Soft.* **11 :**141**.**
Madsen, P., Jensen, J. and Thompson, R. (1994) *5th World Congr. Genet. Appl. Livest. Prod.* **22**:19.
Meyer, K. (1989) *Genetics, Selection and Evolution* **21**:317.
Meyer, K. (1991a) "DFREML User Notes Version 2.0" University of New England, Armidale, Australia.
Meyer, K. (1991b) *Genetics, Selection and Evolution* **23**:67.
Meyer, K. and Smith, S.P. (1996). *Genetics, Selection and Evolution* **28**:23.
Meyer, K. (2005) *Proc. Assoc. Advmt. Anim. Breed. Genet.* **16:** 282.
Meyer, K. (2008) *Genetics, Selection and Evolution* **40:**3.
Misztal, I. (1994) *J. Anim. Breed. Genet.* **111**:346.
Misztal, I. and Perez-Enciso, M. (1993) *J. Dairy Sci.* **76**:1479.
Neumaier, A. and Groeneveld, E. (1998) *Genet. Sel. Evol.* **30**: 3.
Robinson, D.L., Thompson, R. and Digby, P.G.N. 1982. *COMPSTAT 1982, II*, Physica-Verlag, Wien :231.
Patterson, H.D. and Thompson, R. (1971) *Biometrika* **58**:545.
Thompson, R. (1977) *Biometrics* **33**:497.
Thompson, R., Cullis, B.C., Smith, A. and Gilmour, A.R. (2003) *Australian and New Zealand Journal of Statistics* **45**:445.
Smith, S.P. (1995) *J. Comp. Graph Stat.* **4**:134.
Smith, S.P. and Graser, H.-U. (1986) *J. Dairy Sci.* **69**:1156.
Takahashi, K., Fagan, J. and Chin, M.S. (1973). *Proc. 8th Inst. PICA Conf., Minneapolis:* 63.
Verbyla, A. P., Cullis, B. R., Kenward, M. G. and Welham, S. J. (1999) *Applied Statistics* **48**:269.
Welham, S. J. and Thompson, R. (1990) In "Genstat 5, Release 2, Reference Manual Supplement" Numerical Algorithms Group, Oxford p86.
Wilkinson, G.N. and Rogers, C.E. (1973) *Appl. Statist.* **22**:392.
Wray, N.R. (1990) *Biometrics* **46**:197.

STATISTICAL GENETICS TO IMPROVE ROBUSTNESS OF DAIRY COWS

R.F. Veerkamp, H.A. Mulder, M.P.L. Calus, J.J. Windig, and J. ten Napel

Animal Breeding and Genomics Centre, Wageningen UR, PO Box 65, 8200AB Lelystad, The Netherlands

SUMMARY

Rapid changes in farm scale and pressures on costs and farm management together with the negative correlation between production and fitness traits have stimulated the demand for robust cows. We have defined a robust cow as: "a cow that is able to maintain homeostasis in the commonly accepted and sustainable dairy herds of the near future" which clearly contains an element of environmental sensitivity and genotype by environment interaction (GxE). Classical breeding solutions to breed for more robustness are i) avoiding inbreeding, ii) multi-trait selection and iii) allowing more natural selection. Statistical models allow more direct selection for robustness by estimating genetic correlations between environments (discrete macro scale), reaction norms describing genotypes as a function of a continuous environmental parameter (continuous macro scale) and genetic variation in residual variance which is environmental sensitivity for a large number of unidentifiable environmental aspects with a relatively small effect each (unknown micro scale). The application of these complex models is still under development, but first results indicate more GxE for fitness traits than yield traits, albeit evidence for strong re-ranking of animals is still limited, so it is mainly the magnitude of the variance which is affected by environment. These statistical models will contribute, together with classical tools and new phenotypic and genomic measurement tools, to the breeding of cows that also fit in future dairy systems.

INTRODUCTION

Selection for more robust cows has been an important research topic in our group. The aim of this paper is to highlight some of the work we and others have done on robustness, especially in relation to the numerical methods used to select for more robust cows. But first we discuss conceptually why we think that robustness has become important and the definition of robustness Napel *et al.* (2009).

DEMAND FOR ROBUSTNESS

Cannon (1932) first used the term homeostasis to indicate that a body continuously acts to maintain a stable internal environment by responding to external environmental stimuli. In the last two decades, there have been concerns that high-yielding dairy cows struggle to maintain homeostasis. Several studies reported unfavourable genetic correlations between milk yield and reproductive problems, locomotive problems and udder health problems (Pryce *et al.* 1997; Rauw *et al.* 1998; Royal *et al.* 2000), and there is general consensus that selection for milk fat and milk protein yield alone may give an unfavourable correlated response in these traits. The magnitude of these correlated responses is rather small, compared with direct effects of environmental disturbance, albeit when the effects of breeding are accumulated across years these might be substantial. The gradual reduction in genetic levels for fertility and health will put more pressure on management to maintain performance at acceptable levels. Such cows require more management attention. At the same time, the level of management is increasingly under pressure from other directions. For example, due to economic pressure, herd size is increasing and therefore the amount of labour available per animal is decreasing. The shortage of labour is aggravated by the fact that it is increasingly difficult to find suitably skilled labour. Also, pressure on

management increases because previously simple and effective management tools, such as the use of antibiotics, are now perceived as potential risks for human health and therefore regulated much stronger.

These two trends, i.e. negative effects from selection for yield and increasing pressure on management, have fuelled the demand for more robust cows. The demand for robustness in animals is a shift in emphasis between two paradigms both aiming to control the impact of disturbances on an animal (Ten Napel *et al.* 2006). The one approach is called the Control Model and is characterized by maintaining stability through keeping away disturbances. Typically a strategy is used of protecting animals from disturbances as much as possible, constantly monitoring animals, whether a disturbance occurs, and interventions targeted at the disturbance, when it does occur. When taken to the extreme, the homeostasis of the animal is dependent on proper and timely functioning of humans and technical equipment. The other approach is called the Adaptation Model. This approach is characterized by maintaining stability through minimizing the impact of disturbances in the presence of the disturbance. The design of such a production system seeks to utilize the intrinsic capacity of animals to adapt where possible, and use the Control Model approach where necessary (Ten Napel *et al.* 2006).

DEFINITION OF ROBUSTNESS

In the Netherlands, we gradually developed a concept of robustness of farm animals in the course of three to four years, based on discussions with many groups of stakeholders. This process resulted in the following definition of a robust dairy cow: "A robust dairy cow is a cow that is able to maintain homeostasis in the commonly accepted and sustainable dairy herds of the near future."

It is clear from this definition that robustness is not just a matter of the average level of management being suitable for the cow. Dairy herds are dynamic and fluctuations in temperature, air speed, humidity, disease pressure, fodder quality, stocking density, social interaction with other cows, aggression, interaction with stockmen, among other factors, occur. Over time or across herds, common fluctuations largely fit within a certain band width. A cow that is robust is able to maintain homeostasis in a range of production environments with a bandwidth that is wider than the common band width of fluctuations. It does not mean that a robust cow must be able to cope with anything. Some animals respond to a change in environment for some characteristics in a much stronger way than other animals. Such animals are more environmentally sensitive for these traits. In population biology terms, these animals are called 'specialists' as they have a very high fitness only in specific conditions. Less environmentally sensitive animals are called 'generalists'. These qualifications are not absolute, but relative to the range of environments considered. A cow may be a 'specialist' when considering all possible environments, but a 'generalist' when considering the range of acceptable production systems in a country. A robust cow is more of a 'generalist' as it has a reasonable fitness across relevant production environments (Bryant *et al.* 2006).

CLASSICAL TOOLS TO BREED FOR ROBUSTNESS

Three obvious ways to account for robustness in animal breeding are (i) to avoid inbreeding, (ii) to include fitness traits in the breeding goal, breeding value estimation and the selection index and (iii) natural selection. Overcoming or avoiding inbreeding depression and maximizing heterosis is a relatively easy way to improve genetic fitness and is widely utilized in pig and poultry breeding programmes. This effect stems from the observation that characteristics associated with genetic fitness often reveal overdominance, that is the phenotype of heterozygotes is superior to the phenotype of any of the two types of homozygotes. Crossbreeding in a structured manner is less popular in dairy cattle breeding, because the average number of calves per calving is just over one. Furthermore, unfavourable recombination counteracts the favourable heterosis for

milk production traits (Pedersen and Christensen 1989; Dechow *et al.* 2007). Rotational crossing, may be suitable for this purpose (McAllister 2002), but within breed optimizing gain and inbreeding in a breeding program might be as effective (Meuwissen and Sonesson 1998).

Multi-trait selection has been practiced in many dairy countries (Miglior *et al.* 2005). Although multi-trait selection clearly works and is better than single trait selection, there are several reasons why multi-trait index might not be enough to maintain or improve robustness. Firstly, it is difficult for the many aspects of fitness to be clearly defined as a trait to be measured. Secondly, it has been suggested to include traits of the adaptive systems, e.g. immune system, in the multi-trait index (Wagter *et al.* 2000). These adaptive systems, however, are highly integrated life systems with a high degree of unpredictability and ambiguity observed (Tada 1997). Therefore, it is unlikely that genetic selection for changes in the immune system will improve robustness in the foreseeable future, and breeding is likely to be more successful when it concerns the result of adaptation (i.e. effective coping), rather than the adaptation process itself. Finally, fitness traits are expected to exhibit genotype by environment interaction, causing animals to rank differently in different environments. In this way, after any major change in living conditions, there is at least a part of the population that is able to survive in the new environment. Breeding value estimation does not take account of this variation, as generally the estimation models include a correction for heterogeneous variances. Therefore selection does not take into account differences in environmental sensitivity or genotype by environment interaction. When following the definition of robustness given above, robustness is about maintaining homeostasis across environments and environmental challenges, and not only about having a high fitness on average across environments. Clearly environmental sensitivity and genotype by environment interaction play an important role when breeding for robustness.

In natural populations without artificial selection, genetic fitness is maintained through a self-structuring force, called natural selection. If through a change in the environment, variation in genetic fitness arises, then the increase in fitness in the population is equal to the additive genetic variance of fitness at that time (Falconer and Mackay 1996). However natural selection cannot be utilized easily in breeding programmes, for practical reasons. Breeding animals are often kept under strict biosecurity control in order to be able to sell semen or breeding stock. Also, natural selection will be perceived as bad stockmanship or is not acceptable because of the welfare of the animals.

STATISTICAL TOOLS TO BREED FOR ROBUSTNESS

Statistical models that enable to extract information on robustness and environmental sensitivity of animals might be important. In order to estimate genotype by environment interaction, models have developed from (1) environments falling into a limited number of categories (discrete macro scale), (2) production environments differing on a continuous environmental parameter (continues macro scale), and (3) environments differing on a large number of aspects with a relatively small effect each, which are not identifiable (unknown micro scale).

Disturbances by environments on a discrete macro scale. The classical GxE model is when there are distinct discrete environments such as, for example, organic and non-organic dairy herds (Nauta et al. 2006). Grouping herds on prevailing soil type may yield five or six categories. Other examples are grouping by country, geographical region, type of production system or presence or absence of a major disease, and seasons. With these genetic parameters it is possible to carry out a multi-trait breeding value analysis for each trait measured in different groups of environments. In some cases distinct environmental groups, or character states of herds were formed within the environmental levels for yield, heat load index, herd size and altitude in relation to milk

production (Bryant et al. 2007). With these genetic parameters it is possible to carry out a multi-trait breeding value estimation for each trait measured in different groups of environments. If the genetic correlation between the two groups of environments is lower than 0.6-0.7 (Mulder et al. 2006), it suggests that in such a scenario most genetic progress is made for that trait, when breeding specialists rather than breeding for a generalist.

Disturbances by environments on a continuous macro scale

In recent years the reaction norm model has become increasingly important. The performance of a genotype as a function of an environmental parameter is called the reaction norm. If a change in a certain environmental parameter affects some genotypes more than others, then there is genetic variation in environmental sensitivity. Reaction norm models have been applied to environments quantified by the mean performance of all genotypes, which then becomes the environmental parameter. Initially models were used using indirect estimation procedure (Veerkamp and Goddard 1998), but software developments allowing random regression models to be used (Schaeffer and Dekkers 1994; Gilmour *et al.* 2000), have taken over from the indirect approach For example, (Kolmodin *et al.* 2002) used days open and average milk yield as environmental descriptors, others investigated a large number of environmental parameters describing differences between herds in for example production, management, climate and nutrition in relation to production (Calus and Veerkamp 2003; Fikse *et al.* 2003) or in relation to fertility and health (Windig *et al.* 2005b), or looked at a heat stress index in relation to days open (Oseni *et al.* 2004), fat% in relation to milk fat depression (Calus *et al.* 2005), or average herdlife, yield and herd size in relation to length of productive life (Petersson *et al.* 2005).

Despite these relative complex models, for the yield traits genetic correlations across environments are all close to unity albeit variances almost doubled across environments (Calus and Veerkamp 2003; Kolmodin *et al.* 2004). For fitness traits, i.e. health and fertility, larger differences in genetic variances were observed across environments: genetic variances for fertility traits increased in some situations more than twofold, and a threefold increase for genetic variances of somatic cell scores was found (Windig *et al.* 2005a). Genetic correlations of a trait across environments were as low as 0.65 for survival (Windig *et al.* 2005a), while for somatic cell scores, on a test-day level, the lowest genetic correlation was as low as 0.72 (Calus *et al.* 2006). Within country, literature values of genetic correlations for health and fertility ranged from 0.74 (Petersson *et al.* 2005) to unity (Castillo-Juarez *et al.* 2000; Raffrenato *et al.* 2003; Carlen *et al.* 2005). Since most studies that reported environmental variances found changes in environmental variances with a change in environmental predictor, it is important to account for heterogeneous residual variances across environments as well as heterogeneous genetic variances.

The generally high genetic correlations found between environments are somewhat a surprise. One concern was that this might have been caused by the fact that environmental descriptors are often calculated from the same data. For example average milk yield as environmental predictor and milk yield as the trait analysed. Therefore there might be difficulty in disentangling the genetic level of a herd, the phenotypic level of a herd and the genetic level of animals within that environment. (Kolmodin *et al.* 2002) tried to solve this by taking more animals in the herd to calculate the environmental parameter. Another problem typical for animal breeding data is preferential treatment of animals with a high breeding value and selective use of sires in certain environments. Calus (Calus *et al.* 2004) investigated these questions with simulation. He found that there was little effect of the definition of the environmental descriptor on the estimated components. Non-random use of sires, little connectedness and small herd sizes had a large effect on the estimated covariance function. Even with random sire use the less favourable the data, the more the estimated genetic correlation was biased upwards towards unity, i.e. underestimating genotype by environment interaction.

From the initial models used, several authors have made advancements. For example, ASREML allowed further expansion of the single trait reaction norm model to a multi trait reaction norm model (Windig *et al.* 2006). The advantage of the multitrait model is that also the sensitivity of the genetic correlation between traits can be investigated. For example the genetic correlations between milk yield and number of inseminations ranged from 0.18 in large herds to 0.64 in high fertility herds. Also the genetic correlation between yield and SCC differed depending on the success of fertility management in a herd (from 0.25 to 0.47). Another expansion was implemented by including more than one environmental descriptor in the model (Calus *et al.* 2006), e.g. by combining herd bulk tanks SCC, days in milk and their interaction. This model was similar to the covariance function approach used by (Veerkamp and Goddard 1998), albeit Calus et al (2006) were able to estimate the parameters of the covariance function directly from the data. Currently we are developing a model that combines the reaction norm approach with the multitrait approach, which for example, enables to analyse the effect of average mik production in different production systems. Su (2006) developed a Bayesian model that inferred the environmental values simultaneously with the other parameters of the model., which is theoretically more appealing than using a predefined environmental parameters. This model was applied to dairy cattle production data (Shariati *et al.* 2007) and further developed (Su *et al.* 2009) to allow for environmental sensitivity of heterosis. A semi-parametric non-linear longitudinal hierarchical model (Sanchez *et al.* 2009) was suggested by these authors to study longitudinal data, especially when the traits are subject to abrupt changes in due to the environment.

Disturbances by environments on a unknown micro scale. In practice, dairy herds differ in many ways and discrete or continuous environmental parameters describe these differences only in part. These unexplained micro-environmental differences may lead to genetic differences in micro-environmental sensitivity, which is observed as differences in residual variance. When bulls have at least 50-100 progeny, breeding values can be estimated for the size of the residual variance (Mulder et al. 2007). Bulls with progeny that exhibit a large residual variation across herds are the ones that are environmentally sensitive. Bulls with progeny that exhibit a small residual variation across herds are the ones that are not environmentally sensitive.

One of the main problems with analysis of genetic heterogeneity of residual variance is its estimability. Rowe *et al.* (2006) used a two-step approach, in which the first stage used a model to analyze body weight with allowance for differences in residual variance among sire families, and in the second stage the variation among these estimates of the residual variance were analyzed with least squares. In some studies log-transformed squared residuals have been analyzed, also using a two-step approach (Larzul *et al.* 2006; Bolet *et al.* 2007), which gives the flexibility to account for environmental effects on the residual variance at the level of the record that is not possible in the least squares analysis of Rowe *et al.* (2006). Ideally, one would simultaneously fit a model for the mean of the trait and for the residual variance using a structural model. (Sorensen and Waagepetersen 2003; Ros *et al.* 2004; Gutierrez *et al.* 2006) applied these structural models in a Bayesian context, implemented using Markov Chain Monte Carlo sampling (MCMC). Mulder et al (2009) extended the two-step approach of Larzul *et al.* (2006) and Bolet *et al.* (2007) to a bivariate analyses using ASREML to get an estimate of the genetic correlation between the additive genetic effects for the mean and the residual variance. Wolc *et al.* (2009) used a similar approach combined with a generalized linear model for the log-transformed squared residuals, but they were not able to run a bivariate analysis.

There is some empirical evidence that genetic variation in residual variance exists. (SanCristobal-Gaudy *et al.* 2001; Sorensen and Waagepetersen 2003; Ros *et al.* 2004; Gutierrez *et al.* 2006; Ibanez-Escriche *et al.* 2008a,b) used a structural model for heterogeneous residual variance and found substantial genetic variation in residual variance for litter size in sheep, litter

size in pigs, body weight in snails, litter size and weight in mice, body weight traits in mice, and slaughter weight in pigs, respectively. Rowe *et al.* (2006) found substantial genetic heterogeneity of residual variance between sire families in body weight of broilers. Mulder *et al* (2009) found low heritability for residual variance of body weight in broilers (0.03 - 0.05), but a relatively high genetic coefficient of variation (0.35 - 0.57) was estimated. Probably the clearest example is by Mackay and Lyman (2005), who derived 300 isofemale lines of Drosophila melanogaster and found substantial highly significant genetic variance in residual variance between lines under controlled laboratory conditions. Other evidence comes from the selection experiment by Garreau *et al.* (2008) who obtained in a divergent selection experiment with rabbits a clear selection response when selecting on high or low within-litter birth weight. Al these results indicate that environmental sensitivity to unknown environmental disturbances can be improved by means of genetic selection. No estimates or selection experiments have been published for dairy cattle yet. However, if we suppose that also in dairy cattle this genetic variation in residual variance exists, than these genetic differences may be utilized to breed more robust cows, cows that are insensitive to unknown disturbances e.g. for fertility and health traits.

DISCUSSION

A robust dairy cow is able to maintain homeostasis in the commonly accepted and sustainable dairy herds of the near future. Robustness is largely an acquired characteristic through building up experience from exposure to a very large number of minor and major environmental signals. Breeding may give animals an advantage in acquiring robustness. Several methods have been discussed to improve robustness through breeding in practice. Finding re ranking of animals across environments, even with the sophisticated reaction norm models, is difficult, in the scenarios where enough data is available. For production traits it is even questionable if substantial GxE exists, whilst re-ranking for fitness trait has been reported in some studies. The more diverse the systems will be, the more important this GxE becomes for fitness traits. An important strategy will always be to collect data in the same environment we are selecting our breeding animals for. However, this is a practical challenge for dairy cattle breeders, when the speed of change, in for example farm size and scale, continues with the same rate that has been observed over the past 15 years and at the same time the generation interval remains 5 years. Therefore, within the RobustMilk project (www.RobustMilk.eu), there is also attention for measuring robustness traits with mid-infrared spectrometry using a spectra, as part of routine milk analysis. Preliminary analysis indicates that equations developed using the spectra can be used to predict milk fatty acid content (Soyeurt *et al.* 2006) and the question is whether they can be used to also predict dairy cow robustness. i.e. energy balance and immune parameters. Furthermore, in this project rare detailed phenotypes on 2000 cows from four research herds have been brought together, and in the past months all these animals have been genotyped with the 50k SNP array. Statistical analyses of these spectra and genomic data form a challenge on their own, but together with the statistical models describing the genetic variation in environmental insensitivity, they are an essential part in developing a strategy to breed more robust cows suitable for future farming systems.

ACKNOWLEDGMENTS

The RobustMilk project is financially supported by the European Commission under the Seventh Research Framework Programme, under Grant Agreement 211708. The contents of this paper are the sole responsibility of the publishers, and they do not necessarily represent the views of the Commission or its services.

REFERENCES

Bolet, G., Gaffeau, H., Joly, T., Theau-Clement, M., Faheres, J., Hurtaud, J. and Bodin, L. (2007) *Livestock Science* **111**:28

Bryant, J.R., Lopez-Villalobos, N., Pryce, J.E., Holmes, C.W. and Johnson, D.L. (2006) *New Zealand Journal of Agricultural Research* **49**:371

Bryant, J.R., Lopez-Villalobos, N., Pryce, J.E., Holmes, C.W., Johnson, D.L. and Garrick, D.J. (2007) *J. Dairy Sci.* **90**:1538

Calus, M.P.L., Bijma, P. and Veerkamp, R.F. (2004) *Genetics Selection Evolution* **36**:489

Calus, M.P.L., Carrick, M.J., Veerkamp, R.F. and Goddard, M.E. (2005) *J. Dairy Sci.* **88**:1166

Calus, M.P.L., Janss, L.L.G. and Veerkamp, R.F. (2006) *J. Dairy Sci.* **89**:4846

Calus, M.P.L. and Veerkamp, R.F. (2003) *J. Dairy Sci.* **86**:3756

Cannon, W.B. (1932) The wisdom of the body. (Ed. Norton, W.W.)New York).

Carlen, E., Schneider, M.D. and Strandberg, E. (2005) *J. Dairy Sci.* **88**:797

Castillo-Juarez, H., Oltenacu, P.A., Blake, R.W., Mcculloch, C.E. and Cienfuegos-Rivas, E.G. (2000) *J. Dairy Sci.* **83**:807

Dechow, C.D., Rogers, G.W., Cooper, J.B., Phelps, M.I. and Mosholder, A.L. (2007) *J. Dairy Sci.* **90**:3542.

Falconer, D.S. and Mackay, T.F.C. (1996) 'Introduction to quantitative genetics.' Longman: Harlow, UK.

Fikse, W.F., Rekaya, R. and Weigel, K.A. (2003) *Livest. Prod. Sci.* **82**:223

Garreau, H., Bolet, G., Larzul, C., Robert-Granie, C., Saleil, G., Sancristobal, M. and Bodin, L. (2008) *Livestock Science* **119**:55

Gilmour, A.R., Cullis, B.R., Welham, S.J. and Thompson, R. (2000) *NSW Agriculture, Orange Agricultural Institute, Forest Road, Orange, NSW, 2800, Australia.*

Gutierrez, J.P., Nieto, B., Piqueras, P., Ibanez, N. and Salgado, C. (2006) *Genetics Selection Evolution* **38**:445

Ibanez-Escriche, N., Sorensen, D., Waagepetersen, R. and Blasco, A. (2008a) *Genetics* **180**:2209

Ibanez-Escriche, N., Varona, L., Sorensen, D. and Noguera, J.L. (2008b) *Animal* **2**:19

Kolmodin, R., Strandberg, E., Danell, B. and Jorjani, H. (2004) *Acta Agriculturae Scandinavica Section a-Animal Science* **54**:139

Kolmodin, R., Strandberg, E., Madsen, P., Jensen, J. and Jorjani, H. (2002) *Acta Agriculturae Scandinavica Section a-Animal Science* **52**:11

Larzul, C., Le Roy, P., Tribout, T., Gogue, J. and Sancristobal, M. (2006) Canalizing selection on ultimate ph in pigs: Consequences on meat quality. In 'Proc. 8th World Congress Genet. Appl. Livest. Prod.' Belo Horizonte, Brazil.

Mackay, T.F.C. and Lyman, R.F. (2005) *Philosophical Transactions of the Royal Society B-Biological Sciences* **360**:1513

Mcallister, A.J. (2002) *J. Dairy Sci.* **85**:2352

Meuwissen, T.H.E. and Sonesson, A.K. (1998) *J. Anim. Sci.* **76**:2575

Miglior, F., Muir, B.L. and Van Doormaal, B.J. (2005) *J. Dairy Sci.* **88**:1255.

Mulder, H.A., Bijma, P. and Hill, W.G. (2007) *Genetics* **175**:1895.

Mulder, H.A., Hill, W.G., Vereijken, A. and Veerkamp, R.F. (2009) *Animal* **(accepted)**.

Mulder, H.A., Veerkamp, R.F., Ducro, B.J., Van Arendonk, J.A.M. and Bijma, P. (2006) *J. Dairy Sci.* **89**:1740.

Nauta, W.J., Veerkamp, R.F., Brascamp, E.W. and Bovenhuis, H. (2006) *J. Dairy Sci.* **89**:2729.

Oseni, S., Misztal, I., Tsuruta, S. and Rekaya, R. (2004) *J. Dairy Sci.* **87**:3022.

Pedersen, J. and Christensen, L.G. (1989) *Livest. Prod. Sci.* **23**:253.

Petersson, K.-J., Kolmodin, R. and Strandberg, E. (2005) *Acta Agric. Scand., Sect. A, Animal Sci.* **55**:9.

Pryce, J.E., Veerkamp, R.F., Thompson, R., Hill, W.G. and Simm, G. (1997) *Animal Science.* **3**:353.

Raffrenato, E., Blake, R.W., Oltenacu, P.A., Carvalheira, J. and Licitra, G. (2003) *J. Dairy Sci.* **86**:2470.

Rauw, W.M., Kanis, E., Noordhuizen Stassen, E.N. and Grommers, F.J. (1998) *Livest. Prod. Sci.* **56**:15.

Ros, M., Sorensen, D., Waagepetersen, R., Dupont-Nivet, M., Sancristobal, M., Bonnet, J.C. and Mallard, J. (2004) *Genetics* **168**:2089.

Rowe, S.J., White, I.M.S., Avendano, S. and Hill, W.G. (2006) *Genetics Selection Evolution* **38**:617.

Royal, M., Mann, G.E. and Flint, A.P.F. (2000) *Veterinary Journal* **160**:53 .

Sanchez, J.P., Rekaya, R. and Misztal, I. (2009) *Genetics Selection Evolution* **41**.

Sancristobal-Gaudy, M., Bodin, L., Elsen, J.M. and Chevalet, C. (2001) *Genetics Selection Evolution* **33**:249.

Schaeffer, L.R. and Dekkers, J.C.M. (1994) *Proceedings, 5th World Congress on Genetics Applied to Livestock Production, University of Guelph, Guelph, Ontario, Canada, 7 12 August, 1994* **18**:443.

Shariati, M.M., Su, G., Madsen, P. and Sorensen, D. (2007) *J. Dairy Sci.* **90**:5759-5766.

Sorensen, D. and Waagepetersen, R. (2003) *Genet. Res.* **82**:207.

Soyeurt, H., Dardenne, P., Dehareng, F., Lognay, G., Veselko, D., Marlier, M., Bertozzi, C., Mayeres, P. and Gengler, N. (2006) *J. Dairy Sci.* **89**:3690-3695.

Su, G., Madsen, P. and Lund, M.S. (2009) *J. Dairy Sci.* **92**:2204.

Su, G., Madsen, P., Lund, M.S., Sorensen, D., Korsgaard, I.R. and Jensen, J. (2006) *J. Anim. Sci.* **84**:1651.

Tada, T. (1997) *Annual Review of Immunology* **15**:1.

Ten Napel, J., Bianchi, F. and Bestman, M. (2006) In 'Working papers 1: Inventions for a sustainable development of agriculture.' pp. 32. (Transforum: Zoetemeer).

Ten Napel, J., Calus, M.P.L., Mulder, H.A. and Veerkamp, R.F. (2009) In 'Breeding for robustness in cattle'. (Ed. Marija Klopcic, R.R., Jan Philipsson and Abele Kuipers) p. 288. (EAAP Scientific Series - ISSN 0071.

Veerkamp, R.F. and Goddard, M.E. (1998) *J. Dairy Sci.* **81**:1690.

Wagter, L.C., Mallard, B.A., Wilkie, B.N., Leslie, K.E., Boettcher, P.J. and Dekkers, J.C.M. (2000) *J. Dairy Sci.* **83**:488.

Windig, J.J., Calus, M.P.L., Beerda, B. and Veerkamp, R.F. (2006) *J. Dairy Sci.* **89**:1765.

Windig, J.J., Calus, M.P.L., De Jong, G. and Veerkamp, R.F. (2005a) *Livest. Prod. Sci.* **96**:291.

Windig, J.J., Calus, M.P.L. and Veerkamp, R.F. (2005b) *J. Dairy Sci.* **88**:335-347.

Wolc, A., White, I.M.S., Avendano, S. and Hill, W.G. (2009) *Poultry Science* **88**:1156.

Sheep - Meat

MEAT SHEEP BREEDING – WHERE WE ARE AT AND FUTURE CHALLENGES

N. M. Fogarty

NSW Department of Primary Industries, Orange Agricultural Institute, Orange, NSW 2800

SUMMARY

Developments in meat sheep breeding in Australia over the last 40 years are reviewed. This includes the evolution of LAMBPLAN and its implementation in the industry, development of breeding objectives and estimation of genetic parameters. The development of indexes and the importance of major genes for both meat and maternal traits are discussed as well as strategies for combined improvement of wool and meat in Merino enterprises. Opportunities and challenges for breeding in the future are considered.

INDUSTRY BACKGROUND

Performance recording programs to assist meat sheep breeders were first developed in the late 1960s by the NSW, Victorian and SA Departments of Agriculture and the University of NSW (Pattie 1973). Their development was advocated as a means of increasing the rate of genetic improvement in economically important traits, by the Animal Production Committee Expert Panel (APC 1970) and by several workshops over the next decade, although they failed to attract widespread usage by breeders. However in the late 1970s Dorset breeders showed renewed interest in production testing following industry developments that included: demonstration of within flock variation in measured growth and fatness in Dorset production competitions (Fogarty and Harris 1975; Clements and Fogarty 1976); increased focus on carcase weight and fat in carcase classification and lamb marketing (Moxham and Brownlie 1976); strong consumer preference for leaner cuts (Thatcher 1982); use of fleece measurement in Merino breeding (McGuirk 1978); development of technology to accurately measure fat depth in live animals (Thompson *et al.* 1977; Clements *et al.* 1981); and the success of Sheeplan in New Zealand (Clarke 1979). The NSW Meat Sheep Testing Service (MSTS) was implemented in 1980 with widespread support from Dorset and other terminal sire stud breeders in NSW (Harris 1985) and it expanded to testing over 17,000 sheep from 120 studs annually (Fogarty *et al.* 1987). The program of R&D and advisory support was run by the NSW Department of Agriculture from Cowra with financial support from the Australian Meat and Livestock Research and Development Corporation (now Meat and Livestock Australia, MLA). The objective of the project was to "evaluate the development of a viable facility enabling meatsheep stud breeders to objectively test rams for genetic differences in growth rate and fat depth". The project developed practical procedures for measuring fat depth using real-time ultrasound technology, accumulated a large database and provided a model for the development of LAMBPLAN, the national genetic evaluation program that was launched in 1989 (Banks 1990). A procedure was also developed for measuring eye muscle depth using real-time ultrasound in live animals (Gilmour *et al.* 1994) and included in LAMBPLAN (Fogarty *et al.* 1992a).

Genetic improvement of growth and leanness of terminal sire rams using LAMBPLAN was a key element in the Elite Lamb R&D Program (Thatcher 1992) and Strategic Plan adopted by the industry in the 1990s to produce large lean lambs. A national program of central progeny testing of terminal sires in the early 1990s (Banks *et al.* 1995) contributed to the adoption of LAMBPLAN (Banks 1994). Studies also demonstrated that rams with a range of LAMBPLAN estimated breeding values (EBVs) produced lambs with predictable performance (Hall *et al.* 1995; Hall *et al.* 2002) and there was considerably more variation between individual sires than between terminal sire breeds for growth (Fogarty *et al.* 2000a) and carcase (Fogarty *et al.* 2000b) traits. Buyers also began to pay a premium at auctions for flock rams with high EBVs (Ferguson and Fogarty 1997).

More sophisticated software was developed using an animal model and BLUP procedures to estimate EBVs across flocks and years (Gilmour and Banks 1992). For the first time in Australia, breeders were able to compare the genetic merit of sheep in different flocks, which led to a greater uptake of LAMBPLAN, with over 50% of terminal sires in the industry being tested by 1994 (Banks 1994). The larger data sets required new software (Brown *et al.* 2000) and a joint program of MLA and Australian Wool Innovation Ltd (AWI) incorporated Merinos into the national sheep genetic evaluation system (Brown *et al.* 2006, 2007). The Sheep Genetics databases include records on over 1.3m terminal sire, 0.3m maternal and 1.1m Merino animals (Brown *et al.* 2007). There has been substantial improvement in all breeds between 1990 and 2005, with terminal sires increasing by $17/ewe (2.9 s.d.) and an increased genetic trend since 2000 (Swan *et al.* 2009).

While uptake of the new genetic technology was high in the terminal sire sector of the industry in the 1990s, the maternal breeding sector was lagging, despite having EBVs and genetic information available for the economically important reproduction and wool traits (Fogarty *et al.* 1992a). Productivity of the crossbred ewe flock has a major impact on the profitability of lamb enterprises and the task was to achieve greater genetic improvement among the commercial flocks of crossbred ewes. To address this issue the Maternal Sire Central Progeny Test (MCPT) project commenced in 1997 (Fogarty *et al.* 1999), to evaluate and demonstrate the variation in first and second cross progeny performance of maternal and dual purpose (wool and meat) sires and the scope for genetic improvement in the sector. The MCPT demonstrated that there was a range of over $40 gross margin/ewe/year between first cross ewe sire progeny groups (Fogarty *et al.* 2005).

Recently considerable investment has been made in SheepGenomics to "find useful genes and put them to work" (Oddy *et al.* 2007). A resource population has been developed to find quantitative trait loci (QTL) and a greater understanding of functional genomics. The Cooperative Research Centre (CRC) for Sheep Industry Innovation has also developed the Information Nucleus (IN), which progeny tests key young industry sires for an extensive range of traits in widely differing environments (Fogarty *et al.* 2007a). This allows breeders and commercial producers to exploit new technology and genomic information to achieve more rapid genetic improvement.

BREEDING OBJECTIVES AND STATISTICAL DEVELOPMENTS

Studies showed the importance of including liveweight (Stafford and Walkley 1979) and carcase fat (Atkins 1987; Clarke *et al.* 1991) in the breeding objective for meatsheep. LAMBPLAN initially provided within-flock EBVs for weight and fat depth based on live animal measurements and information from correlated traits and relatives (Banks 1990). Eye muscle depth was subsequently included (Fogarty *et al.* 1992a). The breeding objectives for maternal meatsheep breeds include reproduction and wool, in addition to weight, fat and muscle (Fogarty 1987). In 1992 there was a major enhancement of the statistical procedures used for estimation of genetic merit in LAMBPLAN with the development of animal model BLUP software, BVEST (Gilmour and Banks 1992). BVEST was designed to perform on-farm breeding value estimation for individual cohorts of animals and off-farm across flock and across year estimation, using performance information from relatives and all available pedigree records. Muscle depth and maternal traits were also included in this enhancement.

New software (OVIS) was developed in the late 1990s to handle the large data sets, include more traits and breeds and incorporate new features such as an expanded model with maternal and permanent environmental components and genetic grouping (Brown *et al.* 2000). Subsequently, significant development was required to incorporate Merinos to deliver a single national across-flock genetic evaluation system to the Australian sheep industry. These innovations included the amalgamation of databases, data transformations, refinement of analysis models, genetic grouping methodology, updated genetic parameters, multiple trait across-flock linkage assessments, index development, use of a common technical language and changes to reporting (Brown e*t al.* 2007).

GENETIC PARAMETERS

The MSTS dataset was used to validate adjustment procedures (Fogarty and Luff 1985) and provided the first genetic parameter estimates for weight and fat depth in Australian Poll Dorset sheep (Atkins *et al.* 1991) and several other meat and dual purpose breeds (Brash *et al.* 1992). A major program was undertaken to estimate genetic parameters for a range of other economically important traits using research and stud data sets from Border Leicester (Brash *et al.* 1994a), Corriedale (Brash *et al.* 1994b), Coopworth (Brash *et al.* 1994c) and Hyfer (Fogarty *et al.* 1994) breeds, so that breed-specific parameters could be used for calculating EBVs (Fogarty *et al.* 1992a). Genetic parameters in Australian Dorset sheep were also estimated for reproduction traits (Brash *et al.* 1994d), eye muscle depth in live animals (Gilmour *et al.* 1994) and an extensive range of carcase and meat quality traits (Kenney *et al.* 1995). These estimates together with a review of the world literature provided the basis for the early genetic parameter matrix used in LAMBPLAN (Fogarty 1995).

Safari and Fogarty (2003) tabulated 164 reports of sheep genetic parameters for a range of traits published in the world literature over the previous decade. These reports provided weighted means of the parameter estimates for the traits in a review (Safari *et al.* 2005). The review showed there were numerous estimates of heritability, which were reasonably consistent, for wool, growth and, to a lesser extent, reproduction traits, although there were few estimates for carcase and meat traits. The review also showed that while there were several estimates of genetic correlations among the various wool and growth traits, there were very few, if any, among the other trait groups or between the various trait groups. To address this dearth of genetic correlation parameters the Australian Sheep Industry CRC and owners of several research data sets supported a combined analysis. Seven Merino research data sets were combined (Safari *et al.* 2007a) and heritabilities (Safari *et al.* 2007b) and genetic correlations (Safari *et al.* 2007c) among and between a range of wool, growth and reproduction traits were estimated with high accuracy. In addition, genetic parameters were estimated for carcase and meat quality traits (Fogarty *et al.* 2003b; Greeff *et al.* 2008), as well as genetic correlations between carcase and meat quality traits and growth and wool traits (Greeff *et al.* 2008) and ewe reproduction traits (Safari *et al.* 2008). Additional parameter estimates have recently been published for fine wool Merinos (Swan *et al.* 2008) and Merino flocks recorded in the Sheep Genetics database (Huisman *et al.* 2008; Huisman and Brown 2008, 2009a, 2009b). The MCPT data set was also used to estimate genetic parameters for lamb growth, carcase and meat quality, wool production, worm egg count (Ingham *et al.* 2007), feed intake (Fogarty *et al.* 2006a), reproduction (Afolayan *et al.* 2008b), insulin-like growth factor-1 (IGF-1) (Afolayan and Fogarty 2008) and milk production (Afolayan *et al.* 2009c) in first cross animals, together with the genetic correlations between ewe reproduction traits and early growth and wool production (Afolayan *et al.* 2009a), growth and carcase performance of their progeny (Afolayan *et al.* 2008a) and between milk and other production traits of the ewes (Afolayan *et al.* 2009b).

MEAT TRAITS

Indexes. LAMBPLAN initially offered a range of simple indexes to assist in selection of animals for a combination of increased growth rate (post weaning weight) and decreased subcutaneous fat (Banks 1990; Fogarty *et al.* 1992a). The option of eye muscle depth as an additional trait in these indexes for terminal sires was subsequently included (Banks 1994). A wider range of indexes is now available, with two specifically designed dollar indexes for breeding objectives to meet the domestic (20-22 kg carcase weight) and export (24+ kg carcase weight) markets (Brown *et al.* 2000, 2007). Recently the LAMB 2020 dollar index was launched by Sheep Genetics, which as well as combining weaning weight, post weaning weight, leanness and muscle depth, includes a negative emphasis on birth weight and increased resistance to worms (Ball 2008).

Major genes. In the early 1990s a major gene (callipyge) was shown to increase hindquarter muscling and meat toughness in Dorset sheep in the USA (Koohmaraie *et al.* 1995). Sires from Australian Dorset flocks were also shown to carry a similar gene (Carwell), although its effects on loin muscling and tenderness were much smaller than callipyge (Hopkins and Fogarty 1998b). The Carwell allele increases eye muscle weight by 8% and area by 10% (Nicoll *et al.* 1998). The gene is maternally imprinted and has been mapped to a region near callipyge and a marker test is available (Dodds 2007). Another gene has also been found which increases leg muscling and reduces fat in Texel sheep and a marker test is available. It appears to be additive and may be associated with myostatin (Dodds 2007).

MATERNAL TRAITS

Indexes. LAMBPLAN was enhanced in the early 1990s to include reproduction and wool traits, as well as growth, fat and muscle, for maternal and dual-purpose breeds (Fogarty *et al.* 1992a). Reproduction is more important in maternal than Merino breeding objectives and was sensitive to prices and varied in importance for different dual purpose breeds (Fogarty and Gilmour 1993). It is also important to take into account the feed requirements of the maternal flock as well as the lamb progeny in overall enterprise profitability (Fogarty *et al.* 2003a). Dollar indexes are available in LAMBPLAN which are customised for each of the maternal breed groups (Brown *et al.* 2007). Selection on litter size or ovulation rate to increase reproduction was advocated in the 1970s and 1980s, however there is considerable genetic variance for all components of reproduction (Safari *et al.* 2005) and use of a selection index of overall ewe productivity or litter weight weaned may result in a more balanced biological outcome (Snowder and Fogarty 2009). There are several reports demonstrating realised response to selection for litter weight weaned (Fogarty 1994; Ercanbrack and Knight 1998; Cloete *et al.* 2004).

Major genes. A series of experiments in Australia and New Zealand in the early 1980s confirmed the high prolificacy of the Booroola Merino was due to the segregation of a major gene (FecB) (Davis *et al.* 1982; Piper *et al.* 1985), with a molecular test now available (Davis 2004). A recent review of 40 studies in a range of genetic comparisons, environments and production systems (Fogarty 2009) showed the effect of heterozygous (B+) versus non-carrier (++) ewes was +1.1 to +2.0 for ovulation rate (with BB generally being additive) and +0.5 to +1.3 for litter size (with little additional effect for BB). Poor lamb survival and growth, due largely to higher litter size, further reduced the effect for lambs weaned and weight of lamb weaned. Poor lamb survival and associated low birth weight and growth have been major barriers to industry uptake in Australia despite FecB being introgressed into research and commercial flocks. Several other genes that have major effects on ovulation rate have also been reported in overseas breeds (Davis 2004).

COMBINING MEAT AND WOOL

Meat and wool have long been regarded as separate industries, with prime lamb production based on crossbred progeny of terminal sires and first cross dams and apparel wool being the preserve of the Merino (Pattie 1973). There are different breeding objectives for terminal sires (Atkins 1987), prime lamb dams (Fogarty 1987) and Merinos (Walkley 1987), although the Merino has always contributed a majority of genes to the national lamb slaughter through second cross, first cross and straightbred Merino lambs (Fogarty *et al.* 2000a). While Merino and first cross lambs have lower growth rates than second cross lambs (Fogarty *et al.* 2000a; Hopkins *et al.* 2007), there is little difference in their carcase and meat quality performance when grown under the same conditions and compared at the same carcase weights (Hopkins and Fogarty 1998a, 1998b; Fogarty *et al.* 2000b; Ponnampalam *et al.* 2007).

The increasing demand and economic value of lamb and the relative decline in value of wool has meant that more Merino ewes are being mated to terminal sires and many Merino breeders wish to include meat traits in more complex breeding objectives, although development of a dual-purpose sheep is not recommended (van der Werf 2006). There are no major genetic antagonisms between meat and wool traits and improvement can be achieved in both using an appropriate selection index (Fogarty *et al.* 2006b). The genetic parameters (heritabilities, variances and genetic correlations) estimated in Merino and crossbred sheep in the reports noted above now provide comprehensive information for developing more complex breeding objectives and selection criteria for combining meat and wool traits.

OPPORTUNITIES AND CHALLENGES

Meat quality. A key element in the revolution that has occurred in the lamb industry over the last two decades has been improvement in products and quality. Subcutaneous fat has been reduced and muscle size increased while maintaining a high level of eating quality. However selection for muscling can result in structural and biochemical changes to muscle, with less aerobic muscle, less intramuscular fat and sometimes a reduction in tenderness (Pethick *et al.* 2006). It is critical to develop a better understanding of these potentially detrimental effects on meat quality (part of the Meat Program in the Sheep CRC) and ensure appropriate breeding programs are implemented by industry. There will also be increasing consumer interest in meat products that meet specific health standards or confer particular human health benefits (Bermingham *et al.* 2008). An example may be increasing the level of long-chain omega-3 fatty acids in lamb (Pethick *et al.* 2006).

Fitness and reproduction. Net reproduction is low in most Australian and particularly Merino flocks. Improving post natal and embryo survival is critical as ovulation rate is generally not limiting. Poor lamb survival is an animal welfare issue for the industry as well as one of reduced productivity. There is some evidence of genetic variance and heterosis for embryo survival (see Fogarty 2009), although further research needs to be undertaken. While most reports have shown little genetic variance for lamb survival, there is more variation for ewe rearing ability or lamb survival as a trait of the ewe (Safari *et al.* 2005). Recent analysis of a large Merino data set has shown similar results with repeated records of ewe rearing ability, especially for survival to 7 days, being able to increase selection accuracy and improve current generation performance (S Hatcher pers. comm.). There is also genetic variance for adult longevity (Hatcher *et al.* 2009).

Other opportunities for improving reproductive efficiency in the lamb industry include mating ewes for the first time at an earlier age and accelerated lambing systems. Crossbred ewes can be successfully joined in the autumn to lamb in their first year. Puberty and lambing performance in ewe lambs is influenced by both genetic and environmental factors and ewes that rear lambs in their first year rear more lambs in subsequent years (Fogarty *et al.* 2007b). Accelerated lambing systems can increase annual lamb production of ewes (Fogarty *et al.* 1992b) and improvements can be made by selection (Fogarty 1994).

Whole genome selection. Recent advances in sequencing the sheep genome has opened up the opportunity to exploit whole genome selection (Meuwissen *et al.* 2001). Genomic (G) EBVs can be calculated from the information on thousands of single nucleotide polymorphisms (SNP) that will be available in the near future (Oddy *et al.* 2007). In dairy cattle GEBVs have been shown to be highly reliable and are being used to improve the rate of genetic gain (Hayes *et al.* 2009). The task in sheep is to quantify the effects of useful SNP and to validate these in wider industry sheep populations that have relevant phenotypic data. The SheepGenomics (Oddy *et al.* 2007) and CRC Information Nucleus (Fogarty *et al.* 2007a) are important resource flocks in this quest.

Combining quantitative and genomic information. There are a large number of gene markers (van der Werf *et al.* 2007) and DNA tests available (Dodds *et al.* 2007) for disease and production traits in sheep. Strategies need to be developed to effectively combine the genomic and quantitative information. Davis *et al.* (2006) highlighted some of the problems such as a high merit ram for multigenic traits not being indicative of progeny merit in the presence of a segregating major gene. However, the imminent availability of genotype data for a large number of SNP may make the analysis of the data more straight forward (van der Werf *et al.* 2007).

CONCLUSIONS

Genetic improvement, especially in growth, leanness and muscling, has been a key element in the dramatic increase in productivity and profitability of the meat sheep industry over recent decades. These changes have occurred through development of LAMBPLAN, which has been made possible by the ongoing R&D support, and its widespread adoption by industry.

REFERENCES

Afolayan, R.A. and Fogarty, N.M. (2008) *J. Anim. Sci*. **86**:2068.

Afolayan, R.A., Fogarty, N.M., Gilmour, A.R., Ingham, V.M., Gaunt, G.M. and Cummins, L.J. (2008a) *Small Rumin. Res.* **80**:73.

Afolayan, R.A., Fogarty, N.M., Gilmour, A.R., Ingham, V.M., Gaunt, G.M. and Cummins, L.J. (2008b) *J. Anim. Sci.* **86**:804.

Afolayan, R.A., Fogarty, N.M., Gilmour, A.R., Ingham, V.M., Gaunt, G.M. and Cummins, L.J. (2009a) *Anim. Prod. Sci.* **49**:17.

Afolayan, R.A., Fogarty, N.M., Morgan, J.E., Gaunt, G.M., Cummins, L.J. and Gilmour, A.R. (2009b) *Small Rumin. Res.* **82**:27.

Afolayan, R.A., Fogarty, N.M., Morgan, J.E., Gaunt, G.M., Cummins, L.J., Gilmour, A.R. and Nielsen, S. (2009c) *Anim. Prod. Sci.* **49**:24.

APC (1970) *J. Aust. Inst. Agric. Sci.* **36**:30.

Atkins, K.D. (1987) *Proc. Aust. Assoc. Anim. Breed. Genet.* **6**:221.

Atkins, K.D., Murray, J.I., Gilmour, A.R. and Luff, A.L. (1991) *Aust. J. Agric. Res.* **42**:629.

Ball, A.J. (2008) In 'The Breeder's Bulletin', Spring p. 5., MLA and AWI, Armidale.

Banks, R.G. (1990) *Proc. Aust. Assoc. Anim. Breed. Genet.* **8**:237.

Banks, R.G. (1994) *Proc. 5th Wld. Congr. Genet. Appld. Livest. Prod.*, Guelph, Canada. **18**:15.

Banks, R.G., Shands, C., Stafford, J.E., and Kenney, P. (1995) 'LAMBPLAN superior sires', Meat Research Corporation, Sydney

Bermingham, E.N., Roy, N.C., Anderson, R.C., Barnett, M.P.G., Knowles, S.O. and McNabb, W.C. (2008) *Aust. J. Exp. Agric.* **48**:726 .

Brash, L.D., Fogarty, N.M., Gilmour, A.R. and Luff, A.F. (1992) *Aust. J. Agric. Res.* **43**:831.

Brash, L.D., Fogarty, N.M., Barwick, S. and Gilmour, A.R. (1994a) *Aust. J. Agric. Res.* **45**:459.

Brash, L.D., Fogarty, N.M. and Gilmour, A.R. (1994b) *Aust. J. Agric. Res.* **45**:469.

Brash, L.D., Fogarty, N.M. and Gilmour, A.R. (1994c) *Aust. J. Agric. Res.* **45**:481.

Brash, L.D., Fogarty, N.M. and Gilmour, A.R. (1994d) *Aust. J. Agric. Res.* **45**:427.

Brown, D., Tier, B., Reverter, A., Banks, R. and Graser, H. (2000) *Int. J. Sheep Wool Sci.* **48**:285.

Brown, D.J., Ball, A.J., Huisman, A.E., Swan, A.A., Atkins, K.D., Graser, H., Banks, R., Swan, P. and Woolaston, R. (2006) *Proc. 8th Wld. Congr. Genet. Appld. Livest. Prod.* CD-ROM 05-03.

Brown, D.J., Huisman, A.E., Swan, A.A., Graser, H.U., Woolaston, R.R., Ball, A.J., Atkins, K.D. and Banks, R.G. (2007) *Proc. Assoc. Advmnt. Anim. Breed. Genet.* **17**:187.

Clarke, J.N. (1979) *Proc. Aust. Assoc. Anim. Breed. Genet.* **1**:397.

Clarke, J.N., Waldron, D.F. and Rae, A.L. (1991) *Proc. Aust. Assoc. Anim. Breed. Genet.* **9**:265.

Clements, B.W. and Fogarty, N.M. (1976) *Proc. Aust. Soc. Anim. Prod.* **11**:49.

Clements, B.W., Thompson, J.M., Harris, D.C. and Lane, J.G. (1981) *Aust. J. Expt. Agric. Anim. Husb.* **21**:566.
Cloete, S.W.P., Gilmour, A.R., Olivier, J.J. and van Wyk, J. (2004) *Aust. J. Exp. Agric.* **44:**745.
Davis, G.H. (2004) *Anim. Reprod. Sci.* **82-83**:247.
Davis, G., Montgomery, G., Allison, A., Kelly, R. and Bray, A. (1982) *N. Z. J. Agric. Res.* **25**:525.
Davis, G.H., McEwan, J.C. and Dodds, K.G. (2006) *Proc. 8th Wld. Congr. Genet. Appld. Livest. Prod.*, Belo Horizonte, MG Brazil. CD-ROM 04-01.
Dodds, K.G. (2007) *Proc. Assoc. Advmnt. Anim. Breed. Genet.* **17**: 418.
Dodds, K.G., McEwan, J.C. and Davis, G.H. (2007) *Small Rumin. Res.* **70**:32.
Ercanbrack, S.K. and Knight, A.D. (1998) *J. Anim. Sci.* **76**:1311.
Ferguson, B.D. and Fogarty, N.M. (1997) *Proc. Assoc. Advmnt. Anim. Breed. Genet.* **12**:360.
Fogarty, N.M. (1987) *Proc. Aust. Assoc. Anim. Breed. Genet.* **6**:217.
Fogarty, N.M. (1994) *Proc. 5th Wld. Congr. Genet. Appld. Livest. Prod.*, Guelph, Canada. **18**: 79.
Fogarty, N.M. (1995) *Anim. Breed. Abstr.* **63**:101.
Fogarty, N.M. (2009) In 'International Booroola Workshop', Nov. 2008, Pune, India. ACIAR, Canberra, Australia. (in press)
Fogarty, N.M. and Harris, D.C. (1975) *Agric. Gaz. NSW* **86**:32.
Fogarty, N.M. and Luff, A.F. (1985) *Proc. Aust. Assoc. Anim. Breed. Genet.* **5**:225.
Fogarty, N.M., Atkins, K., Harris, D. and Luff, A. (1987) 'Final Report DAN 23S', MLA, Sydney.
Fogarty, N.M., Banks, R.G., Gilmour, A.R. and Brash, L.D. (1992a) *Proc. Aust. Assoc. Anim. Breed. Genet.* **10**:63.
Fogarty, N.M., Hall, D.G. and Atkinson, W.R. (1992b) *Aust. J. Agric. Res.* **43**:1819.
Fogarty, N.M. and Gilmour, A.R. (1993) *Aust. J. Exp. Agric.* **33**:259.
Fogarty, N.M., Brash, L.D. and Gilmour, A.R. (1994) *Aust. J. Agric. Res.* **45**:443.
Fogarty, N.M., Cummins, L.J., Stafford, J.E., Gaunt, G. and Banks, R.G. (1999) *Proc. Assoc. Advmnt. Anim. Breed. Genet.* **13**:78.
Fogarty, N.M., Hopkins, D.L. and van de Ven, R. (2000a) *Anim. Sci.* **70**:135.
Fogarty, N.M., Hopkins, D.L. and van de Ven, R. (2000b) *Anim. Sci.* **70**:147.
Fogarty, N., McLeod, L., Morgan, J. (2003a) *Proc. Assoc. Advmnt. Anim. Breed. Genet.* **15**:314.
Fogarty, N.M., Safari, E., Taylor, P.J. and Murray, W. (2003b) *Aust. J. Agric. Res.* **54**:715.
Fogarty, N.M., Ingham, V.M., McLeod, L., Gaunt, G.M. and Cummins, L.J. (2005) *Proc. Assoc. Advmnt. Anim. Breed. Genet.* **16**:60.
Fogarty, N.M., Lee, G.J., Ingham, V.M., Gaunt, G.M. and Cummins, L.J. (2006a) *Aust. J. Agric. Res.* **57**:1037.
Fogarty, N.M., Safari, E., Gilmour, A.R., Ingham, V.M., Atkins, K.D., Mortimer, S.I., Swan, A.A., Brien, F.D., Greeff, J.C. and van der Werf, J.H.J. (2006b) *Int. J. Sheep Wool Sci.* **54**:22.
Fogarty, N.M., Banks, R.G., van der Werf, J.H.J., Ball, A.J. and Gibson, J.P. (2007a) *Proc. Assoc. Advmnt. Anim. Breed. Genet.* **17**:29.
Fogarty, N.M., Ingham, V.M., Gilmour, A.R., Afolayan, R.A., Cummins, L.J., Edwards, J.E.H. and Gaunt, G.M. (2007b) *Aust. J. Agric. Res.* **58**:928.
Gilmour, A.R. and Banks, R.G. (1992) *Proc. Aust. Assoc. Anim. Breed. Genet.* **10**:543.
Gilmour, A.R., Luff, A.F., Fogarty, N.M. and Banks, R. (1994) *Aust. J. Agric. Res.* **45**:1281.
Greeff, J.C., Safari, E., Fogarty, N.M., Hopkins, D.L., Brien, F.D., Atkins, K.D., Mortimer, S.I. and van der Werf, J.H.J. (2008) *J. Anim. Breed. Genet.* **125**:205.
Hall, D., Luff, A., Fogarty, N. and Holst, P. (1995) *Proc. Aust. Assoc. Anim. Breed. Genet.* **11**:185.
Hall, D.G., Gilmour, A.R., Fogarty, N.M. and Holst, P.J. (2002) *Aust. J. Agric. Res.* **53**:1341.
Harris, D.C. (1985) *Proc. Aust. Assoc. Anim. Breed. Genet.* **5**:120.
Hatcher, S., Atkins, K. and Thornberry, K. (2009) *Proc. Assoc. Advmnt. Anim. Breed. Genet.* **18**: 580.

Hayes, B.J., Bowman, P.J., Chamberlain, A.J. and Goddard, M.E. (2009) *J. Dairy Sci.* **92**: 433.
Hopkins, D.L. and Fogarty, N.M. (1998a) *Meat Sci.* **49**:459.
Hopkins, D.L. and Fogarty, N.M. (1998b) *Meat Sci.* **49**:477.
Hopkins, D.L., Stanley, D., Martin, L. and Gilmour, A.R. (2007) *Aust. J. Exp. Agric.* **47**:1119.
Huisman, A.E. and Brown, D.J. (2008) *Aust. J. Exp. Agric.* **48**:1186.
Huisman, A.E., Brown, D.J., Ball, A.J. and Graser, H.U. (2008) *Aust. J. Exp. Agric.* **48**:1177.
Huisman, A.E. and Brown, D.J. (2009a) *Anim. Prod. Sci.* **49**:283.
Huisman, A.E. and Brown, D.J. (2009b) *Anim. Prod. Sci.* **49**:289.
Ingham, V.M., Fogarty, N.M., Gilmour, A.R., Afolayan, R.A., Cummins, L.J., Gaunt, G.M., Stafford, J. and Edwards, J.E.H. (2007) *Aust. J. Agric. Res.* **58**:839.
Kenney, P.A., Goddard, M.E. and Thatcher, L.P. (1995) *Aust. J. Agric. Res.* **46**:703.
Koohmaraie, M., Shackelford, S.D., Wheeler, T.L., Lonergan, S.M. and Doumit, M.E. (1995) *J. Anim. Sci.* **73**:3596.
McGuirk, B.J. (1978) *Wool Technol. Sheep Breed.* **26:**17.
Meuwissen, T.H.E., Hayes, B. and Goddard, M.E. (2001) *Genetics* **157**:1819.
Moxham, R.W. and Brownlie, L.E. (1976): In 'Proc. Symposium on Carcase Classification'
Nicoll, G.B., Burkin, H.R., Broad, T.E., Jopson, N.B., Greer, G.J., Bain, W.E., Wright, C.S., Dodds, K.G., Fennessy, P.F. and McEwan, J.C. (1998) *Proc. 6th Wld. Congr. Genet. Appld. Livest. Prod.*, Armidale, Australia, **26**:529.
Oddy, V.H., Dalrymple, B., McEwan, J.C., Kijas, J., Hayes, B., van der Werf, J.H.J., Emery, D., Hynd, P.I., Longhurst T. , Fischer, T., Ferguson, D., Forage, R., Cockett, N.E. and Nicholas, F.W. (2007) *Proc. Assoc. Advmnt. Anim. Breed. Genet.* **17**:411.
Pattie, W. A. (1973). In 'The Pastoral Industries of Australia' G. Alexander and O. B. Williams (Eds.) pp. 303-335. Sydney University Press, Sydney.
Pethick, D.W., Banks, R.G., Hales, J. and Ross, I.R. (2006) *Int. J. Sheep Wool Sci.* **54**:66.
Piper, L.R., Bindon, B.M. and Davis, G.H. (1985). In 'Genetics of Reproduction in Sheep', Land R.B. and D. W. Robinson (Eds.) pp. 115-125. Butterworths, London, UK.
Ponnampalam E, Hopkins D, Butler K, Dunshea F, Warner R (2007) *Aust. J. Exp. Agric.* **47**:1147.
Safari, A. and Fogarty, N.M. (2003). Technical Bulletin 49. NSW Agriculture, Orange, Australia.
Safari, E., Fogarty, N.M. and Gilmour, A.R. (2005) *Livest. Prod. Sci.* **92**:271.
Safari, E., Fogarty, N.M., Gilmour, A.R., Atkins, K.D., Mortimer, S.I., Swan, A.A., Brien, F.D., Greeff, J.C. and van der Werf, J.H.J. (2007a) *Aust. J. Agric. Res.* **58**:169.
Safari, E., Fogarty, N.M., Gilmour, A.R., Atkins, K.D., Mortimer, S.I., Swan, A.A., Brien, F.D., Greeff, J.C. and van der Werf, J.H.J. (2007b) *Aust. J. Agric. Res.* **58:**177.
Safari, E., Fogarty, N.M., Gilmour, A.R., Atkins, K.D., Mortimer, S.I., Swan, A.A., Brien, F.D., Greeff, J.C. and van der Werf, J.H.J. (2007c) *J. Anim. Breed. Genet.* **124**:65.
Safari, E., Fogarty, N.M., Hopkins, D.L., Greeff, J.C., Brien, F.D., Atkins, K.D., Mortimer, S.I., Taylor, P.J. and van der Werf, J.H.J. (2008) *J. Anim. Breed. Genet.* **125**:397.
Snowder, G.D. and Fogarty, N.M. (2009) *Anim. Prod. Sci.* **49**:9.
Stafford, J.E. and Walkley, J.R.W. (1979) *Proc. Aust. Assoc. Anim. Breed. Genet.* **1**:337.
Swan, A.A., Purvis, I.W. and Piper, L.R. (2008) *Aust. J. Exp. Agric.* **48**:1168.
Swan, A.A., Brown, D. J. and Banks, R.G. (2009) *Proc. Assoc. Advmnt. Anim. Breed. Genet.* **18**:326.
Thatcher, L.P. (1982) *Proc. Aust. Soc. Anim. Prod.* **14**:47.
Thatcher, L.P. (1992) *Proc. Aust. Soc. Anim. Prod.* **19**:173.
Thompson, J.M., Pattie, W. and Butterfield, R. (1977) *Aust. J. Expt. Agric. Anim. Husb.* **17**:251.
van der Werf, J.H.J. (2006) *Int. J. Sheep Wool Sci.* **54:** 17.
van der Werf, J.H.J., Marshall, K. and Lee, S. (2007) *Small Rumin. Res.* **70**:21.
Walkley, J.R.W. (1987) *Proc. Aust. Assoc. Anim. Breed. Genet.* **6**: 207.

Sheep - Meat

RATES OF GENETIC GAIN IN NEW ZEALAND SHEEP

M.J. Young[1] and P.R. Amer[2]

[1]SIL, Meat & Wool New Zealand, PO Box 39-085, Harewood, Christchurch 8545, New Zealand
[2]AbacusBio Limited, PO Box 5585, Dunedin 9058, New Zealand

SUMMARY

Rates of genetic gain in New Zealand sheep doubled in 140 Dual Purpose flocks, from 30 to 60 cents/yr, and 62 Terminal Sire flocks, from 35 to 60 cents/yr, after establishment of the Sheep Improvement Ltd (SIL) genetic evaluation system in 1999. Further incremental gains of a similar magnitude led to triple the rates of gain pre-SIL after establishment of the Meat & Wool New Zealand Central Progeny Test and the start of SIL-ACE across-flock, across-breed genetic evaluations in 2004. Flocks were classified based on genetic merit for 2004-2006 as; "**Leader**" high merit / high rate of gain; "**Gaining**" below average merit / high rate of gain; "**Slipping**" high merit / low rate of gain; and "**Trailing**" below average merit / low rate of gain. Leader and Gaining flocks were more likely to make use of high merit sires from other flocks. Slipping and Trailing flocks were more likely to make repeated use of older sires of low genetic merit. These data show that high rates of gain are achievable by flocks adopting optimal ram selection and use policies. The challenge is getting more flocks to adopt such policies.

INTRODUCTION

Prior to 1999, genetic evaluations for sheep flocks in New Zealand were largely done within-flock using sire model BLUP, through a number of service providers. In 1999 Sheep Improvement Ltd (SIL) was established as the single national performance recording and genetic evaluation service for the New Zealand sheep industry (Geenty 2000). SIL made individual animal model BLUP evaluations available on-demand leading to more accurate measures of genetic merit, in particular for reproduction measured as litter size (Newman *et al*. 2000).

Since 1999 SIL has continued to evolve, more flocks have joined, and more flocks participate in regular across-flock genetic evaluations. The Alliance (now Meat & Wool New Zealand) Central Progeny Test (CPT) began with matings in 2002 (Campbell *et al*. 2003). This provided genetic connectedness between breeds and breed groups that led to the large-scale, across-flock, across-breed genetic evaluation, SIL-ACE, starting in 2004 (Young and Newman 2009).

Previously no work has been carried out to estimate rates of genetic gain across the New Zealand sheep industry. SIL-ACE data provides the means to assess this. More than 300 flocks, which are estimated to breed greater than 50% of rams used by industry, participate in SIL-ACE.

MATERIALS AND METHODS

Data from the December 2007 SIL-ACE evaluation were made available by SIL. Average genetic merit was calculated for flocks over the period 1990 to 2006 based on data for animals born in these years. One hundred and forty flocks were designated as Dual Purpose (DP) and 62 as Terminal Sire (TS) based on breed. Approximately 1.6 million animals were in the evaluation of which 700 thousand were born in the years 2004-2006, inclusive.

Genetic merit was assessed as an overall index for DP sheep, combining information on Growth, Wool & Reproduction, and as an overall index for TS sheep, based on Growth and Meat (carcass merit based on body weight and ultrasound scanning). Consideration of changes in index components has been reported elsewhere (Amer 2009).

SIL indexes for overall merit combine information across relevant traits in the genetic evaluation (Amer 2000). SIL has reviewed index weightings and added more goal traits since

1999. The latest information on indexes and their component traits is available on the SIL website (www.sil.co.nz). Most SIL flocks use the standard SIL DP Overall or TS Overall indexes. Some flocks use variations on these but few use their own indexes.

Average genetic merit and average rate of genetic progress were calculated for each flock over four (unequal) time periods, where data were available. For some flocks, there were no data in the first period as they were not recording on SIL at that time. The periods were 1990-1994, 1995-1998, 1999-2003 and 2004-2006. The first two periods cover the time before SIL evaluations, while the period 2004-2006 covers the time when SIL-ACE results became available to help breeders compare genetic merit of animals from different flocks and breeds.

Using data from the fourth period (2004-2006) flocks were placed in one of the following four classes, separately for DP and TS flocks, based on average genetic merit and average genetic gain.

- **Leaders** – high merit, fast gain
- **Gaining** – below average merit, fast gain
- **Slipping** – high merit, below average gain
- **Trailing** – below average merit, below average gain

Aspects of flock structure and breeding strategy were studied to explain class differences: flock size, generation intervals (male and female), source of sires, average merit of sires (homebred *vs*. outside), merit of repeat use sires and proportion of lambs with unknown parents.

RESULTS AND DISCUSSION

Average rate of gains for DP and TS flocks are shown in Figure 1. Prior to SIL, gains were in the region of 30 cents/year. In the early years of SIL these increased to close to 60 cents/year and following the establishment of SIL-ACE, which utilized across breed connections from the CPT, this increased by about the same amount again, to close to 90 cents/year.

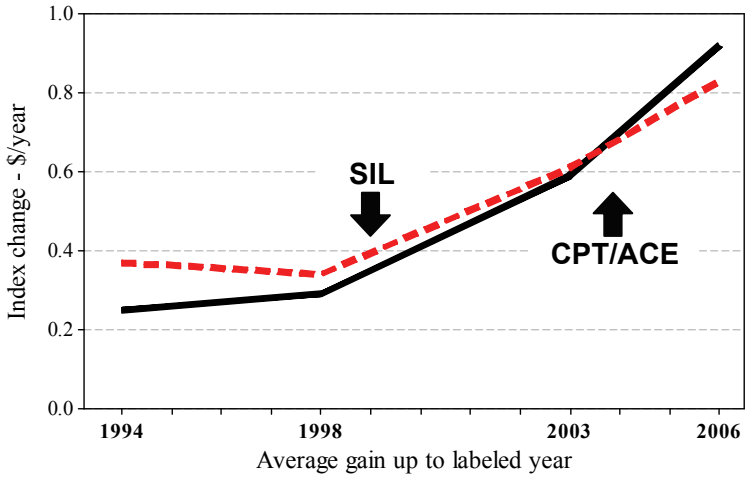


Figure 1. Average rate of gain in SIL-ACE overall indexes for Dual Purpose (solid line) & Terminal Sire (dashed line) sheep. NB: The Y-axis is rate of gain so anything greater than zero is genetic gain. Times that SIL & CPT/ACE were established are indicated.

Not all of these increases in rate of genetic gain can be attributed to SIL or to SIL-ACE. It is likely that breeders made more effective use of SIL figures as they gained confidence with experience. It might be argued that breeders using these systems are more innovative than those that do not. However, a number of flocks considered to be of high genetic merit or to have high

rates of genetic gain fall outside the sample of flocks used in this analysis due to poor genetic connectedness or through not participating in SIL-ACE. Without evidence to the contrary, we assume that the trends reported here are representative of industry trends.

Increased rates of genetic gain have also come from sources other than use of SIL, CPT and SIL-ACE information. Breed substitution, or "blending" has occurred in some flocks so "breed" composition of the industry has changed. Exotic breeds imported into NZ in the last few decades brought new genetics into the mix. No work has been reported on the relative performance of these exotics and local breeds from which we can assess the impact exotic genetics has had on these genetic trends. However, SIL-ACE results (available at www.sil.co.nz) clearly show that many sheep with outstanding genetic merit are purebreds from "traditional" breeds.

Average merit and rates of genetic gain varied over time for all flocks. Using average figures for the four time periods eliminates some of this noise. Table 1 presents average rates of gain for DP and for TS sheep, broken down into the four flock classes and the four time periods. These data clearly show the dramatic increases in genetic gain achieved by the Leader and Gaining flocks following the introduction of SIL and then of SIL-ACE. They also highlight that the Slipping and Trailing flocks made relatively modest gains over the same periods.

Table 1. Annual rates of genetic gain in overall economic index across 4 classes of Dual Purpose and across 4 classes of Terminal Sire flock participating in the SIL-ACE evaluation.

	Dual Purpose sheep index ($)			Overall	Terminal Sire sheep index ($)			Overall
Time period	Trailing	Slipping	Gaining	Leader	Trailing	Slipping	Gaining	Leader
2004-2006	0.57	0.60	1.34	1.41	0.38	0.44	0.86	1.05
1999-2003	0.47	0.75	0.51	0.71	0.27	0.55	0.66	0.71
1995-1998	0.14	0.49	0.08	0.45	0.19	0.32	0.53	0.33
1990-1994	0.17	0.30	0.11	0.39	0.19	0.46	- [1]	0.14

[1]With only one Terminal Sire flock classified as Gaining had animals present in the 1990-1994 time period, so a trend was not calculated.

Only two factors of flock structure and breeding strategy were associated with flock classification. Flocks with high rates of genetic gain were more likely to be using high merit sires from other flocks, while flocks with low rates of gain tended to repeatedly use low merit sires.

It is illuminating to consider the cumulative genetic changes over the years studied relative to industry production statistics for the DP index and two of its components (Amer 2009). These changes were greater than $8 for DP overall index, with more than $5 of this coming after 1999. Associated changes for two component BV traits were 1.4kg for carcass weight, with more than 900g coming after 1999, and more than 5% for number of lambs born (most coming after 1999). The relatively larger change for Reproduction under SIL is expected because genetic merit for litter size was based on very little information from relatives in earlier evaluation schemes.

Overall phenotypic performance of the New Zealand sheep industry improved over the period 1990-2006. Meat & Wool New Zealand Economic Service information (Davison, pers. comm.)

shows carcass weight increased by 2.5kg (14.4 to 16.9) and average lambing percentage by 16% (101.6 to 117.9). These provide the best estimate of actual industry performance against which we can assess the impact of the genetic changes we have described here. We conclude that genetics have contributed significantly to increases in industry performance.

Major differences were seen between individual flocks in rates of genetic gain (Figure 2). While the average flock gained close to $1/year on index from 2004 to 2006, some flocks gained more than $1.50/year and others gained less than 50c/year. This implies that breeders with a long-term commitment to using modern genetic improvement methods are making very good gains.

CONCLUSIONS

Better genetic evaluation methods and large-scale across-flock, across-breed genetic evaluations have each led to dramatic increases in rates of genetic gain in the New Zealand sheep industry. However, there still exists great variation in rate of genetic gain for flocks using these systems. Key features of the breeding strategies of the high merit, fast gaining, "Leader" flocks suggest that making use of high merit rams from other flocks and not using older sires repeatedly has led to higher rates of genetic gain.

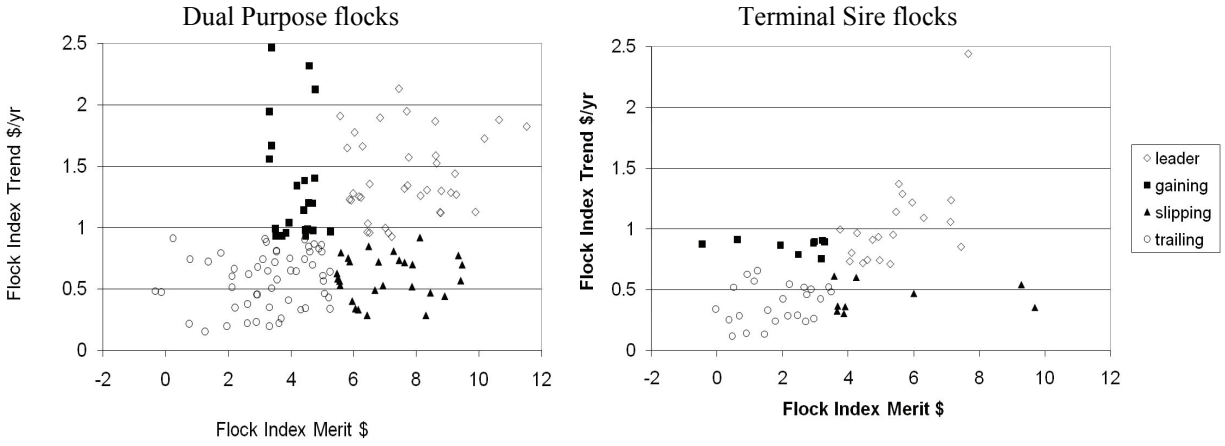


Figure 2. Dual Purpose and Terminal Sire flocks classified into 4 categories based on average index merit and average index trend for the 2004-2006 period.

ACKNOWLEDGEMENTS

Ovita Ltd. for providing the initial funding of this work. Sheep Improvement Limited for assisting with provision of data. John McEwan, Neville Jopson, and Sheryl-Ann Newman for input and comment on this research. Referees for valuable comments on drafts of this paper.

REFERENCES

Amer, P.R. (2000) *Proc. N.Z. Soc. Anim. Prod.*: **60**:189.

Amer, P.R. (2009) *Proc. N.Z. Soc. Anim. Prod.*: **69:**147

Campbell, A.W., Knowler, K., Behrent, M., Jopson, N.B., Cruickshank, G., McEwan, J.C., Wilson, T. (2003) *Proc. N.Z. Soc. Anim. Prod.*: **63**:197.

Geenty, K.G. (2000) *Proc. N.Z. Soc. Anim. Prod*. **60**:180

Newman, S-A.N., Dodds, K.G., Clarke, J.N., Garrick, D.J., McEwan, J.C. (2000) *Proc. N.Z. Soc. Anim. Prod*. **60:**195.

Young, M.J. and Newman, S-A.N. (2009) *Proc. N.Z. Soc. Anim. Prod.*: **69:**145

Sheep - Meat

THE INFORMATION NUCLEUS – GENETICALLY IMPROVING AUSTRALIAN LAMB PRODUCTION

S.I. Mortimer[1], K.L. Pearce[2], R.H. Jacobs[3], D.L. Hopkins[4], R.D. Warner[5], G.H. Geesink[6], J.E. Hocking Edwards[7], D.W. Pethick[2], J.H.J. van der Werf[6] and A.J. Ball[8]

CRC for Sheep Industry Innovation, [1]NSW Department of Primary Industries, Trangie, NSW 2823, [2]Murdoch University, Perth, WA 6150, [3]Department of Agriculture and Food WA, South Perth, WA 6151, [4]NSW Department of Primary Industries, Cowra, NSW 2794, [5]Department of Primary Industries, Werribee, VIC 3010, [6]University of New England, Armidale, NSW 2351, [7]South Australian Research and Development Institute, Naracoorte, SA 5271, [8] Meat & Livestock Australia, University of New England, Armidale, NSW 2351

SUMMARY

Progress with the implementation of the Information Nucleus and Next Generation Meat Quality programs of the CRC for Sheep Industry Innovation is described in relation to the estimation of quantitative genetic parameters for meat production and consumer-relevant traits. The traits have importance for the prediction of lean meat yield, understanding the biochemistry of meat quality traits, consumer acceptability, eating quality acceptability and nutritional value of lamb. Measurements of a comprehensive range of carcass, meat and growth traits are nearing completion on 2007 drop progeny and have commenced on 2008 drop progeny, with records available from up to 2000 animals within each drop. Initial analyses of carcass and meat traits (predicted lean meat yield and tenderness) indicate that sufficient genetic variation exists in novel meat traits which could be used in sheep genetic improvement programs.

INTRODUCTION

The Information Nucleus (IN) is a major initiative and program of the Co-operative Research Centre (CRC) for Sheep Industry Innovation established in 2007 (Fogarty *et al.* 2007). One of the CRC's core research programs that the IN supports is program 3, Next Generation Meat Quality, through the projects 3.1.1 - Range of Meat Phenotypes Measured and 3.2.1 - Biology and Production Pathways for Desired Phenotypes (Pethick *et al*. 2009). The integration of these 2 programs will provide the Australian sheep industry with new information on the quantitative and molecular genetics of new and novel meat traits, as well as growth traits and ultrasound scan measurements on live animals. Because of its linkages to industry flocks, the IN will also support genetic evaluation conducted by Sheep Genetics (Brown *et al.* 2007) as the data from IN progeny are included in the Sheep Genetics database and used in its routine genetic evaluations. This will enable enhanced genetic improvement across the whole sheep industry and in individual flocks as more accurate quantitative Australian Sheep Breeding Values (ASBVs) become available in the short term and later as molecular breeding values are incorporated into ASBVs.

This paper describes progress with the implementation of the IN and Next Generation Meat Quality programs, specifically with reference to the estimation of quantitative genetic parameters for meat production and consumer-relevant traits. As well as the enhancement of ASBVs, the genetic parameter estimates (heritability, phenotypic and genetic correlations) derived from the records of IN progeny will provide an understanding of the consequences of current industry breeding programs aimed largely at improving meat production traits on meat quality traits. The ability to include novel meat traits into sheep breeding programs will also be improved. This information will be critical in maintaining the market acceptability of lamb as a meat that is lean, nutritious, of high eating quality and visually appealing.

IMPLEMENTATION OF THE INFORMATION NUCLEUS

Details of the design, as well as the specific objectives, of the Information Nucleus have been described by Fogarty *et al.* (2007). Briefly, following the first mating in 2007, ewes and wethers have been produced at 8 sites across Australia. These were progeny of key industry sires representative of the major production types in the sheep industry: Merino, Border Leicester X Merino, Terminal X Merino and Terminal X Border Leicester-Merino. The progeny resulted from AI matings to Merino and crossbred ewes, with approximately 4200 ewes inseminated in 2007 at 7 sites and approximately 5000 ewes inseminated in 2008 and 2009 at 8 sites. Similar matings are planned to produce another 2 drops in 2010 and 2011. To date, the progeny born in 2007-08 have been measured and/or sampled for carcass, meat and growth traits, with the potential for records to be available from approximately 2000 animals within each drop. The data for these traits were recorded on the crossbred ewe and wether lambs (except for the Border Leicester X Merino ewe lambs) and half of the Merino wether lambs. Both the crossbred progeny and Merino wether lambs were slaughtered at a target average carcass weight of 21.5 kg, while the Merino wethers were slaughtered following their shearing at 10-11 months of age.

CARCASS AND MEAT MEASUREMENTS

The extensive list of carcass and meat traits being measured on IN progeny through project 3.1.1 of the Next Generation Meat Quality Program is presented in Table 1 (after Pethick *et al.* 2009), together with growth traits being recorded from birth through to slaughter. The traits have importance for the prediction of lean meat yield, understanding the biochemistry of meat quality traits, consumer acceptability, eating quality acceptability and nutritional value of lamb.

Table 1. Summary of growth, carcass and meat traits measured on IN slaughter progeny and their relevance to lamb production (after Pethick *et al.* 2009)

Relevance	Traits
Growth	Weight at birth, weaning, post-weaning, ultrasound scanning, slaughter
Lean meat yield prediction	Hot carcass weight
	GR, Fat C and Fat5th rib depths
	Eye muscle area
	Weight of short loin subcutaneous fat
	Weight of boneless short loin muscle
	Weight of topside muscle
	Weight of round muscle
	Weight of hind leg bone
Biochemistry	ICDH enzyme activity
	Myoglobin content of the loin muscle
Consumer acceptability	GR, Fat C and Fat5th rib depths
	Ultimate pH
	Fresh 24 hour meat colour
	Retail colour stability of the loin muscle
	Skin assessment
Eating quality acceptability	Rate of pH decline
	Ultimate pH
	Shear force of the loin muscle (1 and 5 day aged)
	Compression of the topside muscle
	Connective tissue content of the topside muscle
	Intramuscular fat of the loin muscle
Nutritional value	Iron and zinc content of the loin muscle
	Long chain fatty acids (Omega-3s) of the loin muscle

Pethick *et al.* (2009) have summarised the scope of measurement of traits by project 3.1.1 and highlighted the value of these traits for lamb products. For example, lean meat yield and tenderness were among the traits discussed. Improvements in lean meat yield were suggested to be of benefit to processors (through less carcass preparation time and facilitation of trade on a lean meat yield basis). Further understanding of the genetic relationship between tenderness and muscling was indicated as necessary to improve eating quality of lamb cuts, given earlier evidence that selection for muscling may result in greater toughness of certain lamb cuts (Hopkins *et al.* 2007). The traits listed in Table 1 have relevance to basic aspects of consumer demand for red meat that have been identified (Meat & Livestock Australia market research) as being required for the development of improved red meat products (Pethick *et al.* 2006). For lamb, these aspects of consumer demand included: lamb should be tender and juicy and of good flavour; lamb should provide a good source of lean, high quality protein and nutrients as part of a healthy diet; and that lamb should represent premium quality and value for money given that production systems throughout the supply chain are cost-efficient.

EARLY RESULTS

For these preliminary analyses, available records from 2007 drop animals were analysed for predicted lean meat yield (%) and shear force (Newtons, N) on the loin muscle after ageing for 5 days, as a measure of tenderness. Measurement protocols for the carcass components used to predict lean meat yield and shear force are described by Pearce (2008). Lean meat yield was predicted for each animal using an algorithm based on hot carcass weight, fat depth at the GR site on the hot carcass, eye muscle area between the 12^{th} and 13^{th} ribs and weights of loin fat, loin muscle, topside, round and femur bone (G.E. Gardner unpub. data). The model fitted to the data for each trait included the fixed effects of sex, age of dam, birth-rearing type, site, sire breed, dam breed and the interaction of site by sire breed and a random effect of sire. Hot carcass weight and ultimate pH were fitted as covariates in the model for shear force. The analyses were performed using ASReml (Gilmour *et al.* 2006). Average values for predicted lean meat yield and shear force were 50.3 % and 34.5 N respectively (Table 2). Both traits were at least moderately heritable, with values of estimated heritability of 0.31 for predicted lean meat yield and 0.26 for shear force. Although both estimates had large standard errors, the estimates indicate that genetic variation in these traits exists and may be sufficient to allow inclusion of such traits in sheep breeding programs. Smaller potential responses, however, will be expected from selection for lean meat yield than from selection for shear force given its much lower coefficient of variation (1.8% for lean meat yield versus 21.7% for shear force). Data from Australian flocks have not been previously available to estimate heritability for these traits, but other studies suggest that most carcass composition traits and shear force are of moderate to high heritability (Lambe and Simm 2004).

Table 2. Preliminary estimates of mean, phenotypic variance and heritability, with their standard errors, for predicted lean meat yield and shear force

Trait	Records	Range	Mean	Phenotypic variance	Heritability
Lean meat yield (%)	1781	44.1-52.2	50.3 (0.22)	0.82 (0.03)	0.31 (0.09)
Shear force (N)	773	10.8-75.6	34.5 (1.88)	55.9 (3.29)	0.26 (0.12)

CONCLUSION

The implementation of the IN and Next Generation Meat Quality programs is underway with measurements of a comprehensive range of carcass, meat and growth traits nearing completion on 2007 drop progeny and having commenced on 2008 drop progeny. Initial analyses of carcass traits

have indicated that genetic variation in novel meat traits may exist that could be used in sheep genetic improvement programs.

With completion of measurements on the animals slaughtered already and evaluation of progeny from the remaining 3 matings of the IN, a large body of data will be available for estimation of quantitative genetic parameters. These parameter estimates will enhance the accuracy of current ASBVs provided by Sheep Genetics and extend the range of ASBVs to include novel carcass and meat traits. Accuracy of ASBVs will be further improved by the availability of information on the influence of environmental effects on meat and carcass traits, such as birth-rearing type, dam age and management treatments. It is expected that more accurate ASBVs may be possible if molecular breeding values are used in their prediction for many traits. Research using the IN is now investigating the feasibility of predicting molecular breeding values based on the associations of single nucleotide polymorphisms (SNP) data with performance data from IN progeny (Fogarty *et al.* 2007).

The quantitative genetic parameter estimates will also be used to predict the consequences of selection on current breeding objectives, allowing responses to be monitored in unselected traits important to consumer acceptability, eating quality and nutritional value of lamb. Where responses in these traits are unfavourable, it may be necessary to develop breeding objectives and selection indexes that combine meat production traits with carcass and meat quality traits for use in individual flocks.

ACKNOWLEDGMENTS

The CRC for Sheep Industry Innovation is supported by the Australian Government's Cooperative Research Centres Program, Australian Wool Innovation Ltd. and Meat & Livestock Australia. At each site of the IN, considerable staff and resources are contributed by the participating organisations: NSW Department of Primary Industries, University of New England, Department of Primary Industries Victoria, South Australian Research and Development Institute and Department of Agriculture and Food WA. Staff from CSIRO and Murdoch University have also contributed to the Next Generation Meat Quality Program.

REFERENCES

Brown, D.J., Huisman, A.E., Swan, A.A., Graser, H-U, Woolaston, R.R., Ball, A.J., Atkins, K.D. and Banks, R.G. (2007) *Proc. Assoc. Anim. Breed. Genet.* **17**:187.

Fogarty, N.M., Banks, R.G., van der Werf, J.H.J, Ball, A.J. and Gibson, J.P. (2007) *Proc. Assoc. Anim. Breed. Genet.* **17**:29.

Gilmour, A.R., Gogel, B.J., Cullis, B.R., Welham, S.J. and Thompson, R. (2006) "ASReml User Guide Release 2.0" VSN International Ltd, Hemel Hempstead.

Hopkins, D.L., Stanley, D.F., Martin, L.C., Toohey, E.S., and Gilmour, A.R. (2007) *Aust. J. Exp. Agric.* **47:**1155.

Lambe, N.R. and Simm, G. (2004) In "Encyclopaedia of Meat Science", p.11, editors W.K. Jensen, C. Devine and M. Dikeman, Elsevier, Oxford.

Pearce, K. (2008) "Sheep CRC Program 3: Next Generation Meat Quality project 3.1.1 Phenotyping the Information Nucleus Flocks Operational Protocol Series"1st ed. Murdoch University, Murdoch.

Pethick, D.W., Banks, R.G., Hales, J. and Ross, I.R. (2006) In "Wool Meets Meat – Tools for a Modern Sheep Enterprise", p. 194, editors P.B. Cronjé and D.K. Maxwell, Australian Sheep Industry CRC, Armidale.

Pethick, D.W., Jacob, R.H., McDonagh, M.B., O'Halloran, W.J., Ball, A.J. and Hopkins, D.L. (2009) *Proc. N.Z. Soc. Anim. Prod.* **69**:96.

Sheep - Meat

ECONOMIC EVALUATION OF WHOLE GENOME SELECTION, USING MEAT SHEEP AS A CASE STUDY

R. G. Banks[1] and J.H. J. van der Werf[1]

[1]MLA, c/- Animal Science, University of New England, Armidale 2351
[2]Animal Science, University of New England, Armidale, NSW 2351

SUMMARY

A number of scenarios for meat sheep genetic improvement in Australia are evaluated, incorporating various combinations of additional performance recording via an Information Nucleus (IN), and implementation of Whole Genome Selection (WGS). The unsurprising results are obtained that the increment of economic return depends on the increments of accuracy, speed and cost. Perhaps more interestingly, it is clear that implementation of either IN and or WGS will almost certainly both require but also stimulate evolution of industry structure towards a more clearly defined nucleus:multiplier:commercial base model, and will require further evolution of co-investment, from the current mix of stud breeders, collective industry and taxpayer funds, to models likely to involve investment from processors, retailers and potential from genotyping companies.

INTRODUCTION

Whole genome selection (WGS, van der Werf 2009) is an anticipated outcome from current research into sheep (and other species) genomics, and has begun to be applied to dairy cattle breeding. Initial studies suggest that it offers potential improvement in accuracy of selection and in reducing generation intervals, at least for males (van der Werf 2009). In addition, the sheep industry has established an information nucleus (IN, (Banks *et al.* 2006), which provides capacity to record hard-to-measure traits and thus increase accuracy of selection for the breeding objective. The economic returns from application of WGS in conjunction with the IN will depend on the combination of the increase in rate of genetic progress achieved and the cost of the technology. This paper estimates those returns and the impact of varying the structure of the stud sector, and cost of WGS. The population modelled is the Australian slaughter lamb industry and the terminal sire stud sector which supplies it with rams.

MATERIALS AND METHODS

Model. A simple spreadsheet model was used, with the following parameters and assumptions:

- a commercial population of 10m ewes, mated to rams derived from the stud sector which are used for 4 years, with first progeny at 2 years
- a stud sector of 200,000 ewes with current recording costs in the studs of $10 per ewe, or $2.50 per ewe when WGS is applied (the latter reflecting many breeders choosing not to record if genotyping tools are available)
- current rate of progress in the stud sector of $2 per ewe joined per year (Swan *et al.* 2009)
- IN cost of $1.5m per year, and increase in rate of gain in the stud sector due to the IN of 7.5% (van der Werf 2009)
- increase in rate of progress due to WGS of 40% (van der Werf 2009)
- genotyping costs modelled at $20 per ewe or $50 per ewe
- where additional AI is used in conjunction with WGS, cost per ewe is $25
- costs and returns modelled over 25 years, with a discount rate of 7%

- only additional revenue due to genetic gain in carcase weight and carcase value was included in the returns.

Scenarios modelled:

A: this simply models the current situation, with no IN impact on either costs or returns. This represents what industry is currently achieving, and assumes no IN is established.

B: A plus an IN, and assumes that this is permanently in place.

C1 and C2: B with WGS added within the stud sector, with two genotyping costs modelled ($20 per ewe and $50 per ewe respectively)

C3: C2 but with recording costs per stud ewe reduced to $2.50, reflecting a partial replacement of current performance recording by WGS

D1: C2 with an elite nucleus, such that WGS is applied to only 10% of the stud sector, with the remaining stud ewes becoming a multiplier, lagging 2 years behind the elite 10%. This would be achieved by making all matings in the multiplier to young rams from the elite nucleus, used as ram lambs and used only once.

D2: this is the same as D1 but with all multiplier matings by AI at $25 per ewe

RESULTS AND DISCUSSION

Modelled parameters, costs and returns from each scenario are shown in Table 1.

Table 1. Summary of Economic Evaluation

Scenario	Parameters	Total Stud Recording Cost ($m pa)	Information Nucleus Cost ($m pa)	WGS Genotyping and AI Cost ($m pa)	Total Investment ($m pa) /Total cost per stud ewe ($)	Net Present Value over 25 years ($m)
A	Current	2	0	0	2/10	1,780 (100)*
B	+ IN	2	1.5	0	3.5/17.5	1,896 (107)
C1	+ IN + WGS at $20	2	1.5	4	7.5/37.5	2,432 (137)
C2	+ IN + WGS at $50	2	1.5	10	13.5/67.5	2,356 (132)
C3	+ IN + WGS at $50, reduced stud recording	0.5	1.5	10	12/60	2,375 (133)
D1	+ IN + WGS at $50, elite nucleus, no AI cost	0.5	1.5	1	3/15	1,959 (110)
D2	+ IN + WGS, elite nucleus, AI at $25	0.5	1.5	5.5	7.5/37.5	1,903 (107)

*: value in parentheses is the NPV relative to scenario A

There several aspects of these results, which can be taken in turn. Firstly, the current investment in and utilisation of BLUP technology by the terminal sire stud breeding sector (scenario A) is highly profitable in terms of industry NPV. To put this in context the Net Present Value (NPV) of the gross value of production (GVP) of lamb production over 25 years is approximately $32bn. Thus current genetic improvement in terminal sires adds approximately 5.5% of that GVP. There is very low risk around this NPV – the technology required is already in use, well adopted by both the stud breeder and via ram sales, by commercial producers.

Investment in the IN (scenario B) is predicted to generate an additional 7% NPV (compared to the current situation). This is dependent on the increase in accuracy of selection in the stud sector due to their relationship to animals tested in the IN. An important research question is how to maximise this relationship through optimisation of sire sampling, specifically how many sires should be evaluated, with what relationships amongst them, and how many progeny per sire.

There is low risk around the NPV estimated for scenario B. Essentially it is simply adding data on animals highly likely to make a significant genetic contribution, and via the relationships, on the whole population. The data that makes the difference is that on objective traits, rather than on the selection criteria currently used and which have a lower correlation with objective.

There is an additional benefit of the IN which is not modelled here. The IN includes recording for some traits which do not currently contribute to industry profitability and hence are not included in the objective, such as iron, zinc and omega-3 content. By recording these traits, industry immediately has the ability to at least monitor any correlated responses in them, and very quickly place selection pressure on them in the event that market signals for any become apparent. This provides both insurance and responsiveness benefits, both uncosted in this evaluation. The research questions around the IN which will assist industry to optimise the investment include the relationship with the rest of the stud sector, and how to optimise the traits actually recorded.

Scenarios C1, C2 and C3 add utilisation of WGS to the investment in the IN. Given the predicted increase in rate of progress, these scenarios all generate substantial increases in industry NPV. The differences reflect different assumptions about the price of WGS. The key uncertainty for these scenarios is whether the predicted increase in rate of progress can be achieved, and this in turn depends on whether the numbers of animals recorded for objective traits are sufficient to support the predictions of genetic merit.

A further researchable question regarding WGS is how many animals need to be phenotyped to calibrate a particular WGS prediction to the desired level of accuracy, and how quickly this calibration decays as genetic change proceeds. It will also be important for the sheep industry to build understanding of the linkage disequilibrium (LD) structure within and across the breeds used in slaughter lamb production: the terminal sires used are a mix of relatively recent breeds (eg. Poll Dorset, White Suffolk) and composites being formed from them, whilst the dams are a mix of Merino, Border Leicester-Merino and other breeds and crosses.

The most significant factor impacting industry implementation around the WGS scenarios is simply the increased level of investment required. This leads to discussion of the risk considerations of each scenario.

Risk Considerations. The current approach to genetic improvement in this sector (scenario A) is characterised by decision-making spread over some 400 stud breeders, all competing for ram sales, and with limited capacity to invest in technical or business development. Against this, the number, geographical spread and closeness of connection to ram buyers, provide a measure of robustness and adaptability to the system, coupled with a relatively low risk of total failure to achieve genetic progress. The other risk, or weakness, of this scenario, is that selection is for criteria only moderately correlated with a breeding objective defined from total value chain profitability.

Scenario B, which adds an IN, addresses this latter risk by incorporating measurement of objective traits, as well as of some traits anticipated to become important in the near future (such as the compositional traits). The technical risks here are that the wrong traits will be invested in, or traits will be invested in at the wrong level, and that it is not possible to increase the accuracy of genetic evaluation of stud candidates through their relationship with animals sampled for the IN. Two implementation risks are that animals sampled for the IN become less related to the animals that do end up being widely used, or conversely that too few are sampled and they have too much genetic impact on the total population, leading to inbreeding. These implementation risks are

manageable. The technical risk around trait investment is simply a matter for careful forecasting and risk assessment, while that around impact on accuracy is a matter of calculation and optimisation.

When we consider the WGS scenarios, there are two obvious challenges. Firstly, we don't yet know the accuracy of WGS, how much data is required for calibration, or how quickly predictions decay. All three factors impact on the advantage offered by WGS and how much it costs to achieve that advantage. The second challenge is an implementation challenge, focussed on the level of investment required and the industry structure. The current stud breeder's margin on investment in LAMBPLAN is currently around $25-50 per ewe mated. This margin would be nullified by either of the modelled genotyping costs ($20 or $50 per animal), and accordingly if these costs are realistic, WGS will simply not be taken up by breeders alone. This would likely remain the case even if substantial increases in profitability for processors and/or retailers were generated, since under the present industry structure, price signals from consumer to breeder are essentially averaged at each link in the value chain, and diminish as the number of links increases.

Recognising this potential limitation to adoption of WGS, scenarios D1 and D2 model the establishment of an elite (likely dispersed) nucleus, probably tightly coupled with, or involving, an IN. Given that the current IN is supported by industry and commonwealth funds, it is likely that at least some similar co-investment would be required to initiate scenarios resembling D1 or D2, and such co-investment, but including other partners such as processors, would be required for continuity.

An important point to stress regarding scenarios D1 and D2 is that they limit genotyping to perhaps 20,000 animals, or even less if possible. Indeed, D1 and D2 are quite robust to WGS not offering value at all, since essentially they are the logical next step in evolution of the industry improvement system once an IN is in place. These strategies thus reduce total investment significantly, but at the same time are predicted to deliver considerably lower additional NPV. These considerations around the D strategies highlight the challenge for industry relating to how to achieve the increased investments that seem likely to be needed, along with the uncertainties regarding the realised benefits of WGS.

These risk considerations suggest that the scenarios beyond B are much more uncertain, and to some extent less robust. This reflects the fact that whilst WGS offers scope for faster genetic progress, that scope is totally dependent on price and accuracy. In this sense WGS is simply another form of genetic evaluation, and its value is built on performance recording. Extracting maximum value from the minimum amount of recording is always the challenge.

REFERENCES

Banks R.G., van der Werf J.H.J., and Gibson J.P. (2006*) 8th World Cong. Genet. App. Livest. Prod.* **30:**12

Swan A.A., Brown, D.J., and Banks, R.G. (2009) *Proc. Assoc. Advmt. Anim. Breed. Genet.* **18**:326.

Van der Werf J.H. J. (2009) *Proc. Assoc. Advmt. Anim. Breed. Genet.* **18**:38

Sheep - Meat

PROGRESS IN THE DEVELOPMENT OF BREEDING SCHEMES FOR THE IRISH SHEEP INDUSTRY: THE MATERNAL LAMB PRODUCER GROUPS

T.J. Byrne[1], P. R. Amer[1], P. F. Fennessy[1], R. M. Rohloff[2], A. Cromie[3], P. Donnellan[3], G. Potterton[3], J. P. Hanrahan[4], and B. Wickham[3]

[1]AbacusBio Limited, P O Box 5585, Dunedin, New Zealand
[2]Awareka, 265 Conroys Road, R D 1, Alexandra, New Zealand
[3]Irish Cattle Breeding Federation Society Limited, Highfield House, Shinagh, Bandon, Co. Cork, Ireland
[4]Teagasc, Athenry Production Research Centre, Athenry, Co. Galway, Ireland

SUMMARY

The Irish sheep industry is undergoing major changes in the area of genetic improvement to ensure sheep farming remains competitive with alternative land uses. While the industry faces many challenges economically, environmentally, and socially (farm management), novel and innovative approaches have been developed in the area of sheep breeding to try and change the way ram breeding and purchasing occurs. The Maternal Lamb Producer (MALP) group scheme has been designed to assess different breeds of rams for performance in a range of traits, in a commercial farm situation. This will not only provide valuable data for wider ram genetic evaluations, through a Central Progeny Test (CPT), and demonstrate the use of electronic and DNA technology in sheep breeding, but also provide a relevant transparent demonstration to commercial farmers on the difference in the profitability arising from different rams. It is this demonstration, along with acceptance and uptake by commercial farmers that will drive ongoing investment in genetic improvement in the Irish sheep industry.

INTRODUCTION

The major challenges facing the Irish sheep industry include the part time nature of sheep farming, lack of scale, increasing compliance costs associated with environmental sustainability, high labour input costs, high capital costs associated with management systems, and increasing competition in markets. These factors have contributed to poor profitability in sheep farming operations. The poor profitability has resulted in a steady decline in the number of breeding ewes over the last 10 years, with 2.6 million ewes in 2008, down from 4.3 million in 1999 (Bord Bia 2008).

Sheep Ireland (www.sheep.ie) is the trading name of Sheep Database Limited, a company established by the Irish Sheep Industry. ICBF (the Irish Cattle Breeding Federations Society Ltd - www.icbf.com) provides management and technical expertise for Sheep Ireland and is in the process of implementing a genetic improvement program to help re-establish a highly profitable sheep industry as an important rural industry in the Irish economy. The development of a genetic improvement program provides a starting point for industry-wide efforts to rebuild numbers of breeding ewes and profitability throughout the industry.

Future developments are targeted towards a genetic improvement program that will facilitate sustainable profitability for all sectors of the Irish sheep industry. Ensuring that each component of the supply chain, from the producer to the processor to the market, is viable in the context of the overall sheep industry is critical. Sustainable profitability requires the sheep industry to be competitive with other enterprises. Genetic improvement coupled with advances in management

which increase returns to producers are very important pre-requisites to maintaining or growing the national flock (Conington *et. al.* 2001).

In genetic improvement programs, well-organised breeding schemes are important tools for selection of elite animals and for benchmarking breeding values between breeds and in different environments (McLean *et.al.* 2006). These require creation of good genetic linkages across groups that would not otherwise utilise the same rams.

Breeding schemes such as Central Progeny Tests (CPT), where a selection of rams that are genetically well-linked within the various breeding groups are used as sires, provide a basis for assessing a large number of rams across breeds and for many traits, and linking this information to an even wider number of animals throughout the linked pedigree system (McLean *et. al.* 2006).

This paper outlines the current industry structure, previous genetic evaluation services, previous and existing breeding schemes, and describes the structure and operation of a novel Maternal Lamb Producer (MALP) group scheme.

BREEDING INDUSTRY STRUCTURE

A notable and likely very influential aspect of the Irish sheep industry is the predominance of part-time sheep farmers. Over recent decades, part-time sheep farmers have been able to make significant income off-farm. It is likely that a portion of the recent reduction in sheep numbers can be attributed to a change in the business focus of sheep farmers.

The Irish ram breeding tier is made up of breeders that traditionally use visual appraisal to select for terminal traits within their breeds, with a small proportion of breeders using performance recording for a limited range of growth and carcase traits. Very little recording is carried out on maternal traits. These breeders sell rams to a commercial tier of approximately 35,000 farmers, with increasing competition in a shrinking market. This, combined with falling profitability in commercial sheep farms, has resulted in a difficult trading environment for breeders. This is exacerbated for breeders using performance recording, who have to bear additional costs associated with genetic evaluations, without recognition and financial reward in the market place for these efforts. With many commercial farmers purchasing lower priced rams, a large portion of the revenue received by the majority of pure-bred breeders comes from sales of rams to other breeders at shows and sales.

Natural service is the most common form of mating, and there is only limited use of artificial insemination.

Breeders tend to maintain their individual “pure” breeds with cross or composite breeding rare, and hence breed societies play an important role in ancestry recording, breed promotion and marketing, as well as trait selection.

GENETIC EVALUATIONS

In the past genetic evaluations for rams were based on breeding values of animals within breed, within flock, and within year. Genetic evaluation services via the Pedigree Sheep Breed Improvement Program (PSBIP) (Crosby *et. al.*1998) have been offered by the Irish Department of Agriculture for those breeders collecting performance data; incorporating growth rate, muscle depth and back fat depth to generate breeding values using multiple trait animal model Best Linear Unbiased Prediction (BLUP) methods. These breeding values were weighted and combined into a Lean Meat Index (LMI). Unfortunately participation in the scheme was limited to only a few terminal sire breeds and with the exception of the Texel breed, participation only reached around 10%-15% of the breed society members. A new across breed genetic evaluation system that integrates commercial and pedigree breeder data is under development by ICBF. Across breed genetic evaluations utilising commercial and pedigree breeder data have been very successful in the Irish beef sector (Amer *et. al.* 2001).

NEW APPROACHES AND A NOVEL BREEDING SCHEME

Ireland has lacked a breeding scheme that would provide an accurate assessment of rams across breeds and across multiple traits and that can be linked to the entire pedigree system. There has however been a number of small sire referencing schemes operating, including a sire referencing group (LMI-SR) comprising PSBIP Texel and Charollais breeders. This group was started in 1997, with the aim of improving the accuracy of the LMI and increasing genetic gain for participating breeders.

However these have been within breed rather than across breed evaluations and have been largely unsuccessful in delivering tangible benefits to commercial Irish sheep farmers. A lack of commercial producer uptake is recognised as a major contributing factor (William Hutchinson, pedigree breeders and commercial farmer – personal communication). A consequence of this is that breeders have considered that genetic improvement schemes represent a cost, rather than a benefit to their business.

The focus of the proposed breeding schemes in the Sheep Ireland genetic improvement program is to increase recording by breeders to include more economically relevant traits such as lambing difficulty, and combine this with a novel MALP group scheme beginning in 2008.

The rationale behind the MALP group scheme is to, not only provide valuable data which will be linked to a proposed CPT, but also provide a convincing demonstration of the range in genetic merit among a group of rams. The focus is on the financial and management impacts of using rams with the appropriate traits to improve income and decrease costs. To improve commercial ewe performance, the emphasis must be on maternal traits that will lead to lower cost and easy care sheep (Cottle 1999), and this is therefore the long-term focus of the MALP scheme. However in the short term, the immediate value will come from showing the variation in lambs that are slaughtered.

The objective is to help commercial farmers and processors to appreciate the value of improved genetics, and to help the processors understand the extent of variation among progeny of the same sire. This scheme also presents a great opportunity for breed societies to make significant contributions to the sheep industry through improvement and dissemination of elite breeding animals, and encouragement of performance recording by members.

The scheme involves some 23 commercial farms of sizes ranging from 80 to 1020 breeding ewes, assembled into five groups, involving a total of 224 rams and 8000 ewes. These commercial farms are spread over geographically diverse locations running a range of breeds, to provide regional interest and relevance to a variety of land types and production systems. The farmers involved are supported financially for the additional work undertaken.

Within each of the groups rams have been swapped during the mating season to provide genetic links between farms and therefore enable a comparison of rams across the group. A minimum of 2 rams have been used as links between participating farms to ensure adequate genetic linkage is obtained. Hence of the 224 rams, 70 will provide genetic linkage between flocks and between groups. Rams will also be used to link between years. It is envisaged that in the following years, with the development of a CPT, that data collected in the MALP group scheme will contribute, with the use of link rams, to the evaluation of rams across breed throughout the wider breeding industry.

A very important aspect of this system is that the MALP scheme aims to assess the commercial producers' own rams. This ensures transparency in terms of the results. All of these rams are from pedigree flocks (i.e. they have a pedigree identity). In addition, rams from each of the major recorded breeds (Suffolk, Texel, Charollais, Belclare and Vendeen) have been selected from existing performance-recorded flocks for inclusion in order to facilitate links between the MALP and CPT.

The use of adjunct technologies. The MALP scheme also provides an opportunity to demonstrate the use of electronic identification and DNA technology in sheep breeding (Lewis and Simm 2002).

In order to maintain individual animal identification in a commercial farm environment, all ewes and their progeny are to be electronically-tagged using low frequency electronic identification. This will simplify data collection for all traits and also simplify animal management (drafting, feeding management, record keeping). The real value proposition lies in the use of the data collected via electronic identification in the generation of information that can be applied on-farm to increase efficiency (e.g. culling of poor performers), and reduce costs on farm (cull for persistent lameness).

DNA parentage allows the producer to not only minimise the disturbance of lambing ewes and accurately identify each lamb to a dam, but also provides the opportunity to store blood samples that may be potentially valuable in the future. DNA parentage will be used to allow commercial multi-sire mating, ensure accuracy of parentage recording, and reduce the work load required at lambing. This will mean that all ewes, rams, and lambs will be DNA sampled.

CONCLUSION

A new genetic improvement strategy has been developed for Ireland encompassing a novel breeding scheme aimed at increasing uptake of performance recording by breeders and the use of improved rams by commercial farmers. Underpinning these new breeding initiatives are major developments in national breeding objectives for maternal and terminal sires, and an across breed genetic evaluation system integrating data from both pedigree breeders and commercial farmers. The scheme is built around commercial farms and is focused on practical and producer-driven recording. The MALP group scheme uses leading technology in identification and DNA to increase the accuracy and efficiency of data capture and recording. It is hoped that this scheme will accelerate adoption of genetic improvement technologies by breeders and drive commercial famers to provide accurate purchasing signals when purchasing rams.

REFERENCES

Amer, P. R., Simm, G., Keane, M. G., Diskin, M. G. and Wickham, B. W. (2001) *Livest. Prod. Sci.* **67:**223.

Bord Bia (2007) *Meat and Livestock Review*. Clanwilliam Court, Lower Mount Street, Dublin 2, Ireland.

Conington, J., Bishop, S. C., Grundy, B., Waterhouse, A. and Simm, G. (2001) *Anim. Sci.* **73:**413.

Cottle, D. J. (1999) *Proc. Assoc. Advmt. Anim. Breed. Genet.* **13:**436.

Crosby, E.J., Kelleher, D.L., Olori, V.E. and Reilly, M. (1998) In "Agricultural Research Forum" *Proc. 24th Ann. Meeting Irish Grassland and Anim. Prod. Assoc.* p 39.

Lewis, R. M. and Simm, G. (2002) *Proc. 7th Wld. Cong.Genet. Appl. Livest. Prod* Montpellier, France.

McLean, N. J., Jopson, N. B., Campbell, A. W., Knowler, K., Behrent, M., Cruikshank, G., Logan, C. M., Muir, P. D., Wilson, T. and McEwan, J. C. (2006) *Proc. N.Z. Soc. Anim. Prod* **66:**368.

CHEVERUD REVISITED: SCOPE FOR JOINT MODELLING OF GENETIC AND ENVIRONMENTAL COVARIANCE MATRICES

Karin Meyer[1] **and Mark Kirkpatrick**[2]

[1]Animal Genetics and Breeding Unit*, University of New England, Armidale, NSW 2351
[2]Section of Integrative Biology, University of Texas, Austin, Texas 78712

SUMMARY

Multivariate estimation fitting a common structure to estimates of genetic and environmental covariance matrices is examined in a simple simulation study. It is shown that such parsimonious estimation can considerably reduce sampling variation. However, if the assumption of similarity in structure does not hold at least approximately, bias in estimates of the genetic covariance matrix can be substantial. For small samples and more than a few traits, structured estimation is likely to reduce mean square error even if bias is quite large. Hence such models should be used cautiously.

INTRODUCTION

Accurate estimation of genetic covariances and correlations is inherently problematic as it requires substantial numbers of records on pairs of close relatives for all traits of interest, and as it may impose a considerable computational burden. Examining literature results, Cheverud (1988) found that estimates of genetic correlations for sets of traits such as body measurements are often very similar to their phenotypic counterparts. Others reported corresponding patterns for different natural populations (Roff 1995, 1996), in plants (Waitt and Levin 1998) and livestock (Koots and Gibson 1996; Kominakis 2003). Cheverud's suggestion to substitute estimates of phenotypic for genetic correlations, in particular when sample sizes are small or pedigree information is limited, has met with justifiable criticism (Willis et al. 1991; Kruuk et al. 2008). However, the idea of 'borrowing strength' from the phenotypic covariance matrix in estimating genetic covariances is appealing.

As multivariate analyses involving more than a few traits have become computationally feasible, there has been increasing interest in 'structured' estimation. A modern, mixed model based analogue to Cheverud's proposal might be to estimate genetic and phenotypic or environmental covariance matrices, imposing a common structure on the two matrices. This paper examines three alternatives to do so and their impact on estimates and their sampling properties.

MATERIAL AND METHODS

Structured estimation. Consider a multivariate analysis of q traits, fitting a simple animal model. Let $\mathbf{\Sigma}_G$ and $\mathbf{\Sigma}_E$ denote the covariance matrices for additive genetic and residual effects, respectively.
Unstructured. In most multivariate analyses, we assume covariance matrices are 'unstructured' (US), i.e. we describe the $q(q+1)/2$ distinct elements of each matrix by the corresponding number of parameters. A common parameterisation is to the elements of the Cholesky factor of a matrix.
Common correlation. To fit a common correlation (CORR) matrix, $\mathbf{R}$, we model $\mathbf{\Sigma}_G = \mathbf{S}_G\mathbf{R}\mathbf{S}_G$ and $\mathbf{\Sigma}_E = \mathbf{S}_E\mathbf{R}\mathbf{S}_E$, with $\mathbf{S}_G$ and $\mathbf{S}_E$ the diagonal matrices of genetic and residual standard deviations.
Common principal components. Fitting common principal components (CPC), we assume that both matrices have the same eigenvectors, $\mathbf{V}$, but different eigenvalues, i.e. $\mathbf{\Sigma}_G = \mathbf{V}\mathbf{\Lambda}_G\mathbf{V}'$ and $\mathbf{\Sigma}_E = \mathbf{V}\mathbf{\Lambda}_E\mathbf{V}'$ with $\mathbf{\Lambda}_G$ and $\mathbf{\Lambda}_E$ the diagonal matrices of genetic and residual eigenvalues.
Common GARP model. A related, common structure is obtained by modelling $\mathbf{\Sigma}_G = \mathbf{U}\mathbf{D}_G\mathbf{U}'$ and $\mathbf{\Sigma}_E = \mathbf{U}\mathbf{D}_E\mathbf{U}'$, with $\mathbf{U}$ a unitary, lower triangular matrix. The non-zero off-diagonal elements of $\mathbf{U}$ have an interpretation as regression coefficients in an auto-regressive model, hence the acronym

[†]AGBU is a joint venture of NSW Department of Primary Industries and University of New England

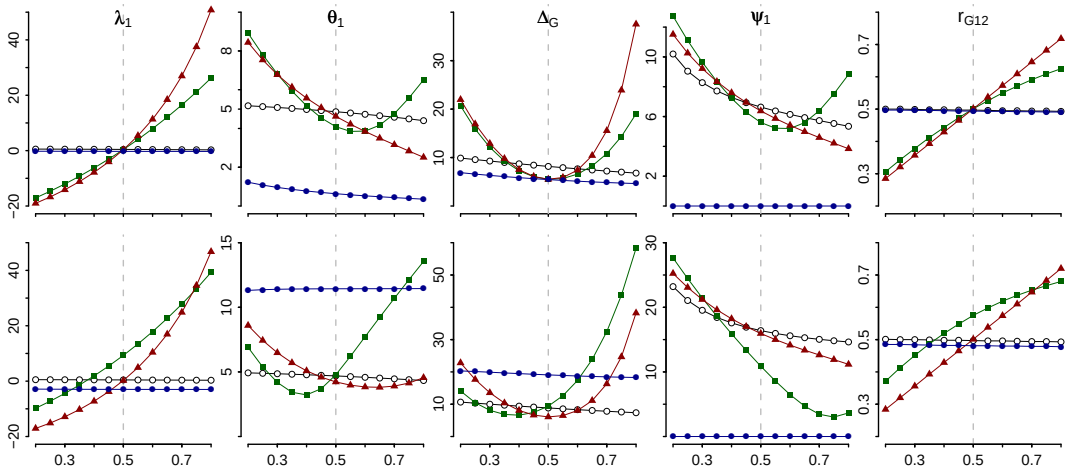


Figure 1. Mean estimates (see text for definitions) for example I, for equal (top row) or unequal (bottom row) heritabilities and r_E = 0.2 to 0.8 (○ US, ▲ CORR, ■ GARP, • CPC).

GARP, standing for generalised auto-regressive parameters (Pourahmadi et al. 2007). Correspondingly, the elements of the diagonal matrices $\mathbf{D}_G$ and $\mathbf{D}_E$ represent the 'innovation' variances, i.e. for variable i the conditional variance given variables 1 to $i-1$.

Parsimony. Each of the 3 structured models reduces the number of parameters to be estimated by $q(q-1)/2$, i.e. from $p = q(q+1)$ in the US case (for $\mathbf{\Sigma}_G$ and $\mathbf{\Sigma}_E$) to $p = q(q+3)/2$.

Simulation. Behaviour of restricted maximum likelihood (REML) estimates for the 4 different parameterisations was examined considering a simple, balanced paternal half-sib design (s sires with n progeny each). This involved sampling of the matrices of mean squares between and within families from appropriate central Wishart distributions, performing 10 000 replicates for each scenario considered. Maximisation of the likelihood, constraining both $\hat{\mathbf{\Sigma}}_G$ and $\hat{\mathbf{\Sigma}}_E$ to be positive definite, was carried out using a Method of Scoring algorithm combined with a derivative-free search.

Summary statistics. Means over replicates were calculated for estimates of genetic correlations ($r_{G\,ij}$), the eigenvalues of $\hat{\mathbf{\Sigma}}_G$ (λ_i), the log likelihood ($\log \mathcal{L}$), and (for $\mathbf{V}_X = \{\mathbf{v}_{X\,i}\}$ in $\mathbf{\Sigma}_X = \mathbf{V}_X\mathbf{\Lambda}_X\mathbf{V}'_X$)

- the angle between i–th eigenvectors of $\mathbf{\Sigma}_G$ and $\hat{\mathbf{\Sigma}}_G$: $\theta_i = (180/\pi)\arccos|\hat{\mathbf{v}}'_{G\,i}\mathbf{v}_{G\,i}|$
- the angle between i–th eigenvectors of $\hat{\mathbf{\Sigma}}_G$ and $\hat{\mathbf{\Sigma}}_E$: $\psi_i = (180/\pi)\arccos|\hat{\mathbf{v}}'_{G\,i}\hat{\mathbf{v}}_{E\,i}|$
- the 'quadratic loss' in $\hat{\mathbf{\Sigma}}_G$: $\Delta_G = \mathrm{tr}(\hat{\mathbf{\Sigma}}_G\mathbf{\Sigma}_G^{-1} - \mathbf{I})^2$
- the mean squared difference in $\hat{r}_{G\,ij}$ and $\hat{r}_{E\,ij}$ (in %): $\Delta_R = \sum_{i=1}^{q}\sum_{j=i+1}^{q}(\hat{r}_{G\,ij} - \hat{r}_{E\,ij})^2/(q(q-1)/2)$
- the 'adjusted' Akaike information criterion: $\mathrm{AIC} = -2\log\mathcal{L} + 2p(1 + \frac{p+1}{qsn-p-1})$

Example I. Example I comprised $q = 2$ traits with a genetic correlation of $r_{G\,12} = 0.5$ and equal phenotypic variances ($\sigma^2_P = 100$), for a moderate sample size ($s = 500$ with $n = 8$). Environmental correlations considered were $r_{E\,12} = 0.2$ to 0.8. Scenario A assumed heritabilities for both traits were equal ($h^2_1 = h^2_2 = 0.3$), while scenario B involved different values ($h^2_1 = 0.36, h^2_2 = 0.24$).

Example II. The second example involved $q = 6$, again using equal parameters of $h^2_i = 0.33$, $r_{G\,ij} = r_{E\,ij} = 0.5$ and $\sigma^2_P = 100$ for all traits to construct population values for $\mathbf{\Sigma}_G$ and $\mathbf{\Sigma}_E$. $\mathbf{\Sigma}_E$ was then replaced by $\mathbf{T}\mathbf{\Sigma}_E\mathbf{T}'$, with $\mathbf{T} = \prod_{i<j}^{q}\mathbf{C}(\alpha)^{ij}$ and $\mathbf{C}(\alpha)^{ij}$ a rotation matrix with elements $c_{ii} = c_{jj} = \cos(\alpha)$, $c_{kk} = 1$ for $k \neq i, j$, $c_{ij} = \sin(\alpha)$, $c_{ji} = -c_{ij}$ and zero otherwise. Rotation angles from $\alpha = 0°$ to 6° (equal for all i, j) were used to generate $\mathbf{\Sigma}_G$ and $\mathbf{\Sigma}_E$ with increasingly different eigenvectors. Three sample sizes, $s = 1000, n = 20$, $s = 500, n = 10$ and $s = 250, n = 8$, were examined.

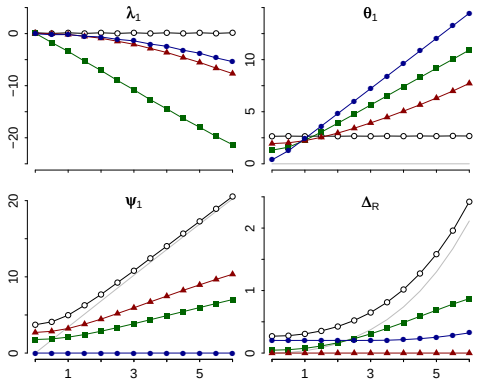


Figure 2. Means statistics for example II (s=1000, n=20; see Figure 1 for legend).

RESULTS

Means of summary statistics for example I are summarised in Figure 1. For equal heritabilities, eigenvectors of $\mathbf{\Sigma}_G$ and $\mathbf{\Sigma}_E$ are collinear regardless of the value of r_{E12} (shown along the horizontal axes). Hence CPC is the correct model throughout, and estimates of λ_1 (expressed here as % deviation from population value) and $r_{G\,12}$ for CPC and US are virtually the same. Fitting CPCs thus reduces sampling variation in the direction of the first genetic eigenvector (θ_1) substantially, and yields a consistently lower loss in $\hat{\mathbf{\Sigma}}_G$ (Δ_G) than the US model. For CORR and GARP, estimates of $r_{G\,12}$ are dominated by the population value for $r_{E\,12}$ (i.e. $\hat{r}_{G\,12}$ closely follows $r_{E\,12}$), with a corresponding bias in estimates of λ_1. The angle between the first eigenvectors of $\hat{\mathbf{\Sigma}}_G$ and $\hat{\mathbf{\Sigma}}_E$ (ψ_1) is similar to that for US, except for high values of $r_{E\,12}$. With a moderate sample size, there is a narrow range of $r_{E\,12} \neq r_{G\,12}$ ($\approx 0.35 - 0.60$) for which these structured models reduce Δ_G compared to US.

For $h_1^2 \neq h_2^2$, however, CPC is no longer the correct model, with the angle between the first eigenvectors of $\mathbf{\Sigma}_G$ and $\mathbf{\Sigma}_E$ ranging from 22.7° ($r_{E\,12} = 0.2$) to 14.6° ($r_{E\,12} = 0.8$). Fitting CPC for this case, estimates of λ_1 and $r_{G\,12}$ are little affected, but estimates of the direction of eigenvectors are heavily biased, with a correspondingly large loss Δ_G. Nevertheless, CPC appears advantageous over both CORR and GARP for larger differences between $r_{G\,12}$ and $r_{E\,12}$.

Results for example II are given in Figure 2, with different values of the rotation angle α along the horizontal axes. The population value for ψ_1 (shown as smooth gray line) increases linearly with the α used, causing estimates of θ_1 to increase similarly when fitting CPC. Again, estimates of λ_1 are relatively little biased, even if the CPC model is grossly wrong. True differences in Δ_R (gray line) increase quadratically with α. For US analyses, estimates of Δ_R are consistently larger, reflecting marked sampling variation. All three structured models underestimate differences in genetic and environmental correlations for values of α larger than $\approx 3°$.

Corresponding values for Δ_G together with the proportion of replicates for which each model fitted 'best', based on the value of AIC, are shown in Figure 3. With a difference of 15 parameters between US and structured models, the latter can provide estimates of $\mathbf{\Sigma}_G$ with substantially lower quadratic loss than US, especially for small samples. While CORR and GARP appeared advantageous over CPC in terms of Δ_G, model selection on the basis of AIC generally favoured CPC over the other structured models, decreasingly so as α increased. Bias in both the individual parameters and Δ_G increased faster with α for CPC than for CORR or GARP. AIC is derived from $\log \mathcal{L}$ and thus

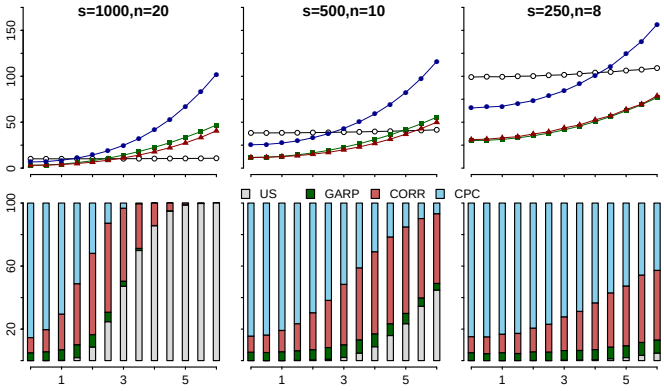


Figure 3. Estimates of Δ_G and proportion of samples (in %) for which each model fitted 'best' for example II (○ US, ▲ CORR, ■ GARP, • CPC).

dominated by $\hat{\mathbf{\Sigma}}_E$, i.e. such model selection by and large does not aim at minimising loss in $\hat{\mathbf{\Sigma}}_G$.

DISCUSSION

It has been demonstrated that joint modelling of the genetic and environmental covariance matrix in mixed model analyses is readily feasible, and can result in reduced sampling variation. This resulted in 'improved' estimates of $\mathbf{\Sigma}_G$, i.e estimates with a smaller quadratic loss than unstructured estimates in a range of scenarios. Of the three alternative parameterizations considered, none proved best overall. The common principal components model tended to yield least biased estimates of the first genetic eigenvalue. Reduction in mean square errors and loss generally comes at the price of bias in estimates. Disconcertingly, standard likelihood based model selection procedures (AIC) appeared to favour parsimonious models imposing a common structure for a range of cases where this was not the appropriate model, in spite of accounting for the number of parameters estimated. For small samples in particular, Δ_G somewhat higher than in the US case and thus potentially non-negligible bias seemed to be tolerated. Further work is necessary to determine the best strategy for model selection in practical applications.

A less rigid alternative to the assumption of a common structure may be a 'shrinkage' of the estimated genetic towards the phenotypic covariance matrix. While this does not reduce the number of parameters to be estimated, it can reduce sampling variation in $\hat{\mathbf{\Sigma}}_G$ and thus Δ_G. For instance, we could maximise $\log \mathcal{L}$ subject to a penalty which measures the divergence between $\hat{\mathbf{\Sigma}}_G$ and $\hat{\mathbf{\Sigma}}_G + \hat{\mathbf{\Sigma}}_E$. This is similar in spirit to the 'bending' procedure proposed by Hayes and Hill (1981). Preliminary analyses have been promising, showing a marked reduction of loss in $\hat{\mathbf{\Sigma}}_G$ even for mild penalties accompanied by relatively small bias.

CONCLUSION

Structured estimation provides a powerful tool to increase the accuracy of genetic parameter estimation, especially for multivariate analyses comprising more than a few traits and smaller sample sizes, and is readily implemented in a mixed model framework. However, as always, there is the trade-off between a reduction in sampling variation and bias. The utility of such analyses depends very much on the underlying assumption of a common structure to hold at least approximately – while parsimonious estimation may yield estimates with reduced loss or mean square error, this may be at the expense of substantial bias. Structured estimation appears most promising when sample sizes are small, but such models are not a substitute for using data sets of sufficient size.

ACKNOWLEDGEMENTS

This work was supported by Meat and Livestock Australia under grant BFGEN.100B (KM) and National Science Foundation grants EF-0328594 and DEB-0819901 (MK).

REFERENCES

Cheverud, J. M. (1988) *Evolution* **42**:958.
Hayes, J. F. and Hill, W. G. (1981) *Biometrics* **37**:483.
Kominakis, A. P. (2003) *J. Anim. Breed. Genet.* **120**:269.
Koots, R. and Gibson, J. P. (1996) *Genetics* **143**:1409.
Kruuk, L. E. B., Slate, J. and Wilson, A. J. (2008) *Ann. Rev. Ecol. Evol. System.* **39**:525.
Pourahmadi, M., Daniels, M. J. and Park, T. (2007) *J. Multiv. Anal.* **98**:569.
Roff, D. A. (1995) *Heredity* **74**:481.
Roff, D. A. (1996) *Evolution* **50**:1392.
Waitt, D. E. and Levin, D. A. (1998) *Heredity* **80**:310.
Willis, J. H., Coyne, J. A. and Kirkpatrick, M. (1991) *Evolution* **45**:441.

FACTOR-ANALYTIC MODELS TO REDUCE COMPUTATIONAL REQUIREMENTS IN INTERNATIONAL GENETIC EVALUATION OF BEEF CATTLE

Karin Meyer

Animal Genetics and Breeding Unit*, University of New England, Armidale, NSW 2351

SUMMARY

There has been considerable recent interest in factor-analytic models for the analysis of genotype by environment type problems. It is shown that, compared to 'standard' multivariate analyses, such models can substantially reduce computational requirements for international genetic evaluation of beef cattle, fitting an animal model and treating performance in different countries as separate traits.

INTRODUCTION

Factor-analytic (FA) structures provide a powerful means to model multivariate covariance matrices for 'similar' traits parsimoniously, and are readily implemented in our standard mixed model framework for estimation and prediction. FA models are used increasingly, in particular in the analysis of data from plant breeding trials subject to genotype by environment (G×E) interactions; see Meyer (2009) for a recent review. This paper demonstrates that FA models can provide an attractive alternative to standard, multi-trait models for international genetic evaluation of beef cattle.

MATERIAL AND METHODS

Alternative models. Consider an animal model analysis of q traits where each individual has a record for a single trait ('country' or 'location') only. Assume covariance matrices for genetic effects have a factor-analytic structure, i.e. can be written as $\mathbf{\Sigma} = \mathbf{\Gamma\Gamma}' + \mathbf{\Psi}$, with $\mathbf{\Psi}$ diagonal. We then have 3 equivalent models.

i) Multivariate (MV): In a multi-trait analysis we estimate genetic effects ($\mathbf{u}$) in each country directly invoking the standard, linear model $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Zu} + \mathbf{e}$, with $\mathrm{Var}(\mathbf{u}) = \mathbf{\Sigma} \otimes \mathbf{A}$ and $\mathbf{A}$ denoting the numerator relationship matrix (NRM). This requires $\mathbf{\Sigma}$ to have full rank.

ii) 'Extended' factor analytic (XFA): If $\mathbf{\Sigma} = \mathbf{\Gamma\Gamma}' + \mathbf{\Psi}$, we can partition genetic effects into m common ($\mathbf{c}$) and q specific ($\mathbf{s}$) factors, $\mathbf{u} = (\mathbf{\Gamma} \otimes \mathbf{I})\mathbf{c} + \mathbf{s}$, with $\mathrm{Var}(\mathbf{c}) = \mathbf{I} \otimes \mathbf{A}$ and $\mathrm{Var}(\mathbf{s}) = \mathbf{\Psi} \otimes \mathbf{A}$. Fitting these separately gives XFA model $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}^\star\mathbf{c} + \mathbf{Zs} + \mathbf{e}$, with $\mathbf{Z}^\star = \mathbf{Z}(\mathbf{\Gamma} \otimes \mathbf{I})$ (Thompson et al. 2003).

iii) Principal components (PC): If specific effects are assumed absent, i.e. $\mathbf{\Psi} = \mathbf{0}$, the XFA model reduces to a PC model. An equivalent model to MV and XFA is obtained if we decompose $\mathbf{\Gamma\Gamma}' + \mathbf{\Psi} = \mathbf{\Gamma}^\star\mathbf{\Gamma}^{\star\prime}$ (to form $\mathbf{Z}^\star = \mathbf{Z}(\mathbf{\Gamma}^\star \otimes \mathbf{I})$) and consider all factors, i.e $m = q$. For $m < q$, we obtain a reduced rank model, i.e., the PC model can accommodate $\mathbf{\Sigma}$ which does not have full rank.

Transformations. Solutions for genetic effects from one model are readily transformed to those from another. Fitting an XFA or full rank PC model, $\hat{\mathbf{u}} = (\mathbf{\Gamma} \otimes \mathbf{I})\hat{\mathbf{c}} + \hat{\mathbf{s}}$ or $\hat{\mathbf{u}} = (\mathbf{\Gamma}^\star \otimes \mathbf{I})\hat{\mathbf{c}}$. Conversely, fitting a MV model, $\hat{\mathbf{c}} = (\mathbf{\Gamma}'\mathbf{\Sigma}^{-1} \otimes \mathbf{I})\hat{\mathbf{u}}$ and $\hat{\mathbf{s}} = (\mathbf{\Psi\Sigma}^{-1} \otimes \mathbf{I})\hat{\mathbf{u}}$ (Smith et al. 2001).

Factor rotation. Fitting m common factors, $\mathbf{\Gamma}$ of size $q \times m$ has $m(m-1)/2$ elements given by orthogonality constraints. $\mathbf{\Gamma}$ is not unique and can be rotated, i.e. replaced by $\mathbf{\Gamma T}$ with $\mathbf{T}$ an orthogonal matrix. A particular rotation yields $\mathbf{\Gamma}^+$ with all above diagonal elements equal to zero, which can be interpreted as the matrix derived from the first m columns of the Cholesky factor of $\mathbf{\Sigma}$. Such 'trian-

*AGBU is a joint venture of NSW Department of Primary Industries and University of New England

gular' $\mathbf{\Gamma}^+$ yields a less dense $\mathbf{Z}^\star$. Let $\boldsymbol{\gamma}'_j$ denote the $j-$th row of $\mathbf{\Gamma}^+$. For an individual with a record in location j, $\boldsymbol{\gamma}'_j$ represents the respective part of the design matrix, contribution to the coefficient matrix ($\mathbf{C}$) in the mixed model equations (MME) is proportional to $\boldsymbol{\gamma}_j\boldsymbol{\gamma}'_j$. As elements $j+1$ to m of $\boldsymbol{\gamma}_j$ are zero, for a single record per individual, the corresponding $m\times m$ diagonal block of $\mathbf{C}$ then consists of a dense sub-block for rows and columns 1 to j, while the remaining $m-j$ rows and columns are diagonal (with non-zero diagonal elements due to the NRM). For instance, for $j=2$ only row and column 1 and 2 are linked by a non-zero element, and only for $j=q$ are all m^2 elements in the diagonal block non-zero.

Data. Case I comprised simulated records for a half-sib structure, considering $q = 8, 12$ or 16 countries, with 100 'global' sires used in all countries, 900 'local' sires used in a particular country only, and 50 progeny per sire and country. Both sires and progeny were assumed to have a single record, with Global sires belonging to country 1. The model of analysis was a simple animal model, fitting country means as the only fixed effects. Data were simulated assuming heterogeneous variances, moderate heritabilities and high genetic correlations between countries, using population values of covariance matrices in solving the MME.

Case II considered 865 129 weaning weight records for Australian and New Zealand Angus calves in 1 336 herds, pre-corrected for the effects of age at weaning, birth type and dam age, and $N = 998\,479$ animals in the pedigree. Records in individual herds were assigned to 10 different 'countries', considering herds as they occurred in the data, resulting in 69 875 (country 7) to 104 903 (country 2) records per trait. The model of analysis fitted direct and maternal additive genetic effects and maternal permanent environmental effects (324 613 dams) as random effects, and contemporary groups (between 9 243 and 12 415 per 'country') as fixed effects. Covariance matrices used for the mixed model analyses were constructed from standard variance components for this trait and breed, with genetic correlations assumed to be about 0.8 (0.78 to 0.82).

Case III involved weaning weight records for Hereford calves in 9 countries, as collated for a global international evaluation feasibility study (Graser 2008). There were 4 281 659 records, 2 678 762 dams with progeny in the data, and 6 648 388 animals in the pedigree. For efficiency, countries were renumbered in descending order of the number of records (c.f. Meyer 2009), with 2.67, 0.55, 0.52, 0.22 and 0.16 million for countries 1 to 5, and less than 50 000 for the remainder. The model of analysis was as for case II. Covariance components utilised were similar to those in the Hereford pilot study, but ignored three small within country direct-maternal genetic covariances.

Analyses. Estimates of fixed effects and predictions for random effects were obtained by solving the corresponding MME iteratively, using a pre-conditioned conjugate gradient (PCG) algorithm, as implemented in the iterative solutions module of WOMBAT (Meyer 2007) which holds the complete MME in core, using sparse matrix storage techniques. All calculations were carried out using double precision (8 Byte) floating point variables. XFA and PC models were parameterised using the rotation of factors to 'triangular' $\mathbf{\Gamma}^+$. Strategy A used a simple, diagonal pre-conditioning matrix, obtained as the reciprocals of the diagonal elements of $\mathbf{C}$, and stored in core. Strategy B employed a block pre-conditioner for random effects. For MV models this involved the inverse of the $q\times q$ diagonal block for each level, treated as dense. For XFA and PC models blocks were of size $m+q$ and m, respectively, i.e. jointly considered common and specific factors (if fitted) for an individual. Inversion of these blocks exploited their known sparsity structure. Dense blocks of all inverses were written to disk and re-read for each PCG iterate. The algorithm was assumed to have converged when the ratio of the sum of squared deviations in solutions between iterates dived by the sum solutions squared was less than 10^{-14}. Computations for analyses requiring less than 4 Gb of memory (RAM) were carried out on a single user, 64-bit machine with a dual-core processor rated at a speed

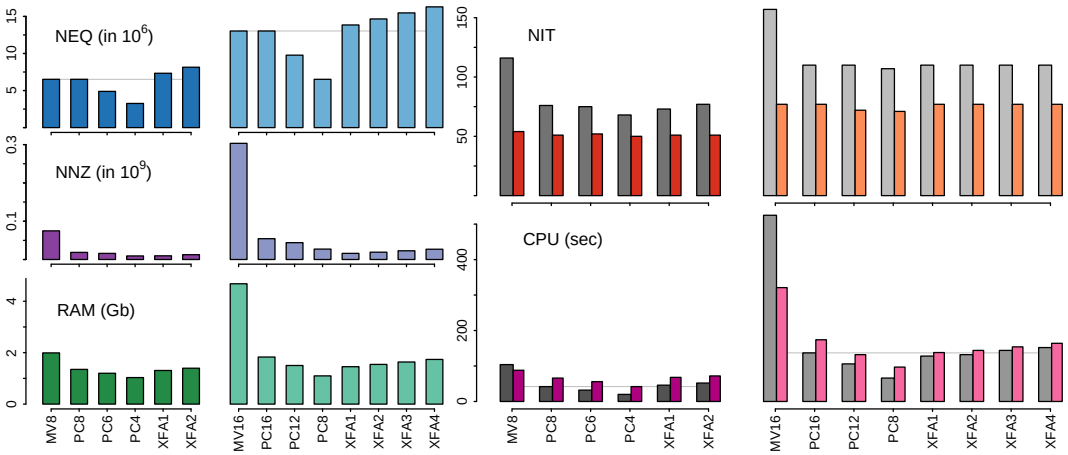


Figure 2. Results for simulated data for 8 (dark) and 16 (light) traits (see text for definitions)

of 2.2 GHz. The remaining analyses were performed on a multi-user, 64-bit machine with dual-core processors rated at 2.6 Ghz.

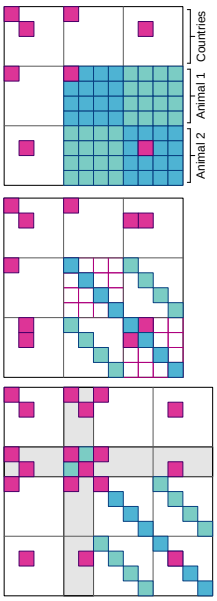


Figure 1. Toy case

RESULTS

Figure 1 shows the sparsity pattern of $\mathbf{C}$ for the 3 alternative models for a toy example comprising $q = 4$, means for each trait (rows/columns 1 to 4) and effects for 2 animals, with records in country 1 and 2, respectively, and animal 1 the sire of animal 2. For MV (top), each element of the NRM inverse contributes q^2 coefficients, resulting in dense animal $\times$ animal blocks. In contrast, for PC (middle, $m = q$) or XFA (bottom, $m = 1$), each such element contributes only m or $m + q$ coefficients. This is off-set, in part at least, by a denser design matrix $\mathbf{Z}^\star$, i.e. more contributions from the 'data part'. However, for 'triangular' $\mathbf{\Gamma}^+$, contributions are limited. For XFA, there are extra equations (gray background) and additional 'data' terms linking equations for $\mathbf{c}$ and $\mathbf{s}$, but diagonal blocks of $\mathbf{C}$ are sparser still.

Case I. Characteristics of the MME and computational requirements to solve them for the simulated data are summarised in Figure 2. To alleviate scale problems, values for $q = 8$ (except NIT) are plotted at twice their actual value. While MVq and PCq involved the same number of equations (NEQ), the number of non-zero off-diagonal elements in one triangle of $\mathbf{C}$ (NNZ) and, consequently, the memory required (RAM) differed substantially. Using a simple, diagonal pre-conditioning matrix (left bars) required marked more PCG iterates (NIT) than a block-diagonal matrix (right bars), but, except for MV, required less time in total (CPU). Increasing the number of common factors m in XFAm analyses augmented NEQ linearly, but appeared to have relatively little impact on computational requirements.

Case II. Corresponding results for the Angus data are shown in Figure 3. A number of combinations PCm_1/m_2 or XFAm_1/m_2 are examined, with m_1 and m_2 the number of common factors fitted for direct and maternal genetic effects. For NIT and CPU, values pertain to strategy B for MV10 and strategy A otherwise. With multiple random effects, differences between MV and either of the FA

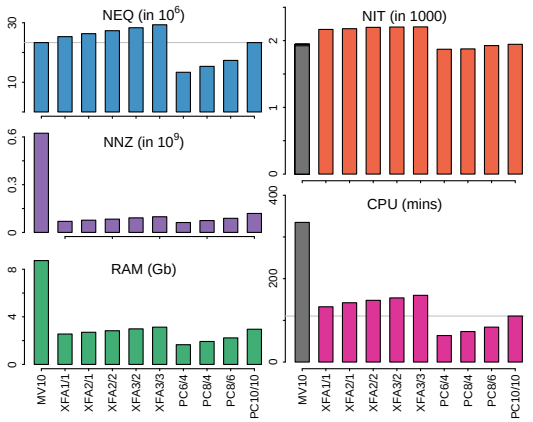


Figure 3. Results for Angus analyses (see text)

models were even more pronounced than for case I. Again, the number of common factors had little effect on NIT, but increased time per iterate and thus CPU proportionally. CPU time required for full rank PC was less than for any of the XFA analyses and only a third of that for the MV model. Resources required for reduced rank analyses are even less. E.g., omitting 2 direct and 4 maternal genetic factors reduced CPU to 75% of that for a full rank model.

Case III. Results for the large scale analyses for Herefords are given in Table 1. 'Best' values for NNZ (in million) pertain to countries numbered as described above, 'worst' used the reverse order. Using $\mathbf{\Gamma}$ with *mq* non-zero elements rather than $\mathbf{\Gamma}^+$, NNZ for PC9/9 would have been as large as $1\,005.6 \times 10^6$. A MV analysis holding MME in core was not feasible for this case. A corresponding analysis using an 'iteration on data' type strategy required more than 60 hours CPU time (Tier 2008; *pers.comm.*). Again PC analyses appeared advantageous over XFA implementations, and reducing the rank of fit even slightly had a dramatic impact on computational requirements.

DISCUSSION

Phocas et al. (2005) showed that animal model analyses would be preferable for international genetic evaluation of beef cattle. Computational requirements for such analyses can be large. However, as results clearly demonstrate, judicious choice of an equivalent model can greatly aid with this task. Parameterisation to a FA model, combined with a rotation of the factor matrix to triangular form, is especially beneficial for G×E scenarios where individuals have records in a single location only. Moreover, separation of genetic effects into common and specific factors is appealing, as estimates have a direct interpretation as global breeding values and local deviations. Results suggest that PC models are most advantageous computationally, even if little or no rank reduction is feasible. As outlined, there is a direct relationship between solutions from the different alternative models, i.e. estimates of 'global' breeding values are readily determined from those from a PC analysis.

Table 1. Global Hereford evaluation study

Model	NEQ	NNZ		RAM	NIT	CPU
		worst	best	(Gb)		(h)
XFA1/1	157.7	407.3	407.3	14.7	3051	29.8
XFA2/2	171.0	489.4	460.3	16.2	3110	32.0
PC9/9	144.4	899.8	386.7	13.9	3002	27.2
PC8/7	124.4	725.8	354.9	12.0	2690	20.6
PC7/6	111.1	599.9	294.1	10.5	2598	17.6

REFERENCES

Graser, H.-U. (2008) World Hereford Conf., Copenhagen, Denmark, June, 30 - July, 1.
Meyer, K. (2007) *J. Zhejiang Uni. SCIENCE B* **8**:815.
Meyer, K. (2009) *Genet. Select. Evol.* **41**:21.
Phocas, F., Donoghue, K. and Graser, H.-U. (2005) *Genet. Select. Evol.* **37**:361.
Smith, A. B., Cullis, B. R. and Thompson, R. (2001) *Biometrics* **57**:1138.
Thompson, R., Cullis, B. R., Smith, A. B. and Gilmour, A. R. (2003) *Austr. New Zeal. J. Stat.* **45**:445.

APPLICATION OF RANDOM REGRESSION TECHNIQUES TO DISSECT AGE-DEPENDENT QUANTITATIVE TRAIT LOCI FOR GROWTH IN LAMBS

G. Hadjipavlou[1,2] and S.C. Bishop[1]

[1]Roslin Institute and R(D)SVS, University of Edinburgh, Roslin, Midlothian EH25 9PS, UK
[2]Institute of Evolutionary Biology, University of Edinburgh, Edinburgh EH9 3JT, UK

SUMMARY

A systematic procedure for selecting random regression (RR) models that include quantitative trait loci (QTL) as time-dependent random effects, in addition to polygenic and environmental effects, was investigated using live weight data measured at 4-week intervals on lambs from 9 half-sib families of Scottish Blackface sheep. Microsatellite marker information was used to calculate gametic relationship matrices. The RR model allowed the detection of apparently significant QTL on chromosomes 14 and 20 when a RR polynomial of 2nd order or higher was fitted for the QTL effect. However, generally the choice of RR models with relevant variance components and the appropriate order for the random coefficient parametric curves proved complex for models both with and without QTL. With the given data structure, the RR approach had poorer ability to describe growth QTL over time compared to a growth model approach previously used to dissect the expression of age-dependent QTL for growth rates and live weights.

INTRODUCTION

Growth is a longitudinal trait that is a composite of growth rate phenotypes over time. Patterns of genetic correlations among live weights at different ages often demonstrate the trait complexity. For example, using sheep data Riggio *et al.* (2008) showed that inter-age genetic correlations for live weight, whilst strongly positive, are often less than unity, with the correlation decreasing as the time between weight measurements increases. Thus, it is likely that distinct loci act on live weights at different growth stages. For the detection of quantitative trait loci (QTL) that are associated with growth or live weight, it would be beneficial to simultaneously analyze multiple measurements and take account of the correlation structure among measurements ordered across time. Random regression (RR) methodology, in principle, provides a means of longitudinal trait analysis, while accounting for the covariance structure among measurements and allowing flexible model fitting. The main objective of this study was to examine the steps and issues relating to the use of RR to dissect QTL and to utilise RR models for chromosome-wide detection and quantification of QTL that influence longitudinal live weights, using data from Scottish Blackface sheep. We also sought to compare the results from applying RR for detecting age-dependent QTL with those of an alternative technique used previously for growth QTL mapping by analysing descriptors of growth derived from fitting a growth curve (Hadjipavlou and Bishop 2009).

MATERIALS AND METHODS

Data. The data comprised live weights at birth and at 4-week intervals up to 24 weeks, and genotypes for microsatellite markers on OAR 1-3, 5, 14, 18, 20, and 21, for 788 lambs from 9 half-sib families of Scottish Blackface sheep. A total of 135 markers were used (range=8-34, median= 15). Number of progeny per family ranged from 34 to 154 individuals, bred over 3 years (2001–2003). Standard records (parentage, day of birth, sex, rearing type) were collected. The pedigree comprised 1119 animals, including two generations of sire and dam ancestors.

Random Regression (RR) model. A RR model was used to fit the longitudinal live weight data first without accounting for a QTL and subsequently with the inclusion of a QTL effect. The generalized RR equation used for the full model is: $\mathbf{y= X\beta + Wq + Z_1u + Z_2p + Z_3l + Z_4e}$
where **y** is a vector of observations taken at several time points for each lamb; **β** is a vector of age-dependent fixed effects; **X** is the design matrix connecting fixed effects with records; $\mathbf{Wq}$, $\mathbf{Z_1u}$, $\mathbf{Z_2p}$, $\mathbf{Z_3l}$, $\mathbf{Z_4e}$ are the systematic time-dependent deviations from the fixed curve, modelled as random effects, due to allelic effects of the QTL, polygenic, permanent animal, litter (common environment) and residual effects, respectively. Vector **q** is of dimensions $2N_gp_1$, where N_g is the number of animals included in the gametic matrix. Vector **u** is of dimension N_ap_2, N_a being the number of animals in the relationship matrix (i.e. pedigree). The permanent animal vector **p**, the common environment vector **l** and the residual vector **e** are of dimensions N_pp_3, N_pp_4 and N_pp_5, respectively, where N_p is the number of animals with records. Parameters p_1, p_2, p_3, p_4, p_5 correspond to the number of random regression coefficients used to model the associated random effect. Matrices **W**, $\mathbf{Z_1}$ and $\mathbf{Z_2}$, $\mathbf{Z_3}$ are design matrices of the RR curve, including covariates. The elements of these matrices are $\Phi_i=\forall_i(t^*_{ij})$, where $\forall_i$ are coefficients of the chosen Legendre polynomial for lamb i at age j (t^*_{ij}). The age values are standardized between -1 and +1. Because heterogeneous residual variance was assumed, $\mathbf{Z_4}$ was fitted as a diagonal matrix of distinct variance for each age class. The random vectors **q**, **u**, **p**, **l**, **e**, are assumed to be mutually independent and to follow multivariate normal distributions; $\mathbf{q_i}|M, c_i \sim MVN(0,\mathbf{K_{0i}}\otimes\mathbf{Q_i}|M,c_i)$, $\mathbf{u} \sim MVN(0,\mathbf{G_0}\otimes\mathbf{A})$, $\mathbf{p} \sim MVN(0, \mathbf{P_0}\otimes\mathbf{I})$, $\mathbf{l} \sim MVN(0, \mathbf{L_0}\otimes\mathbf{I})$, $\mathbf{e} \sim MVN(0, \mathbf{I\sigma^2_{ek}})$ where $\mathbf{K_{0i}}$, $\mathbf{G_0}$, $\mathbf{P_0}$ and $\mathbf{L_0}$ are covariance matrices among random regression coefficients. Matrix **A** is the additive genetic relationship matrix and $\mathbf{Q_i}|M,c_i$ is the gametic relationship matrix of the allelic effects at the QTL, conditional on marker data (M) and the position (c_i) on the chromosome.

Model Choice and Statistical Testing. Prior to fitting a QTL, the Likelihood Ratio (LR) test was used to assess the significance of polygenic, permanent animal and litter effects across nested mixed RR models with varying order of the random regression polynomials for each effect. In all models, relevant identifiable fixed effects and all two-way interactions of month with these fixed effects were fitted. A fixed regression of 5th order, analogous to the population mean in single time-point analyses, and heterogeneous residual variances for the 7 monthly intervals were also fitted. The optimal RR model was then used to fit a QTL effect for each chromosome. The QTL effect was also modelled with random polynomial curves to allow for systematic effects of the QTL on the deviation of the animal phenotype from the expected value over time, and hence allowed for changes in the QTL variance with age. The gametic relationship matrix was calculated across all animals using Loki (Heath 1997), and it was utilised in the model to include the allelic effects of the QTL conditional on marker data and the position on the chromosome. LR tests were used to assess the significance of QTL with different order for the random regression polynomial on chromosomes 14 and 20. All RR model analyses were performed using ASReml.

RESULTS AND DISCUSSION

No-QTL RR model choice. Table 1 shows the log likelihood (LogL) estimates from a group of models in which a polygenic effect was fitted with a Legendre polynomial of order 1. Nested LR testing can only be performed across rows or down columns, and the df associated with each order of fit are 1, 3, 6, 10, 15 and 21, respectively. Because RR describes individual deviations from the average curve of the population, a 2nd order RR polynomial can be significant when orders 0 or 1 are not. Some models did not converge. In other cases in which the LogL did not improve when the order of either litter or permanent animal effect was increased, the higher order model was rejected as the likelihood maximisation process had probably failed to reach a global maximum.

Based on the above, all models with either of the two environmental effects fitted to order higher than 3 were rejected. From the remaining models, a model with a 2nd order RR polynomial fitted for both environmental effects was chosen to reduce the possibility of model overfitting. Inter-age genetic correlations estimated from this model were positive yet decreased as the interval between weight measurements increased, and were in agreement with those reported by Riggio *et al.* (2008).

Table 1 Log-likelihood estimates for no QTL RR model fitting live weights over time

Random regression polynomial order[1,2]		Permanent animal effect					
		0	1	2	3	4	5
Litter effect	0	-6299.87	-6335.83	-6131.82	-6130.64	NC[3]	NC
	1	-6302.58	-6393.58	-6192.8	-6194.48	NC	NC
	2	-6118.01	-6224.90	-6189.98	-6163.19	-6159.74	-6180.33
	3	-6130.09	-6243.11	-6178.88	-6194.03	-6201.74	NC
	4	-6172.57	-6286.41	-6177.15	-6199.01	-6231.62	NC
	5	-6219.32	-6332.14	-6195.83	-6211.04	-6249.87	NC

[1]A polygenic effect with a random regression polynomial of order 1 was fitted in each model.
[2]RR polynomial order is equivalent to order+1 random coefficients estimated.
[3]NC=Model fit did not converge.

QTL RR model choice and estimates. Chromosomes (OAR) 14 and 20 were explored as age-dependent QTL were previously found on these chromosomes, using live weights predicted from growth curves (Hadjipavlou & Bishop 2009). For each chromosome and each order of RR for the QTL effect, the likelihood of the full model was maximised every 5cM. Formal model testing included LR test and a conservative chi-squared test statistic were employed to assess the overall QTL significance, as well as the significance of increasing the order of the RR polynomial fitted for the QTL, at the chromosome position with the highest likelihood. The LogL and test statistics for an OAR 14 QTL are shown in Table 2. A first order Legendre polynomial for the QTL effect resulted in a model for which the QTL was not significant. Additionally, the QTL variance was not estimable (or estimated to be zero) at certain chromosomal positions. A significant QTL was supported by models with polynomial order of 2 or higher. Further, the statistic for increasing the order of the RR polynomial provided justification for fitting a QTL with a 3rd order RR.

Table 2 Log-likelihood estimates and test statistics for RR model with chromosome 14 QTL effect fitted using random coefficient polynomials

QTL RR order in model	LogL	Test statistic when compared to no QTL model[1]	Test statistic of increasing QTL order[1]
1	-6188.91	2.14 (3 d. f.)	NA
2	-6183.47	13.02 (6 d. f.)*	10.88 (3 d. f.)**
3	-6158.9	62.16 (10 d. f.)**	49.14 (4 d. f.)**
4	-6156.43	67.10 (15 d. f.)**	4.94 (5 d. f.)

[1]QTL effect: *significant (P<0.05); ** highly significant (P<0.01)

Figure 1 shows the proportion of phenotypic variance partitioned to the QTL effect across time in a model with a QTL RR of order 3. Since the estimated variance ratios across time differ only marginally, it is difficult to come to a conclusion regarding the trend in variance change across time. A time-dependence trend may be speculated, with a maximum effect around 60 days of age, but it is masked further by inflated variance estimates at the end age points, which arise as a

consequence of prediction using polynomial regression. A similar pattern of results was seen for OAR 20. Our results also showed confounding of the permanent animal and QTL effect, presumably because both terms are correlated with an animal's Mendelian sampling term.

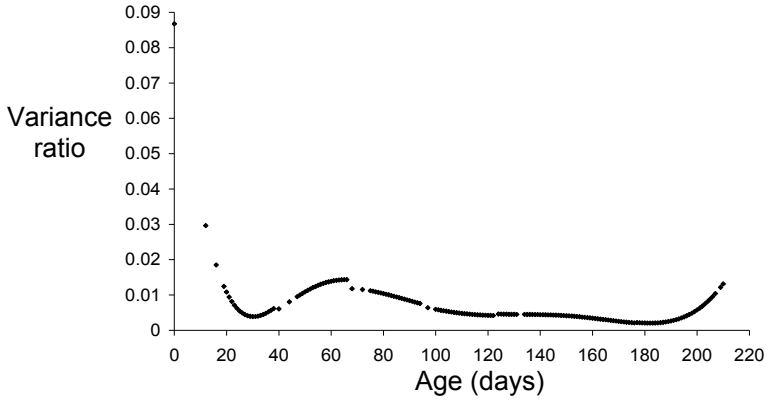


Figure 1. Proportion of phenotypic variance explained by the QTL, across age, on OAR 14.

Comparison between growth model and RR approaches and results. We have previously demonstrated the use of growth models to dissect the expression of age-dependent QTL. An OAR 14 QTL, which was only significant around 8 weeks of age, was detected and an OAR 20 QTL displayed a significance trajectory across age, and was significant from 8-16 weeks, with maximum significance at 12 weeks. The RR results support the above findings since on both chromosomes a QTL effect of order 2 or higher was deemed significant in a model in which polygenic and environmental effects were also modelled using RR. The fact that higher polynomial orders proved necessary to identify the QTL suggests a complex time dependence of QTL effects, as assessed by deviations from the expected age-dependent live weight. The pattern of QTL variance over time remains to be resolved, as results from our analysis are ambiguous.

CONCLUSIONS

Random regression techniques have been previously proposed for dissecting time-dependent polygenic and environmental effects, and could potentially be used to identify and describe the expression of age- or time-dependent QTL. Our study explored a systematic procedure for selecting random regression models that included a QTL as a time-dependent random effect, in addition to polygenic and environmental effects. High order RR models apparently detect time-dependent QTL, even though the pattern and magnitude of the change in QTL variance across time remains unclear. Using simulations, future work will assess the power of RR models for detection of age-dependent QTL, and explore the data requirements for such analyses.

ACKNOWLEDGEMENTS

This work was conducted as part of the GENACT Project, funded by the Marie Curie Host Fellowships for Early Stage Research Training, as part of the 6th Framework Programme of the European Commission. Defra funded the resource flock and genotyping.

REFERENCES

Hadjipavlou, G. and Bishop, S.C. (2009) *Anim. Genet.* **40**:165.
Heath S. (1997) *Am. J. Hum. Genet*. **61**:748.
Riggio, V., Finocchiaro, R. and Bishop S.C. (2008) *J. Anim. Sci*. **86**:1758.

COMPARISON OF GENETIC PARAMETERS OBTAINED FROM AN ORDINAL CANINE HIP PHENOTYPE DATA SET BY LINEAR AND ORDINAL ANALYSES

B.J. Wilson[1], F.W. Nicholas[2], J.W. James[2] and P.C. Thomson[1].

[1] Faculty of Veterinary Science, University of Sydney, PMB3 Camden NSW 2570
[2] Faculty of Veterinary Science, University of Sydney NSW 2006

SUMMARY

Many traits upon which selection may be desired are categorical traits recorded on an ordinal scale. Due to the relative ease and accessibility of linear analysis, researchers often set aside the categorical nature of these data sets and analyse them on the observed scale. Canine hip dysplasia is a common developmental disorder of the canine coxofemoral joint which results in significant pain and dysfunction in many affected animals. It is a multifactorial disease and genetic control is highly desirable. The hip dysplasia records used most commonly for selection in Australia comprise nine, radiographically assessed, ordinal traits with varying distributions, scored on each side of a dog's hips. Hip dysplasia and pedigree records from 13124 German Shepherd Dogs were analysed to compare heritability estimates and estimated breeding values obtained by analysing original or log-transformed scores with those obtained using ordinal logistic regression as well as binary logistic regression at each possible cut-point. All models incorporated the same fixed and random effects including pedigree structure. The results demonstrate that the ability of linear models to predict results of ordinal traits is variable and appears to be related to the distribution of the scores. It is therefore recommended that, where practical, the categorical nature of ordinal traits be accounted for in calculation of estimated breeding values and heritability estimates, particularly when the distributions of traits demonstrate significant skewness.

INTRODUCTION

Many important traits in animal breeding are recorded using discrete categories which have a natural order. Often these categories are assigned numerical labels based upon their order, and the categorical nature of these traits is often ignored when the traits are being analysed for use in a selection program. This could potentially lead to inaccuracies in the evaluation of the suitability of candidates for breeding and compromise the success of selection programs.

Selection against the multifactorial developmental joint disease, canine hip dysplasia (CHD), in Australia has traditionally followed the UK in the use of a hip score phenotype. This phenotype is based on radiographic examination of the dog's hips in young adulthood (typically around 19 months). Nine radiographic traits from each hip are each assigned a numerical score between the best (0) and worst (5 for one trait; 6 for the other traits). The nine scores for each hip are then added together to give a total for that hip, and the scores for left and right hip are then added together to give each dog a score between the ideal (0) and the worst (106). Selection advice is typically based on either the total score or the score of the worse hip, neither of which takes into account the ordinal nature of the data, nor the potential for underlying differences of distribution in each of the nine radiographic traits (Lawson 2000; Wood *et al*. 2004).

In this study we evaluate the extent to which the nine radiographic traits have similar distributions and compare heritability estimates and estimated breeding values (EBVs) calculated from the original and log-transformed data and from ordinal logistic regression as well as binary logistic regression at each possible cut-point.

MATERIALS AND METHODS

Data. A total of 13124 hip dysplasia scores collected from Australian registered German Shepherd Dogs (GSDs) born between 1980 and 2004 were collected and matched to pedigree information. The depth of pedigree information available for each dog with a record was variable, depending on when the dog was born and how recently its ancestors were imported into Australia. In addition, the animal's sex, age at radiography and year of birth were collected. Two ordinal scores (one from each hip) were available from each animal for each of the nine traits, resulting in 26248 scores for each trait.

Analysis. Genstat 10 was used to calculate the skewness of the distribution of the raw and log-transformed scores for each trait. ASReml 2 and beta versions of ASReml 3 were used to estimate the variance components and to calculate EBVs by fitting mixed models using residual maximum likelihood (REML) techniques (Gilmour *et al.* 2006). Models included fixed effects for gender, age (in months) at which radiography for scoring was undertaken, year of birth and hip (left or right). Random effects were a pedigree effect (additive genetic effect) and a permanent environmental effect to account for non genetic similarity in the left and right hip score of each dog. A linear model was fitted using untransformed (uLIN) and log-transformed scores (tLIN) for each of the 9 traits. A binary model (BIN) was fitted at each feasible cut-point for each of the traits. Finally, a multi-threshold model (ORD), which required the use of ASReml 3, was fitted for each trait, using as many thresholds as the program was able to include and reach a solution. Heritability estimates were obtained from ASReml for each analysis. Correlations between EBVs calculated by different methods were estimated using R software.

RESULTS

Distribution: All nine traits demonstrated positive skew, with traits 1 through 3 showing less skewness than traits 4 through 9 (see Figure 1). Log transformation decreased the skewness in all traits except trait 2, although some remained very significantly skewed (See Table 1).

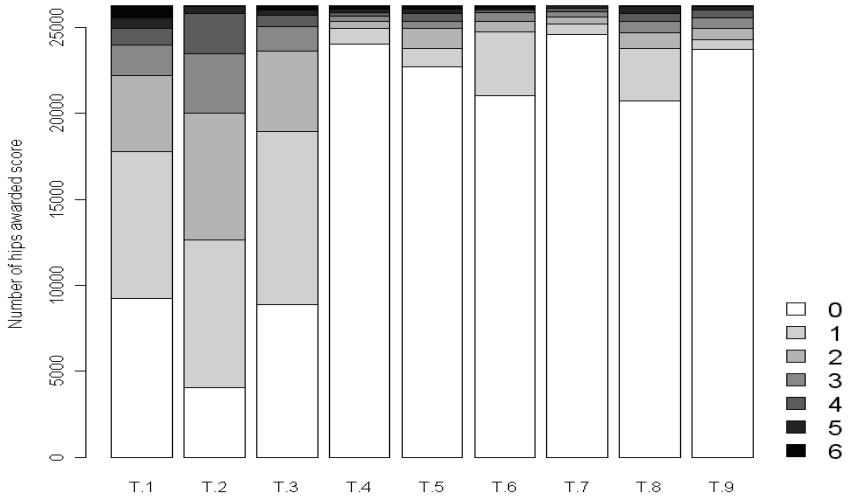


Figure 1: Distribution 26248 ordinal scores (from 13124 dogs) awarded for each of nine hip dysplasia traits.

Heritability Estimates: Heritability estimates are shown in Tables 1 and 2. Table 1 demonstrates that linear methods can result in estimates which are significantly different from estimates obtained using a multi-threshold analysis. Transformation appears generally to increase the similarity between linear and multi-threshold estimates, although the estimates may remain significantly different, e.g. for T.7 which remained very skewed despite the transformation. There does not seem to be a particular binary cut point that results in heritability estimates similar to that of the multi-threshold analysis.

Table 1. Heritability estimates plus standard errors for ordinal scores of nine radiographic traits analysed using raw and log-transformed linear analyses and a multi-threshold analysis (and skewness parameters for the score distribution)

Trait	Untransformed linear analysis	Log-transformed linear analysis	Multi-threshold ordinal analysis
T.1	0.30 ± 0.02 (1.45)	0.27 ± 0.02 (0.08)	0.26 ± 0.02
T.2	0.29 ± 0.02 (0.56)	0.23 ± 0.01 (-0.72)	0.25 ± 0.01
T.3	0.33 ± 0.02 (1.51)	0.26 ± 0.02 (-0.03)	0.27 ± 0.02
T.4	0.24 ± 0.02 (4.81)	0.18 ± 0.02 (3.54)	0.16 ± 0.02
T.5	0.27 ± 0.02 (3.47)	0.24 ± 0.02 (2.49)	0.24 ± 0.02
T.6	0.28 ± 0.02 (3.73)	0.22 ± 0.02 (1.99)	0.25 ± 0.02
T.7	0.40 ± 0.02 (5.13)	0.32 ± 0.02 (4.05)	0.24 ± 0.03
T.8	0.28 ± 0.02 (2.98)	0.26 ± 0.02 (1.87)	0.24 ± 0.02
T.9	0.21 ± 0.02 (3.70)	0.19 ± 0.02 (3.02)	0.21 ± 0.03

Table 2. Heritability estimates for nine ordinally scored radiographic traits obtained by binary logistic regression using different score cut points to define the outcome variable

	0and1	1and2	2and3	3and4	4and5	5and6
T.1	0.21[sd]	0.21[sd]	0.23	0.25	0.27	0.27
T.2	0.16[sd]	0.20[sd]	0.27	0.28	0.35	0.18
T.3	0.18[sd]	0.24	0.27	0.28	0.30	0.37
T.4	0.14	0.24	0.23	0.28[sd]	0.33[sd]	0.35
T.5	0.23	0.24	0.25	0.25	0.30	0.39
T.6	0.22	0.22	0.22	0.32	0.40	0.30
T.7	0.24	0.24	0.23	0.24	0.27	-
T.8	0.21	0.21	0.20	0.20	0.30	0.30
T.9	0.21	0.17	0.18	0.26	0.35	0.32

[sd] Significantly different ($P < 0.05$) from multi-threshold ordinal estimate

Estimated Breeding Values: The relationship between the skewness of the scores and the correlation between EBVs obtained by the ORD and uLIN analyses and those obtained by the ORD and tLIN analyses is shown in Figure 2. It demonstrates a clear trend for the utility of linear analyses to decrease as skewness increases.

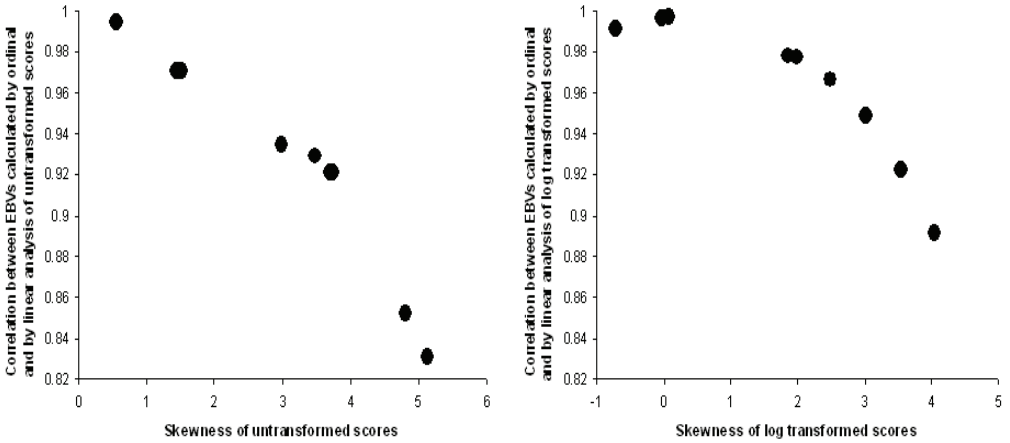


Figure 2: Relationship between the skewness scores and the correlation between EBVs calculated by linear and multi-threshold methods for untransformed linear scores (left) and log transformed linear scores (right).

DISCUSSION

The results show that while calculation of EBVs and heritability estimates for ordinal scores using linear mixed models may result in reasonable estimates in some cases, there may also be appreciable differences in the results. Analysis of this real data set showed that a log transformation resulted in heritability estimates more similar to those estimated by multi-threshold procedures. For one especially skewed trait, however, the result remained significantly different. The binary model heritability estimates were not demonstrably better than the transformed linear model estimates. The correlation between EBVs from linear and multi-threshold models showed a strong negative association with skewness. While correlations for traits with minimal skew were very good, further work is needed to assess the extent to which there might be individual animals for which substantially different EBVs are obtained from the two approaches; and the effect which this may have on the effectiveness of a selection program. It is therefore recommended that when calculating EBVs and heritabilities for ordinal traits, a multi-threshold model be implemented using an ordinal logistic mixed model. If linear models must be used, then a correction for skewness appears to be advantageous.

REFERENCES

Gilmour, A.R., Gogel, B.J., Cullis, B.R. and Thompson, R. (2006) "ASReml User Guide Release 2.0" VSN International Ltd, Hemel Hempstead.

Lawson, D.D. (2000) In "Hereditary Bone and Joint Diseases in the Dog", p.267, editors J.P. Morgan, A. Wind and A.P. Davidson, Schlütersche GmbH & Co, Hannover.

Wood, J.L.N., Lakhani, K.H. and Henley, W.E. (2004) *The Veterinary Journal* **168**:14

A FRAMEWORK TO LINK WHOLE GENOME SNP ASSOCIATION STUDIES TO SYSTEMS GENETICS

S.J. Goodswen [1, 2], H.N. Kadarmideen [1], C. Gondro [2] and J.H.J. van der Werf [2]

[1] CSIRO Livestock Industries, JM Rendel laboratory, PO Box 5545, Rockhampton Mail Centre, Rockhampton, QLD 4702
[2] School of Environmental and Rural Sciences, University of New England, Armidale, NSW 2351

SUMMARY

One of the main outcomes from whole genome association studies (WGAS) are statistically significant single nucleotide polymorphisms (SNPs). This outcome would be enhanced if either biological meaning or potential functional roles for these SNPs are revealed. Such knowledge would enhance the process of marker- or gene-assisted selection (MAS/GAS), or at least would provide more understanding of the achieved genetic change. We develop and present an R package called *FunctSNP*, interfaced to a relational database, to add potential functional information to SNPs that have shown to be statistically significant in a WGAS. The approach and developed software has been designed in such away that the database can contain any model species.

INTRODUCTION

Whole genome association studies with genotyped SNPs are a set of methods to identify which SNPs are associated with variation in a particular complex trait of interest. There are essentially two reasons why WGAS has gained in popularity. Firstly, the huge increase in the availability of SNPs. For example, the dbSNP database housed by the National Center for Biotechnology Information (NCBI) contains millions of SNPs for many model species. Secondly, the falling cost and increased capacity of high-throughput SNP genotyping chips has made WGAS feasible and affordable.

The significant SNPs derived from WGAS can in effect be classified into 4 types: (1) a SNP that contributes to variation in the complex trait; (2) a SNP with no known biological effect but in linkage disequilibrium (LD) to an untagged SNP (not genotyped) that contributes to variation in the complex trait; (3) a SNP with an association only - no known biological effect or linkage to a causal SNP; (4) a SNP *not* associated with the complex trait (a false positive). These significant SNPs are currently used in two conventional, albeit alternative, approaches depending on the requirements of the researcher: (1) a molecular biologist usually searches for candidate genes in the region of the significant SNP and then does functional experiments, which could include targeted sequencing; (2) an animal breeder or geneticist uses significant SNPs as DNA markers for associated trait selection (MAS/GAS).

Although WGAS are the best methods we currently have, literature states that the most notable limitation of WGAS is the potential for false-positive results (Pearson and Manolio 2008). Since the type of significant SNP following WGAS is not distinguishable to the researcher, there is a potential to perform functional experiments based on a false-positive SNP and/or use a SNP as a marker that is neither the causal variant nor in LD with the causal variant.

At present, the only approach to truly validate a significant SNP is to repeat the WGAS on a different population. *FunctSNP*, however, can assist the researcher in distinguishing the type of significant SNP by providing information as to whether a SNP is: (1) a potential causal variant; (2) close to a SNP that is a potential causal variant; (3) close to a gene that has the potential to affect the complex trait [see Program functionality section for more details].

DESIGN AND IMPLEMENTATION

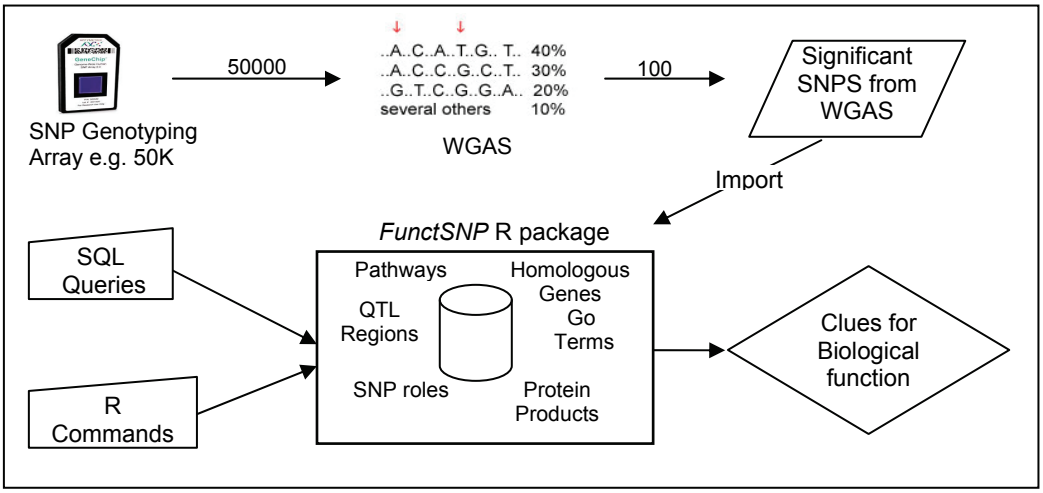


Figure 1. Schematic of *FunctSNP* R package post-WGAS

The *FunctSNP* R package (Figure 1) interfaces to a localised relational database that contains the most pertinent SNP related information for a chosen model species from all the sites listed in Table 1. Most of these sites are intuitive and informative. A researcher, through the use of Internet addresses, hyperlinks, and cross-linked data can navigate from site to site with ease; however, they cannot integrate related data in one query. So there are essentially two main reasons for using *FunctSNP*: The power of a statistical computing program, interfaced with a *relational* database. R is a free software environment for statistical computing and graphics (http://www.r-project.org/). A relational database is structured data in the form of two-dimensional tables. Each table is linked by a defined relationship and consequently all data is integrated. The power comes from a language designed to query the database called Structured Query Language (SQL).

Program creation. Our first database design consideration was which Relational Database Management System (RDMS) to use. We opted for the open source SQLite which stores data in a file that is platform independent and is the preferred RDMS for R (SQLite available from http://www.sqlite.org/). Next, we were faced with the challenge that there is currently no standardized format for transporting data from one database to the next, although extensible markup language (XML) and BioMart are gaining prominence. The data formats from each site were: NCBI - Microsoft SQL Server, GO – MySQL, KEGG – XML, UniProt – XML QTLdb – ASCII text file, OMIA – MySQL, and Homologene - ASCII text file. Generic C++ programs and Perl scripts were written to convert the data formats to a format compatible with SQLite.

There are essentially 4 manual steps to create a local database: (1) download the data from the online resource; (2) decompress file(s) if required; (3) convert data into a format that can be imported to a relational database; (4) import the data into the database. These 4 steps need to be repeated for each online resource. We have developed software to provide a framework in which these 4 steps are automatically repeated to create a local database with data extracted from any number of resources. To ensure the local database is always up-to-date the automatic creation can

be scheduled to run on a weekly or even daily basis. The principle idea is that once the schedule is set up no further human intervention is required.

At the time of writing this paper, *FunctSNP* is not publically available. It is anticipated that *FunctSNP* will be available for download from a CSIRO HTTP server. To the best of our knowledge there are only a few SNP software tools designed specifically for the livestock industry. Therefore it is expected that the first version of *FunctSNP* will be for *Bos taurus.*

Table 1. Internet sites containing biological data

Acronym	Name	Link	Resource
dbSNP	Single Nucleotide Polymorphism database	http://www.ncbi.nlm.nih.gov/	SNPs
GO	Gene Ontology	http://www.geneontology.org/	Genes and gene product attributes
KEGG	Kyoto Encyclopaedia of Genes and Genomes	http://www.genome.jp/kegg/	Biological pathways
UniProt	Universal Protein Resource	http://www.uniprot.org/	Protein sequences and functional information
QTLdb	Animal Quantitative Trait Locus database	http://www.animalgenome.org/QTLdb/	Quantitative Trait Loci data (QTL)
OMIA	Online Mendelian Inheritance in Animals	http://omia.angis.org.au/	Genes, inherited disorders and traits
HomoloGene		http://www.ncbi.nlm.nih.gov/homologene	Homolog detection

Program functionality. The primary function of *FunctSNP* is to provide information about the potential biological functions of SNPs or genes close to SNPs. Information such as SNP chromosomal location, exon/intron status, synonymous/non-synonymous effect, SNPS in Quantitative Trait Loci (QTL) regions, and biological pathways, GO terms, and protein products for related genes. There are two ways to access the information: (1) Direct access to the database using SQL queries through the SQLite program, and (2) using a set of R functions (currently 20) to answer preset questions. For example, "Give me all gene IDs associated with significant SNPs" and the output is a 2 dimensional array for further analysis with R commands.

The database is both SNP ID (using NCBI's dbSNP rs# cluster ID) and gene ID (using NCBI's gene ID) centric. For example, given a SNP ID as an input, the output is the gene ID on which the SNP resides. The gene ID provides the link to *every* gene attribute in the database such as name, function, protein products, and biological pathways. Alternatively, given a gene ID as an input, the output is a list of SNPs located on the gene. The SNP ID is then the link to every SNP attribute such as exon/intron status, and chromosomal location.

Not all significant SNPs from WGAS are located on genes. Therefore *FunctSNP* provides the functionality to find the nearest genes to such SNPs. The search distance is base pairs (bp) and is user defined. For example, enter 100 bp and the program returns the gene ID for any SNP residing less than 100 bp (e.g. a promoter region) from the gene's transcription start site. In a similar manner to searching for nearest genes, *FunctSNP* provides information as to whether a SNP is close to another SNP that is a potential causal variant. For example, search for all non-synonymous SNPs within a user defined distance from a significant SNP.

There is much more publically available data on some species than others. So for some species there may be no insightful data on the genes identified as linked with significant SNPs. *FunctSNP* provides the ability to find homologous genes across all species or a specified species. For example, for identified *Bos taurus* genes, we can obtain gene IDs for homologous genes from

Homo sapiens. These gene IDs can then be uploaded into other analysis programs such as DAVID (http://david.abcc.ncifcrf.gov/summary.jsp).

To aid interpretation of the output from *FunctSNP*, we intend to provide a scoring system for WGAS significant SNPs e.g. SNPs which reside on an exon region with a non-synonymous effect will be given the highest rank. A normalised score (taking into account such factors as sample size, SNP chip capacity) will be adopted to make it comparable with WGAS of the same population. There will also be the option to search for the highest scoring SNPs within a user defined distance from a significant non-causal SNP.

As part of the ongoing development, the Internet will be monitored for new or overlooked sites which contain informative publically available biological data that may be incorporated into *FunctSNP*. Also, future development will be focused on how best to interpret and present the results. For example, taking the genes associated with the SNPs and applying a gene set enrichment analysis (GSEA).

Program testing. To test *FunctSNP*, we used real 10k SNP genotype data from a yet to be published WGAS. From this 10K SNP data, 165 SNPs had been identified as significantly affecting a particular trait of economic interest. The NCBI reference numbers (rs#) for the 165 significant SNPs were imported into our database and in less than five minutes a series of reports using *FunctSNP* R functions were generated. From the 165 SNPs, 49 were located in a gene region. One of these 49 SNPs was located on an exon and was synonymous. For the 49 genes, there were 81 GO terms, 20 KEGG pathways, 355 homologous genes from 17 different species, and 32 homologous genes from the same species. A total of 12 genes downstream and 10 genes upstream were found less than 10,000 bp from the non-coding significant SNPs.

CONCLUSIONS

To take full advantage of WGAS, we need to make that essential link between the outcomes from WGAS and the information that exists about the function of genes and pathways. We conclude that *FunctSNP* is a post-WGAS tool that provides an opportunity to screen and select for SNPs that have a higher likelihood to be related to variation in a particular complex trait of interest.

ACKNOWLEDGEMENTS

Office of the Chief Executive (OCE) of CSIRO for postgraduate scholarship.

REFERENCES

Pearson, T. A. and Manolio, T. A. (2008) *JAMA* **299**:1335.

STATISTICAL CONSIDERATIONS IN THE ANALYSIS OF GENE EXPRESSION DATA FROM HETEROGENEOUS SOURCES

P.C. Thomson, M. Singh, and H.W. Raadsma

ReproGen, Faculty of Veterinary Science, University of Sydney, Camden NSW 2570

SUMMARY

Combining gene expression data from heterogeneous sources has the potential to increase our understanding of comparative genomics, amongst other things. However, the statistical analyses must address inherent differences across experiments before a combined analysis can be undertaken. A procedure to achieve this is described here, and the methods are illustrated by an analysis of gene expression data in the sheep and cow looking at three stages of lactation.

INTRODUCTION

Increasingly gene expression arrays are being used as a tool for understanding the genetic architecture of complex traits. Whilst initial work was mainly focused on human and model animal species, there is now considerable attention to the application of these techniques in livestock. One particular area that has not received much attention hoverer is the integration of gene expression data from heterogeneous sources, such as across different species, or across different platforms within the same species. However, there are many potential advantages if a combined analysis using gene expression data from multiple sources can be conducted. For example, a combined analysis across species will lead to a better understanding of comparative genomics, and may also elucidate how useful a model species is, as in the case of a sheep as a small-ruminant model for the cow. Further, situations arise where researchers have access to gene expression data from more than one platform, and it would be desirable to explore these data to get an overall picture of gene expression, as well as to understand possible differences between platforms, and incorporate these in formal meta-analyses.

Combining data from heterogeneous sources requires careful statistical consideration, and shares some of the same issues involved with meta-analysis of genomic data. The following describes a three-stage process that was developed for the analysis of a component of a lactation genetics study conducted by the CRC for Innovative Dairy Products. In particular, the methods developed will be illustrated by an application to a comparative study into patterns of gene expression over the lactation cycle of cows and sheep, with values recorded pre-peak, peak, and post-peak lactation from samples derived from mammary gland tissues using a common bovine gene expression array developed by Affymetrix. This array has been shown to have utility in both cattle and sheep and represents over 20,000 gene targets.

METHODS

The first stage involves normalisation of the raw expression data across the series of expression of arrays. In some situations, it may be possible to perform a joint normalisation using all data simultaneously by processing this through a standard procedure, such as RMA-normalisation (Irizarry *et al*. 2003). This is necessary to align the distributions (across genes) of all expression chips to be similar, and hence comparable. However, an additional manual step of quantile-normalisation will usually need to be applied to complete this process. Firstly, to introduce some notation, assume that there are two heterogeneous groups to consider, ovine and bovine in the current situation, but this may be generalised to multiple heterogeneous data sets. Assume we have expression arrays $i = 1, 2, \ldots, n_1, n_1 + 1, n_1 + 2, \ldots, n_1 + n_2$, where there are n_1 and n_2 arrays

respectively for group 1 and 2, and that y_{ij} represents the (normalised) expression data for array i at gene j, $j = 1, \ldots, g$. Then proceed as follows:

1. For each expression array, sort the (normalised) expression values from smallest to largest, say $y_{i(1)}, y_{i(2)}, \ldots, y_{i(g)}$, $i = 1, \ldots n_1 + n_2$.
2. Obtain the means for each of these g order statistics by averaging over all the arrays, i.e. $\overline{y}_{(1)}, \overline{y}_{(2)}, \ldots, \overline{y}_{(g)}$. Similarly, obtain these g mean order statistics for each group by averaging over the n_1 and n_2 arrays for Groups 1 and 2, say, $\overline{y}_{1(1)}, \overline{y}_{1(2)}, \ldots, \overline{y}_{1(g)}$ and $\overline{y}_{2(1)}, \overline{y}_{2(2)}, \ldots, \overline{y}_{2(g)}$.
3. Adjustment is made separately for each group, by means of linear interpolation for each expression level, y_{ij}. The interpolation uses the following before-after (x,y) pairs: $\{(\overline{y}_{1(1)}, \overline{y}_{(1)}), (\overline{y}_{1(2)}, \overline{y}_{(2)}), \ldots, (\overline{y}_{1(g)}, \overline{y}_{(g)}), \}$ for Group 1, and a similar set for Group 2. So if the value of y_{ij} in Group 1 lies between $\overline{y}_{1(k)}$ and $\overline{y}_{1(k+1)}$, then the adjusted expression value is calculated as $y_{ij}^{(\text{adj})} = \overline{y}_{(k)} + (y_{ij} - \overline{y}_{1(k)}) \times (\overline{y}_{(k+1)} - \overline{y}_{(k)}) / (\overline{y}_{1(k+1)} - \overline{y}_{1(k)})$, with a similar procedure for Group 2 arrays. A special contingency is required to handle the extreme expression values below $\overline{y}_{1(1)}$ or $\overline{y}_{2(1)}$, or above $\overline{y}_{1(g)}$ or $\overline{y}_{2(g)}$.

The result is that the overall distribution of expression levels for the two (or more) groups will be the same, thus allowing direct comparisons. Code to undertake this step has been written in R with the `approx()` function an efficient means to perform the interpolations.

The second stage involves estimating the effects of each gene across the different states, and this is achieved by fitting a single large-scale linear mixed model to all the expression data. Specifics of fixed and random effects to be included in the model will depend on the particular study, but the general rule is to include all relevant sources of variation in the model. This in an extension to the linear mixed model technique described in Sharp *et al*. (2008). A typical linear mixed model will be of the form

$$y = \mu + \text{Array} + \text{Gene} + \text{Gene.Group} + \text{Gene.State} + \text{Gene.Group.State} + \varepsilon$$

where $y = \log_e(\text{RMA})$, the normalised expression values; μ = overall mean expression value; and Array = fixed effect of array i, $i = 1, \ldots, n_1 + n_2$. All the remaining terms are random, namely, Gene = effect of gene j, $j = 1, \ldots, g$; Gene.Group = effect of gene in a particular group; Gene.State = effect of gene in a particular state; Gene.Group.State = effect of gene in a particular group-state combination; and ε = random error. Here, State refers to one of the comparisons of interest, e.g. expression levels at pre-peak, peak or post-peak of lactation. The sum of the Gene.State and Gene.Group.State BLUP solutions of the random effects will be used to assess the effect of a gene for subsequent analysis. ASReml (Gilmour *et al*. 2006) can be used to fit such models.

The third stage involves assessing which genes are differentially expressed (DE) in different states. Histograms of these effects for each Group×Stage combination show a heavy-tailed distribution, suggesting the existence of two groups of effects, one with a large variance (DE genes) and one with small variance (non-DE genes) (Figure 1). This is modelled is a two-component mixture distribution, with the DE component being $N(0, \sigma_1^2)$ and the non-DE component $N(0, \sigma_0^2)$ with $\sigma_1^2 > \sigma_0^2$. The (prior) probability of any gene being DE is π_1. The mixture model is fitted to each Group×Stage set of gene effects using the E-M algorithm (McLachlan and Basford 1988) and returns estimates of π_1, σ_1^2 and σ_0^2. The posterior probability (τ_j) of the gene being DE in that particular Group×Stage combination is calculated by an application of Bayes rule, $\tau_j = \pi_1 f_1(z_j) \Big/ \left[\pi_1 f_1(z_j) + (1 - \pi_1) f_0(z_j) \right]$, where $f_1(\cdot)$ and $f_0(\cdot)$ are the

normal probability density functions for the DE and non-DE components, respectively, and z_j is gene j effect. Values of $\tau_j > ½$ indicate the gene is more likely DE than non-DE. However, in order to reduce false positives, it is preferable to select a higher threshold, such as $\tau_j > 0.8$.

Having identified the subset of genes that are DE in a particular Group×Stage combination, comparisons across groups may then be of interest. For example, when groups may represent different species, it is of interest to see which genes are DE in both species, and to investigate the direction (up- or down-regulation of effects).

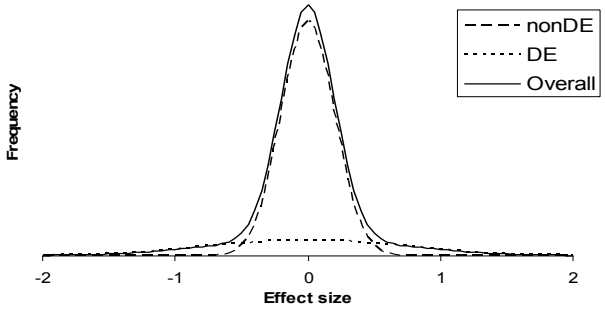


Figure 1. Mixture model for the distribution of DE and non-DE genes.

RESULTS AND DISCUSSION

As indicated above, the methodology outlined here will be illustrated by two gene expression studies conducted into lactation in sheep and cattle. Both studies examined three time points (states), pre-peak, peak, and post-peak lactation, and bovine Affymetrix arrays were used in both studies, For the sheep experiment, four arrays were used per time point, and for cattle five were used (although one array for the post-peak was not usable).

Despite joint RMA-normalisation of all 26 arrays, differences were still apparent between the ovine and bovine arrays, as shown in by the boxplots in Figure 2A. The additional normalisation step using the linear interpolation method was applied and the resultant distributions of adjusted expression values are shown in Figure 2B. Clearly the distributions are now very similar.

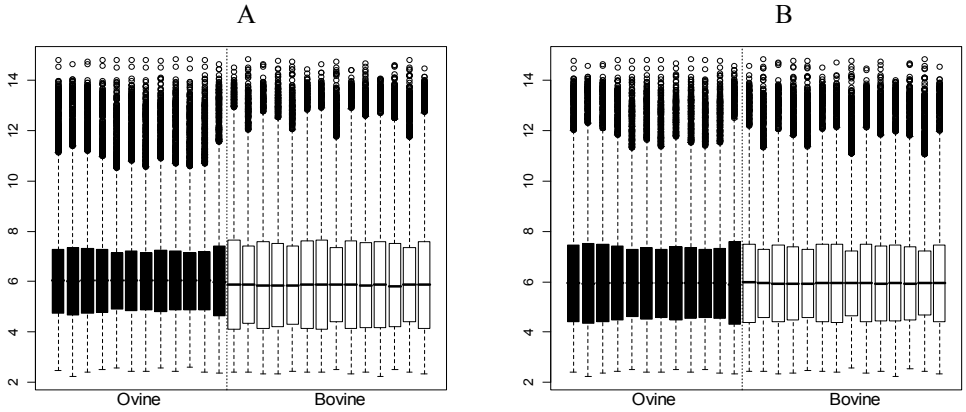


Figure 2. Boxplots of the distributions of expression values before (A) and after (B) the additional linear interpolation normalisation step.

Using the large-scale linear mixed model, BLUP solutions of the random effects were obtained, and the six sets (species × period) of genetic effects determined. In the process of fitting

the mixture model, the probability of it being DE is calculated (τ_j), and these values for sheep in the pre-peak are shown in Figure 3. This shows a typical pattern with a small cluster of genes with a very high probability of being DE ($\tau_j \sim 1$), although any gene with $\tau_j > 0.8$ is considered DE.

These estimated effects (BLUP values) were obtained for both sheep and cow samples, and Figure 4 shows a smoothed scatter plot of sheep and cow genetic effects in the pre-peak phase. It is seen that many genes have similar expression patterns across sheep and cow, as indicated by the large strong positive association.

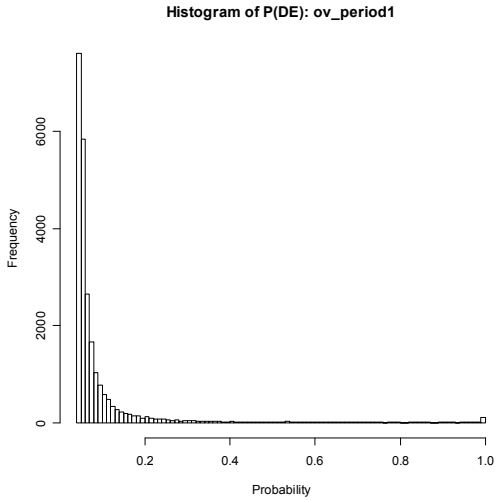


Figure 3. Distribution of the probability of a gene being DE for sheep in the pre-peak lactation stage

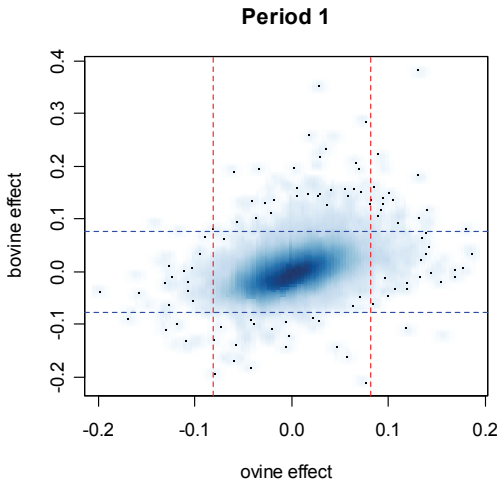


Figure 4. Joint distribution of genetic effects for sheep and cow in the pre-peak lactation stage. Vertical and horizontal dashed lines indicate thresholds for genes being DE.

CONCLUSION

Combining expression data from heterogeneous sources, be they from different species, tissues, or perhaps experimental platforms, has the potential to add considerable insight into our understanding of gene function, and will add additional value over what can be provided by studying gene expression patterns from a single isolated experiment. The methods outlined here provide a three-stage procedure that will allow these meta-analyses of expression data to be undertaken to move to an entire transcriptome analysis of a particular physiological state or organism.

REFERENCES

Gilmour, A.R., Gogel, B.J., Cullis, B.R., and Thompson. R. (2006) "ASReml User Guide. Release 2.0". VSN International Ltd, Hemel Hempstead.

Irizarry, R. A., Hobbs, B., Colin, F., Beazer-Barclay, Y.D., Antonellis, K.J., Scherf, U., and Speed, T.P. (2003). *Biostatistics* **2:**249.

McLachlan, G.J., and Basford, K.E. (1988) "Mixture Models: Inference and Applications to Clustering". M. Dekker, New York.

Sharp, J.A., Mailer, S.L., Thomson, P.C., Lefevre, C., and Nicholas, K.R. (2008) *Molecular Cancer* **7**:1.

THIRTY YEARS OF AAABG

Ian Franklin

CSIRO Livestock Industries, Belair, SA

SUMMARY

AAABG was founded to stimulate animal breeding research and its adoption by the various livestock industries. To this extent, it has been successful; since the inaugural conference 30 years ago, over 2000 papers, both invited and contributed, cover most the major issues. This paper presents a brief and personal overview of issues and successes, most of which were resolved during the first fifteen years, and of the challenges ahead.

INTRODUCTION

> In many ways the lot of the theoretical population geneticist … is a most unhappy one. For he is employed, and has been employed for the last 30 years, in polishing with finer and finer grades of jeweller's rouge, those three colossal monuments of mathematical biology, *The causes of evolution*, *The genetical theory of natural selection* and *Evolution in Mendelian populations*. (Lewontin, 1964)

Lewontin, of course, was talking about the contributions of Haldane, Fisher and Wright to population genetics theory. Equally, the reference could apply to quantitative genetics. Fisher, and to a lesser degree Wright, laid the foundations of quantitative genetics: Fisher in his 1918 and many subsequent papers on design and analysis of experiments, and Wright for his work on genetic relationships, the effects of small population size and his influence on J. Lush. However, in contrast to population genetics, quantitative genetics always had a strong empirical foundation, and clear commercial applications.

Population genetics is concerned with changes in the frequency of genes in populations and the mechanisms underlying the maintenance of the standing variation. Quantitative genetics, on the other hand, focuses on genetic and phenotypic means and variances of traits in populations. Clearly, the two disciplines are related. By concentrating only on the mathematical and statistical aspects, however, both run the risk of ignoring the underlying biology. In recent years, molecular genetics has reinvigorated population genetics; quantitative genetics is approaching a similar renaissance. I shall return to this later.

For much of the 20th century, animal breeding research concentrated on refining the statistical methodology for estimating components of variance and developing breeding programs based on these components. By the 1970s, despite years of effort by researchers, extension officers and some producers, penetration of modern animal breeding technology to the extensive industries of Australia was still poor. Advances were evident in the pig and poultry industries, and in dairy, as most states had developed dairy herd improvement programs. Nevertheless, even in these industries, estimates of rates of realised genetic progress were rare. In the extensive industries, there were none. BLUP methodology, now incorporating the inverse genetic relationship matrix, offers a means to estimate realised progress, given accurate estimates of genetic parameters.

It was clear that an effective means to present an already well established animal breeding theory to extension officers and, especially, to producers, was desirable. Neither the Genetics Society nor the Australian Society of Animal Production was an adequate forum for this.

1979 -- THE INAUGURAL MEETING

Stuart Barker opened the inaugural conference with an historical account of animal breeding research in Australia, the motivation for establishing the Association, and established the

groundwork for future meetings. Paddy Cunningham delivered a summary of the current understanding of quantitative genetics (in many respects just as relevant today), and its relation to animal breeding. In this, and in almost meetings that followed, special sessions were devoted to the current status of breeding programs for each of the most commercially significant species.

This conference, and most of those that followed, had a theme (Table 1). Here, the theme was measurement. Of course, animal improvement is possible without objective measurement, and some degree of visual appraisal continues to this day, but the introduction of objective measurement (and reliable pedigrees) to the various extensive industries had in the past been a major stumbling block to the application of quantitative genetics theory.

Table 1. Location and themes of each AAABG conference over 30 years

Year	Volume	Site	Theme
1979	1	Armidale	Measurement and recording
1981	2	Melbourne	Selection and mating programs
1982	3	Brisbane	Efficiency
1984	4	Adelaide	Implementation
1985	5	Sydney	Various
1987	6	Perth	Various (breeding objectives)
1988	7	Armidale	Economics
1990	8	Hamilton, NZ	Technology transfer
1991	9	Melbourne	Profit and Prophets
1992	10	Rockhampton	International trade
1995	11	Roseworthy	Quality and profit
1997	12	Dubbo	Responding to client needs
1999	13	Mandurah	Breeding for 21st century
2001	14	Queenstown, NZ	Biotechnology
2003	15	Melbourne	50 Years of DNA
2005	16	Noosa Lakes	New Genetic Technologies
2007	17	Armidale	Making it happen
2009	18	Barossa Valley	Matching genetics and environment

Over the next 15 years, most of the major issues were resolved. Statistical methods and algorithms to take advantage of rapid changes in computer technology were developed and made available to researchers, and to the industries, for estimating genetic parameters, or for managing and implementing breeding programs. A National Dairy Herd Improvement Scheme was established (Jones 1991), programs such as Breedplan (Nicol *et al*. 1985), Woolplan (Brien 1990) and Lambplan (Banks 1990) were in place. Sire evaluation programs were created for sheep and cattle and made available to producers (James 1979; Roberts 1979). These were crucial for improving productivity and quality in these species.

The following years saw discussion of some of the finer points of implementation, the status of emerging industries, trade and the potential for the new DNA based technologies. Clearly, I cannot discuss in any detail the 2000 odd papers presented over the 30 years of AAABG. In any case, to many of these I can add no worthwhile opinion. Rather I restrict myself to a subset of issues; those that I believe represent the essential elements of any breeding program. I can do no better than quote Bill Hill at the 1981 meeting:

> With *well defined* breeding objectives, *reliable* estimates of genetic parameters such as heritabilities and genetic correlations and an *adequate* knowledge of the biology of the species, it is not difficult to construct a feasible breeding program *predicted* to have near-optimal rates of progress over a *few* generations.

Readers may note that I have inserted an emphasis of my own.

WELL DEFINED BREEDING OBJECTIVES

The first of these is the most problematic, and numerous AAABG sessions have been devoted to this topic. Many authors have discussed the complexities of this issue (eg Barlow 1987). However, since considerable progress can be achieved with an objective not based on detailed economic analysis, it is not a precondition for beginning a breeding program. Indeed, it may be the last element to be put in place. The primary goal of the commercial livestock producers is to improve enterprise efficiency. Each suffers from a deteriorating environment, as they have to compete for market share not only with members of their own industry but also with other products, with pressure from retailers and with the changing demands from consumers. In other words, producers need, like the Red Queen, to run as hard as they can to stay in the one place. Often, the major benefit accrues to the retailers and the consumers.

Ideally, one needs to develop an economic model for each industry - one that allows inputs for each circumstance and helps to define the variables that are critical for economic survival. In practise, a truly predictive model is almost certainly non-linear and probably impossible, as it is difficult to predict vicissitudes in consumer demand (*e.g.* changing fashions for egg colour or egg size), in ethical attitudes and of course, the weather. The model needs, therefore, to accommodate risk management (see Anderson 1988). Then, this economic model needs to be linearised at the current status of the enterprise, or at some future goal. It is no wonder that breeders and their advisors often choose to adopt a desired gains model, or more dangerously, choose a set of traits that they believe covers their needs. In practise, then, the breeding objective is commonly defined as a linear sum of traits that we wish to improve, each weighted by a set of economic values.

The greatest dangers arise when all relevant traits (including those not easily measurable) are not included in the objective (Hill 1981), or when the objective is based on individual performance rather than the enterprise (Cartwright 1982). Finally Smith (1988), following Smith *et al.* (1986), argues that, "genetic improvement should not be used to correct inefficiencies in the system". It seems that bounds need to be set on this principle. Some issues, such as stocking rates, are clearly managerial decisions. Others are less clear. Should we, for example, assume optimal parasite control when the relationship between resistance and cost of control is a step function? Also, consider a flock producing fibre of a diameter that markedly mismatches market demand for apparel wool. It could be argued cogently that this is an example of a managerial inefficiency, and that the producer should engage in breed substitution (or upgrading) rather than waste valuable selection differential on reducing fibre diameter.

While not entirely resolved, the debate over this issue during the early life of AAABG, has certainly had an effect – compare, for example, Ponzoni (1979) with Ponzoni (1988).

RELIABLE ESTIMATES OF GENETIC COMPONENTS

Perhaps this is the area where we have seen the greatest advance over the last 30 years. Initially, genetic variances and covariances were estimated using carefully designed experiments (*eg* half sib families) and the analysis of variance. Now we have an armoury of statistical and computing tools, using maximum likelihood, Monte Carlo, and other optimisation algorithms, to estimate more directly the desired genetic components. Most readers will be aware of packages such as ASREML and WOMBAT. I do not wish to imply here that design is irrelevant – any estimation procedure works better with well-designed data sets.

AN ADEQUATE KNOWLEDGE OF THE BIOLOGY OF THE SPECIES

It is a common practise in statistics to avoid drawing conclusions about within population relationships from between population comparisons. To do so even further isolates quantitative geneticists from the functional relationships inherent in the underlying biology. The components

of wool weight provide a good example of such functional relationships. For example, the well described inverse relationship between density and fibre diameter ($N \propto D^{-2}$) reveals an underlying functional relationship, related to the developmental events leading to the initiation and formation of the follicle. Similarly, the inverse relationship between length growth rate and diameter, such that the ratio L/D^2 is approximately constant within the animal, has been much trumpeted by nutritional scientists as genetic (and indeed it has a high heritability). Finally, the supply of nutrients to the follicle, and to the fibre (reflected in the product LD^2) is a consequence of a complex set of processes that reflect feed intake, feed efficiency and partition of nutrients and energy between various components of body growth. The fact that the quantities above are ratios should deter no one; I have already argued that all of these variables should be log transformed. Whether these components prove useful as selection criteria is a moot point, but they may help to improve our understanding of physiological changes underlying gains. Similar functional relationships exist, I am sure, for other complex traits in diverse species.

At a deeper level, many animal scientists have searched, often in vain, for reliable physiological indicators of production traits (see Blair *et al*. 1990). However, one appears very promising as an indicator of growth rate and feed efficiency. The IGFs, and in particular IGF1, were first suggested by Salmon and Daughaday (1957) as mediators of growth. IGF1 is expressed throughout the organism, and its receptors are ubiquitous. It has now been shown to correlate well with body size in a variety of species, and with feed efficiency in cattle and pigs.

PREDICTED PROGRESS OVER A FEW GENERATIONS

It is well known from single trait selection experiments that, in populations of reasonable size, response is undiminished and approximately linear over tens or even hundreds of generations. Loss of genetic variation is usually not a problem. However, we have few data on the stability of genetic correlations over long periods; this is potentially important since most practical breeding programs depend on multivariate selection indices. This is, perhaps, another reason for choosing criteria that are, as much as is possible, functionally unrelated.

It is as important as ever to monitor realised progress. A departure from that predicted indicates something is seriously amiss, taking into account, of course, the fact that there may be considerable variation between replicate lines.

A DIGRESSION ON SCALING

Despite the fact that most statistical packages available to the breeder offer the option to transform the observed data, I have seen little evidence that such transformations are common. The topic is rarely discussed; the one exception is a note by James (2007). Since many important commercial traits involve growth processes, logarithmic transforms are often appropriate. Such transformations have the additional benefit of removing the irksome problem of ratios in objectives and selection criteria. We accept the need to transform data such as faecal egg counts, but not fibre diameter, and use measurements such as the coefficient of variation, which is an explicit admission that a logarithmic scale is appropriate.

A suitable scale of measurement is not merely desirable to ensure the independence of variances upon the mean, as we know linear analysis is quite robust to such deviations. The importance, I argue, is to ensure an appropriate genetic scale. Failure to choose one can lead to apparent dominance and non-linearity of response, if measured inappropriately, and may be particularly important in interpreting genetic crosses. QTL analyses, in particular, often involve crosses between divergent species, and it seems prudent to remove apparent effects that are merely a consequence of choosing an inappropriate scale.

ON THE NATURE OF QUANTITATIVE VARIATION

At the core of contemporary animal breeding theory is the nature of quantitative inheritance. In general, most traits of interest to the breeder appear approximately additive. For heritability at least, this assumption is well justified by experiments in drosophila, mice, corn and other species; the behaviour of genetic correlations is less clear. Furthermore, genetic variation is not limiting; it seems that new variation is generated at such a rate (by mutation or recombination) that progress can continue almost indefinitely. These observations provide a strong empirical foundation for applying the theory to livestock improvement. However, if we wish to understand the genetic basis of this variation, we enter difficult territory – one that requires that we reconcile the above observations with our increasing awareness of the complexity of gene action.

Through much of the 20th century, deep divisions beset population genetics over the origin and maintenance of the standing variation in gene frequencies. One, the "classical" view, maintains that most variation is either deleterious and maintained by recurrent mutation, or neutral and maintained by drift. Hence, phenomena such as inbreeding, and its converse, heterosis, are due the covering or uncovering of these harmful mutations. The alternative "balanced" view is more mystical, full of terms such as co-adapted gene complexes, assumes widespread epistasis and that variation can be maintained by over-dominance. Undoubtedly, the truth lies somewhere in between, but I tend to sit in the latter camp; one of the so-called "naïve pan-selectionists". The classical view still dominates population genetics, and adherents tend to treat genes as entities, not imbedded in a complex interacting system. Concepts of the "the Selfish Gene" and terms such as junk DNA have arisen from such thinking. I suggest that those who see value in identifying genes of large effect that can improve selection response for commercial traits also reflect the "bean bag genetics" approach, as Mayr disparagingly called single gene selection models.

I am much encouraged by the revelations of molecular genetics, which support the notion that even simple phenotypes are a consequence of a complex interactions at all levels of gene expression. The mammalian genome has roughly 20,000 genes, not much more than that found in far less complex organisms. Complexity in development arises, I believe, through greater interaction between genes and between their gene products.

The observable properties of quantitative variation, together with the very high rates of generation of new variation each generation, despite the underlying interactions, are then a paradox that can only be resolved, I believe, by concluding that many genes can affect each trait, and that each of these genes contribute, on average, a small affect.

BIOTECHNOLOGY AND THE FUTURE

In recent years, and especially in the three meetings spanning 2001-5, attention has focused on biotechnology, and particularly on marker assisted selection (MAS). For the reasons alluded to above, I have always been sceptical of the promises of MAS, as I was earlier of claims by some genetic engineers that large changes in productivity could be achieved benignly. I do not imply that genes of large effect cannot be found, or that identifying such genes is unimportant. In populations undergoing intense selection, rare alleles or new mutations that have a large effect on the trait will increase in frequency, despite the fact that they may be deleterious in other respects. Also, false positives are common in QTL screens with small data sets leading the researchers (and their funding bodies) to the conclusion that a significant fraction of the variation for a trait can be attributed to a few genes. The current state of MAS in cattle was recently summarised by Van Tassell *et al*. (2007), who concluded that the application of QTL has been, as yet, limited. What disturbs me is not the search for QTL, nor for the genes that affect livestock in useful or adverse ways. Rather, it is the promise to funding agencies, and the diversion of research funds from other important research areas that is problematic.

DNA technology's first important application was to identify parentage, initially using microsatellites, now being replaced by SNPs. Accurate pedigrees are an important part of modern animal breeding technology, but DNA sampling is an enabling technology for other genetic tests, such as testing for carriers of deleterious genes. Pedigrees define genetic relationships, but with a wide SNP coverage of the genome, it is possible to estimate these relationships directly. Since these would be realised rather than expected relationships, they will be more accurate.

Just as the progression of Moore's Law had, perhaps, the largest impact on animal breeding technology over the last 30 years, DNA technology is progressing at a similar rate. For example, the human genome project cost millions of man-hours and billions of dollars to complete; the task has recently been repeated for less than $50,000. Complete DNA sequencing is now within the reach of all species. We can analyse the whole transcriptosome, or identify transcripts in a selected chromosomal region, and profile expression patterns for each. SNP chips containing 100,000 or more polymorphic bases are now available for many species.

While advances are made in annotating the known genes (identifying their protein product and possible function), we are still far from interpreting the complex interactions and predicting the phenotypic consequences of gene substitutions, alone or in combination. We are still ignorant of the developmental genetics of body size, or even of organelles such as the wool follicle or the mammary gland, but the rate of gain in our knowledge is truly staggering. Comparative genomics offers us a chance to ask, for example, why some species are susceptible to parasites, but closely related ones are not. The challenge to the modern quantitative geneticist is to capture this new knowledge to identify pathways and genes that can be used to improve productivity.

REFERENCES

Anderson, J.R. (1988) *Proc. Aust. Assoc. Anim. Breed. Genet.* **7**:32
Banks, R. (1990) *Proc. Aust. Assoc. Anim. Breed. Genet.* **8**:237
Barlow, R. (1987) *Proc. Aust. Assoc. Anim. Breed. Genet.* **6**:162
Barker, J.S.F. (1979) *Proc. Aust. Assoc. Anim. Breed. Genet.* **1**:2
Blair, H.T., McCutcheon, S.N. and Mackenzie, D.D.S. (1990) *Proc. Aust. Assoc. Anim. Breed. Genet.* **8**:133
Brien, F.D. (1990) *Proc. Aust. Assoc. Anim. Breed. Genet.* **8**:241
Cartwright, T.C. (1982) *Proc. Aust. Assoc. Anim. Breed. Genet.* **3**:5
Cunningham, E.P. (1979) *Proc. Aust. Assoc. Anim. Breed. Genet.* **1**:8
Cunningham, E.P. (1979) *Proc. Aust. Assoc. Anim. Breed. Genet.* **1**:18
Hill, W.G. (1981) *Proc. Aust. Assoc. Anim. Breed. Genet.* **2**:3
James, J.W. (1979) *Proc. Aust. Assoc. Anim. Breed. Genet.* **1**:47
James, J.W. (2007) *Proc. Assoc. Advmt. Anim. Breed. Genet.* **17**:150
Jones, L.P. (1991) *Proc. Aust. Assoc. Anim. Breed. Genet.* **9**:9
Lewontin, R.C. (1964) *Proc. XI Int. Congr. Genetics* Oxford, Pergamon Press *p571*
Nicol, D.C., Graser, H.-U., Tier, B. and Hammond, K. (1985) *Proc. Assoc. Advmt. Anim. Breed. Genet.* **5**:151
Ponzoni, R.W. (1979) *Proc. Aust. Assoc. Anim. Breed. Genet.* **1**:320
Ponzoni, R.W. (1988) *Proc. Aust. Assoc. Anim. Breed. Genet.* **7**:55
Roberts, E.M. (1979) *Proc. Aust. Assoc. Anim. Breed. Genet.* **1**:48
Salmon, W.D and Daughaday, W.H. (1957) *J. Lab. Clin. Med.* **49**: 825.
Smith, C., James, J.W. aand Brascamp, E.W. (1986) *Anim. Prod*. 43: 545
Smith, C. (1988) *Proc. Aust. Assoc. Anim. Breed. Genet.* **7**:42
Van Tassell, C.P. Sostegard, T.S. Liu, G. and Matukmalli L.K. (2007) *Proc. Assoc. Advmt. Anim. Breed. Genet.* **17**:461.

ASSOCIATION OF MICROSATELLITE MARKERS AND NRAMP1GENE WITH BOVINE TUBERCULOSIS TRAITS IN ZEBU CATTLE

A. Ali [1,2], H.N. Kadarmideen[1], P. C. Thomson [2], C. Flury[3], B. Müller[4] and J. Zinsstag[4]

[1] CSIRO Livestock Industries, JM Rendel Laboratory, Rockhampton, QLD 4701, Australia.
[2] Faculty of Veterinary Science, University of Sydney, Australia.
[3] Berner Fachhochschule, Scheweizerische Hochschule Für Landwirtschaft, Länggasse 85, Zollikofen , CH-3052, Switzerland. [4] Swiss Tropical Institute, Basel, CH 4002, Switzerland.

SUMMARY

The main objective of this study was to detect association between microsatellite genetic markers and a candidate gene with tuberculosis-related traits in African *zebu* cattle. A total of 249 of Chadian *zebu* cattle was genotyped for 23 microsatellites and for a known candidate gene *NRAMP1* (natural resistance associated macrophage protein 1). These animals were measured for two tuberculosis-related traits, namely, single intra-dermal comparative cervical tuberculin (SICCT) test on live animals and lung lesion (LL) from the same animals at slaughter. A generalised linear mixed model (GLMM) treating both traits as binomially distributed was fitted using *probit* link function. Eleven out of 21 microsatellite markers tested were significantly associated with presence of LL (P-value < 0.001 to P-value <0.01). For SICCT trait, only BM2113 marker was significant (P-value = 0.012). NRAMP 1 gene (chr 1) was significantly associated with LL at P-value = 0.006, but not associated with SICCT test (P-value = 0.488). Reasons for this disagreement are considered. These results show that these genetic markers and *NRAMP1* gene could be potentially used in marker-assisted selection (MAS) strategies in breeding programs to control the spread of *Mycobaterium. Bovis,* which is a causative agent of bovine tuberculosis.

INTRODUCTION

Mycobacterium bovis is a member of the group of *Mycobacterium* genus classified as the *Mycobacterium tuberculosis complex* and the cause of bovine tuberculosis (bTB). *Mycobacterium bovis* (*M. bovis*) has a broad range of hosts including livestock, wildlife and humans. The bTB is a disease of socio-economic and public health importance and constraining international trade of animals and their products. In Africa, the disease is present virtually on the whole continent (Ayele *et al.* 2004), and due to lack of financial resource, very few countries are able to apply a standard control measurement. In previous studies from Chad and Cameroon, significant differences in bTB infection prevalence were observed between Arab and Mbororo breeds in African zebu cattle (Diguimbaye-Djaibe *et al*. 2006). Resistance and susceptibility to infectious diseases is often influenced by the genetics of the host. Based on this consideration, the present study investigates possible association between 23 microsatellite markers and a known candidate gene *NRAMP1* with bTB infection of *zebu* cattle in Chad. This study is first of its kind to investigate association between microsatellite markers and known a candidate gene *NRAMP1* with bTB in African cattle.

MATERIAL AND METHODS

Animals. A total 249 animals of Arab (n=162) and Mbororo (n=87) breeds from African zebu cattle were measured for several phenotypes of bTB-related traits between July and November 2005 at abattoirs in Southern Chad, which included SICCT tests and lung lesions. Animals were raised in a long distance transhumant livestock production system with frequent trans-border movement of herds between the Central African Republic and Chad. These animals are considered a representative sample from a large number of different herds and big area in southern Chad.

Determination of bTB infection. SICCT testing results were available for all animals investigated in this study. To be able to perform SICCT on animals to be slaughtered, an arrangement was made with the slaughterhouse management to maintain animals three days prior to slaughter in the animal confinement area. SICCT testing and reading was carried out according to standard protocols. After slaughter, post-mortem examination was carried out by inspecting lungs and other organs for the presence of TB-like lesion. If a TB like lesion was detected, a specimen was collected for bacteriological culture. *Mycobacterium bovis* was confirmed by detection of acid fast bacilli (AFB) using Ziehl Neelsen method.

Genotype data. Blood samples were collected and genomic DNA was extracted using the QIAamp® DNA Blood Kit (QIAGEN, Cat. No. 51106).A total of 249 animals were available for genotyping. Twenty three markers:BM1818, BM1824, BM2113, CSRM60, CSSM66, ETH10, ETH225, ETH3, ETH152, ETH185, TGLA122, TGLA126, TGLA227, TGLA53, ILSTS005, ILSTS006, HEL5, HAUT27, SPS115, INRA32, INRA35, INRA23, MM12 and a known candidate *NRAMP1* gene was investigated in this study. Markers were selected from genetic diversity markers of African zebu cattle recommended by Domestic Animal Diversity (DAD) Information System of the Food and Agriculture Organisation (http://dad.fao.org/). All of the loci under investigation were polymorphic. The ranges of genotypes and alleles were 10 to 51 and 4 to 16 respectively. Markers with missing fractions of more than 20% were excluded from the analysis (see table 1).

Table 1. Number of genotypes and percent of missing genotypes for the Arab and Mbororo breeds of African zebu cattle.

Arab breed (n=162)			Mbororo breed (n=87)		
Marker	# genotypes	% missing	Marker	# genotypes	% missing
BM1818	35	1.5	BM1818	29	6.8
BM1824	14	0.0	BM1824	14	0.0
BM2113	32	0.0	BM2113	31	0.0
CSRM60	30	0.0	CSRM60	30	0.0
CSSM66	42	0.0	CSSM66	42	0.0
ETH10	30	1.5	ETH10	30	0.0
ETH225	23	1.5	ETH225	23	0.0
ETH3	20	0.0	ETH3	20	0.0
ETH152	10	0.0	ETH152	9	1.4
ETH185	45	1.4	ETH185	41	0.0
HAUT27	19	13.7	HAUT27	17	10.8
HE5	22	26.7	HE5	22	29.7
ILSTS006	30	3.8	ILSTS006	30	5.4
ILSTS005	18	7.6	ILSTS005	15	4.1
INRA23	32	0.8	INRA23	32	0.0
INRA32	32	3.8	INRA32	32	2.7
INRA35	21	2.3	INRA34	21	4.1
SPS115	14	0.0	SPS115	16	0.0
TGLA122	37	3.8	TGLA122	37	2.7
TGLA126	27	0.0	TGLA126	28	0.0
TGLA227	31	0.0	TGLA227	31	0.0
TGLA53	50	29.0	TGLA53	51	19.9
MM12	51	0.0	MM12	51	0.0
NRAMP1	10	0.0	NRAMP1	10	0.0

Statistical analysis. Association analyses were restricted to SICCT and LL. These response variables were recorded as discrete data (1 recorded as positive and 0 as negative). A generalized linear mixed models (GLMM) treating both traits as binomially distributed was fitted using the probi*t* link function and dispersion parameter was fixed at 1 in GenStat 11.1 package, terms included in the model were sex, age and breed and genotypic effects of each marker and NRAMP1 gene as fixed effects. The marker-trait association analysis included one marker at a time, hence there were 22 analyses. A marker effect was considered significance if a P-value ≤ 0.01.

Results and discussion. Of the 249 animals tested, the proportions of visible lesion detected and positive SICCT reaction observed were 6.8% and 28.0% respectively, more Mbororo breed was effected by lung lesion and positive SICCT reaction than Arab breed (data not shown). Microsatellite markers, chromosomal location and P-value for association with lung lesion and SICCT traits are shown in Table 2. Overall, eleven out of 21 microsatellite markers tested were significantly associated with presence of lung lesion. For SICCT, only BM2113 marker was significance associated with positive skin tuberculin test (P-value =0.012). For lung lesion trait, BM1824 (chr 1), ETH10 (chr 5), ETH185 (chr 17) and INRA35 (chr 16) markers show high significance association (table 2).

Table 2. Association of microsatellite markers with lung lesion and tuberculin tests (SICCT) from cattle screened for bovine tuberculosis (P-values indicate significance of marker effects).

Lung lesion trait (n=249)			SICCT trait (n=246)		
Marker	Chr.	P-value	Marker	Chr.	P-value
BM1818	23	0.005	BM1818	23	0.085
BM1824	1	<0.001	BM1824	1	0.182
BM2113	2	0.001	BM2113	2	0.012
ETH3	19	0.002	ETH3	19	0.707
ETH10	5	<0.001	ETH10	5	0.308
ETH152	5	0.002	ETH152	5	0.357
ETH185	17	<0.001	ETH185	17	0.343
ILST005	10	0.008	ILST005	10	0.133
ILST006	7	0.003	ILST006	7	0.063
INRA35	16	<0.001	INRA35	16	0.102
TGLA126	20	0.003	TGLA126	20	0.147
NRAMP1	1	0.006	NRAMP1	1	0.488

SICCT = single intra-dermal comparative cervical tuberculin test Chr = chromosomal location of markers

At a genotypic level of within each marker, we observed some genotypes whose effects were either positive or negative on SICCT and LL (data not shown here); these genotypes are of interest to geneticists to implement marker assisted selection (MAS) of animals based on their genotypes. *NRAMP 1* gene (chr 1) was significantly associated with LL (P=0.006), but not associated with SICCT test (P=0.488). This may be due to some false negative results involved in the SICCT test as it can only detects an early stage of bTB infection. A considerable number of animals that were sampled had an advanced stage of bTB as previously described (Ngandolo *et al*. 2009), consequently, some SICCT tests gave false negative results., It is important to know that SICCT detects only a cell-mediated immune (CMI) response which predominates in the early stage of the disease, but this response declines and is eventually superseded by antibody-mediated response as

the disease progresses. Hence, incidence of positive SICCT tests may have been underestimated and subsequently also the marker association effects. Genetic variation in cattle hosts is an important determinant of the manifestation of infection with *Mycobacterium spp*. In previous study, an allelic variant of several genes have been implicated in the genetic susceptibility to tuberculosis infection both in human and mouse model, but little is known in cattle. These results show that these genetic markers and *NRAMP1* gene could be potentially used in MAS strategies in breeding programs to control the spread of *M. bovis* infection by increasing the frequency of desirable genes in the herds. A whole genome scan study using high density single nucleotide polymorphisms (SNPs) on a large number of genetically structured populations (i.e. backcross, F2, half-sibs and full-sibs of large families) and well recorded animals would be an ideal situation, but this may not be realistic in Africa. Although, microsatellite markers used here were for genetic diversity study, we included them in the association analyses with an expectation that some of the markers might turn out to be significant. Results showed that this is true and could be used in MAS or gene assisted selection (GAS); these MAS / GAS strategies would be more economical to apply in less developed countries where there are no comprehensive phenotypic and/or pedigree recording as in developed countries. A statistical model that fits allele substitution effects instead of genotypic effects at markers will reduce the number of parameters to be estimated. Such a model might allow fitting all 24 markers at once in multiple-marker regression. However, the peculiarity of this dataset was that traits were non-normal or binary and sample sizes were too small; both pose statistical challenges in terms of fitting threshold-liability models for data with small sample sizes (Kadarmideen *et al*. 2004). These approaches or problems are currently being investigated.

CONCLUSIONS

This study provides evidence for genetic variation between African cattle at DNA marker level for susceptibility to tuberculosis infection. Although microsatellite markers used here were for study of genetic diversity, we found that some markers and *NRAMP1* gene are significantly associated with bTB related traits of SICCT and LL. Those markers and *NRAMP1* could be potentially used to implement MAS/GAS strategies for breeding program to control the spread of *M. bovis* by increasing the frequency of a favourable genotype(s) in the herds.

AKNOWLEDGMENTS

The authors would like to thank Swiss National Science Foundation (SNSF) for providing financial support for the project leaders H. Kadarmideen and J. Zinsstag. Laboratoire de Recherches Veterinaires et Zootechniques de Farcha, Chad for their contribution. A. Ali was funded by Australian CSIRO Livestock Industries post-graduate fellowship.

REFERENCES

Ayele, W.Y., Neill, S.D., Zinsstag, J., Weiss, M.G., Pavlik, I. (2004) *Int J Tuberc Lung Dis*. **8**:924.

Diguimbaye-Djaibe, C., Hilty, M., Ngandolo, R., Mahamat, H.H., Pfyffer, G.E., Baggi, F., Hewinson, G., Tanner, M., Zinsstag, J., Schelling, E. (2006) *Emerg Infect Dis*. **12**:769.

Kadarmideen, H.N., Schworer, D., Ilahi, H., Malek, M., Hofer, A. (2004) *J Anim Sci*. **82**:3118.

Ngandolo, B.N., Muller, B., Diguimbaye-Djaibe, C., Schiller, I., Marg-Haufe, B., Cagiola, M., Jolley, M., Surujballi, O., Akakpo, A.J., Oesch, B., Zinsstag, J. (2009) *Prev Vet Med*. **89**:81.

Domestic Animal Diversity of the Food and Agriculture Organisation (http://dad.fao.org/).

Proc. Assoc. Advmt. Anim. Breed. Genet. 18:472-475

THE ROLE OF ANIMAL GENETIC IMPROVEMENT IN REDUCING GREENHOUSE GAS EMISSIONS FROM BEEF CATTLE

P. F. Arthur[1], K. A. Donoghue[2], R. M. Herd[3] and R. S. Hegarty[3]

[1]Elizabeth Macarthur Agricultural Institute, NSW Department of Primary Industries, Camden, NSW 2570
[2]Agricultural Research Centre, NSW Department of Primary Industries, Trangie, NSW 2823
[3]Beef Industry Centre, NSW Department of Primary Industries, Armidale, NSW 2351

SUMMARY

In Australia emissions from the livestock industries represent 10.9% of the net national greenhouse gas (GHG) emissions in 2006, and most of these were from sheep and cattle. With the government signalling its commitment to reduce emissions, industries need to develop emissions reduction strategies. This paper identifies some of the current genetic improvement practices in beef cattle that reduce GHG emissions and also identifies new areas for further research with potential for GHG reductions. Current GHG emission reduction strategies in beef cattle are reliant on improving productivity of cattle in order to reduce emissions per unit of product. Hence emissions reduction at the national level is largely reliant on there being a cap or reduction in animal numbers. In the long term it is important that strategies that directly reduce GHG emissions per unit of feed intake be developed.

INTRODUCTION

The agricultural sector is a source of greenhouse gas (GHG) emissions worldwide, with the magnitude of its contribution differing from country to country. A recent FAO report estimates that globally livestock are responsible for 18 percent of greenhouse gas emissions (Steinfeld et al. 2006). In Australia, the emission from the livestock industries is estimated at 61.0 Mt CO_2 –e, which represented 11.3% of the net national GHG emissions in 2007 (DCC 2009). Over 90% of the livestock emissions are from ruminants, predominantly sheep and cattle.

In December 2008, the Commonwealth Government of Australia released its White Paper on "Carbon Pollution Reduction Scheme: Australia's Low Pollution Future"**,** and it signalled the government's commitment to reduce greenhouse pollution in Australia in the short and long term. With this comes the need for all industries to examine and develop strategies to reduce their contribution to GHG emissions. The objective of this paper was to identify some of the current genetic improvement practices in beef cattle that reduce GHG emissions and to identify for research new areas with potential for GHG reductions.

GHG EMISSIONS FROM CATTLE

Greenhouse gases generated by cattle production include methane (CH_4) and nitrous oxide (N_2O), which have global warming potentials 21 and 310 times that of CO_2 respectively, making them very potent GHG. Methane primarily arises from enteric fermentation but also small amounts derive from manure stores (Fig 1). Manure emissions are least from extensive grazing enterprises and greatest from manure stockpiles and slurries reflecting the need for a moist anaerobic environment for methanogenesis.

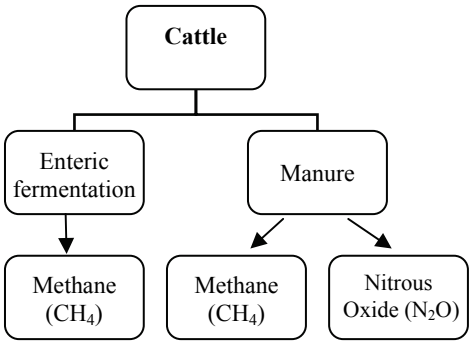


Figure 1. Greenhouse gas emissions from cattle.

BEEF BREEDING PRACTICES THAT REDUCE GREENHOUSE GAS EMISSIONS

A strong positive relationship between feed intake and methane production in ruminants is recognized in most algorithms predicting methane production rate (Blaxter and Clapperton 1965, Pelchen and Peters 1998). On grazing or forage diets, the higher the feed intake the higher the daily methane output by the ruminant, on the same feedstuff. However a strategy of reducing daily feed intake to achieve lower levels of methane production has received little attention because of concerns over reduction in productivity of the ruminant. In beef production such a strategy would mean that slaughter cattle will take longer to reach market weight, young replacement heifers will take longer to reach puberty and weaning rates in cows will reduce. Therefore to achieve the same level of productivity, this will mean that all classes of cattle (steers, heifers, cows etc) have to be kept longer resulting in potentially higher total feed intake. A life cycle analysis approach to consider the broader impacts of such mitigation strategies is important, rather than look only at mitigation effects on individual animals.. Any strategy that reduces feed intake per unit of product would also result in reduction of GHG emission per unit of product, which can be used as an indicator of improved emissions intensity of the production system. Some of the GHG reduction strategies for which genetic improvement can play a major role are discussed below.

Reducing age at slaughter. The endpoint for slaughter cattle is determined by the specification of the target market, usually for weight and fatness. Achieving these slaughter specifications at a younger age will result in a lower total GHG emission per unit of product relative to a higher age at slaughter. Impacts can be substantial, with feedlot finishing of cattle in northern Australia for 2-5 months calculated to reduce lifetime methane production of slaughter cattle by 34–54%, largely through reduced time to slaughter (McCrabb *et al.* 1998).. Similarly, Howden and Reyenga (1999) showed that methane emission per unit liveweight gain reduces as average daily gain (ADG) increases. This is primarily due to the fact that in beef cattle over 50% of feed intake is used for maintenance, hence the faster an animal grows the lower the total feed requirements for maintenance over the growing period, leading to lower methane emission per liveweight gain. In recognition of this, GHG quantification protocols (Beef feeding – reducing days on feed, and, Beef lifecycle) have been developed and approved for the carbon offset trading in Alberta, Canada (Alberta Environment 2009).

Improved ADG can be achieved through improvement in nutrition and the environment, but can also be achieved through genetic improvement either by crossbreeding or selection for growth traits.

Efficiency of feed utilisation. By definition beef cattle that are efficient in feed utilisation will eat less per unit of product. In Australia, residual feed intake is the feed efficiency trait used for genetic improvement of feed efficiency. It has the unique characteristic that low RFI cattle consume less feed than high RFI cattle for the same level of productivity (Arthur *et al.* 2001). Theoretical calculations based on the reduction in feed intake showed that low RFI cattle have 15% - 21% reduction in methane emissions, 15% reduction in methane from manure and 17% reduction in nitrous oxide from manure, relative to high RFI cattle (Okine *et al.* 2001; Herd *et al.* 2002). These results were confirmed by empirical evidence from two studies where actual methane emissions were measured. The results of the two studies indicate that there is a 15% - 30% reduction in methane emissions and 15% - 20% reduction in manure production from low RFI relative high RFI cattle (Nkrumah *et al.* 2006; Hegarty *et al.* 2007). In recognition of this, selection for low RFI is being considered for potential GHG protocol development for the carbon offset trading in Alberta, Canada (Alberta Environment 2009).

Use of adapted, high producing cattle. The main purpose of the breeding herd, in beef cattle, is to produce progeny. Hence a dry cow consumes feed (hence produces GHG) but with no product to show for. A herd with a higher calving rate is therefore desirable, not only in terms of the higher number of weaned calves to be sold, but also in terms of lower feed intake per unit of product, leading to lower GHG emissions per unit of product. It has been shown that female fertility is improved by the use of crossbred females for breeding (Arthur *et al.* 1999). Selection for reduced days to calving in beef cattle also improves reproductive performance in females.

Productivity of most cattle breeds in the tropics is low due to stressors (heat, poor nutrition, disease) imposed on the animal by the harsh environment. Hence most tropical breeds have low productivity. The use of tropically-adapted composites in northern Australia is one of the practical genetic improvement strategies used successfully to improve productivity. Replacement of shorthorn cattle with composite-breed cattle by the North Australia Pastoral Company's "Alexandria" station, was associated with reduction in the methane/kg liveweight weaned from 1.25 to 0.86 t CH_4/t LW weaned. This advantage came largely from increased weaning rate (55 to 80 calves/100cows) but increased slaughter weight may also be expected (Bentley et al. 2008). When planned properly, such tropically-adapted composites offer less GHG per unit of product relative to purebreds due to higher female fertility, progeny achieving lower age at slaughter and higher feed efficiency.

Uniqueness of genetic improvement. It is apparent from the outlined GHG reduction strategies that genetic improvement programs in general will result in a reduction of GHG emissions per unit of product, but not always reduce total emissions. Genetic improvement has been widely adopted in the beef industry to improve production efficiency, with the resulting improvement in enterprise profitability a continuing driver for such change. The fact that some of these breeding decisions also help reduce GHG emissions are unexpected secondary benefits. Further, any change in an animal's characteristics achieved through genetic improvement can be passed on to the next generation, so any reduction in GHG which is associated with such genetic improvement is also permanent. This is in contrast to other strategies to reduce GHG emissions where a particular treatment, such as grain (instead of pasture) feeding, and feeding of edible oils, need to be applied on a regular basis. For some extensively managed cattle, application of such treatment on a regular basis is not practical.

CHALLENGES AND OPPORTUNITIES FOR THE FUTURE

Comprehensive assessment of the contribution of genetic improvement strategies to real or potential reductions in GHG emissions in the Australian context need to be carried out. One such assessment was reported by Alford *et al.* (2006) on potential GHG reductions from selection for low RFI in Australia. Using a modest base scenario of 0.76% rate of genetic improvement and 30% maximum adoption, Alford *et al.* (2006) reported that the cumulative reduction in national emissions was 568,100 t of methane over 25 years, with annual emissions in year 25 being 3.1% lower than in year 1. Any increase in the rate of genetic improvement and/or the maximum adoption level increases the cumulative reduction in methane emission.

The strategies outlined above are important first steps. However, most of the GHG emission reduction strategies are reliant on improving productivity of cattle in order to reduce emissions per unit of product. Achieving greater productivity frequently results in increased feed intake hence greater GHG emissions per head of cattle. Hence GHG emissions reduction for beef cattle at the national level can only be achieved only if there is a cap or reduction in animal numbers. A number of other industries are also faced with this dilemma. The automobile industry, for example, is reducing emissions per vehicle, however any reductions in the national emissions can only be achieved if the number of cars on the road is capped or reduced. Therefore, it is important that in

the long term, strategies to directly reduce the GHG emissions per unit of feed intake need to be developed.

One of the long term strategies which need investigating is the genetic improvement in methane emission per unit of feed intake. A review by Herd and Hegarty (2007) gives an indication that there is genetic variation in many of the biological processes inside the animal that contribute to the level of methane production. The challenge now is to develop accurate and repeatable methods of measuring individual animal intake and methane production, which can be applied to large numbers of cattle.

REFERENCES

Alberta Environment (2009) www.carboboffsetsolutions.ca (Downloaded on 2009-03-26).

Alford, A. R., Hegarty, R. S., Parnell, P. F., Cacho, O. J., Herd, R. M. and Griffith, G. R. (2006) *Aust. J. Exp. Agric.* **46**:813.

Arthur, P.F., Archer, J.A., Johnston, D.J., Herd, R.M., Richardson E.C. and Parnell, P.F. (2001) *J. Anim. Sci.* **79**:2805.

Arthur, P. F., Hearnshaw, H. and Stephenson P. D. (1999) *Lives. Prod. Sc.* **57**:231.

Bentley, D., Hegarty, R.S. and Alford, A.R. (2008) *Aust. J. Exp. Agric*. **48:**60.

Blaxter, K. L. and Clapperton, J. L. (1965) *Br. J. Nutr.* **19**:511.

DCC (2009) National Greenhouse Gas Inventory (2009) Australia's National Greenhouse Gas Accounts, Aust. Govt. Dept. Climate Change. http://www.climatechange.gov.au/inventory/2007/index.html (Downloaded on 2009-07-09).

Hegarty, R. S., Goopy, J. P., Herd, R. M. and McCorkell, B. (2007) *J. Anim. Sci*. **85**:1479.

Herd, R. M., Arthur, P. F., Hegarty, R. S. and Archer, J. A. (2002) *7th World. Congr. Genet. Appl. Livest. Prod.* **31**:281.

Herd, R. M. and Hegarty, R. S. (2007) UK Department of Environment, Food and Rural Affairs sponsored workshop on "Genetics of livestock nitrogen and carbon emissions", Edinburgh, UK, October 2007. http://randd.defra.gov.uk/Document.aspx?Document=AC0204_7642_FRA.pdf (Downloaded on 2009-03-27).

Howden, S. M. and Reyenga, P. J. (1999) *Aust. J. Agric. Res*. 50:1285.

McCrabb, G.J., Kurihara, M. and Hunter, R.A. (1998) *Proc. Nutr. Soc. Aust*. **22**:55.

Nkrumah, J. D., Okine, E. K., Mathison, G. W. Schmid, K., Li, C., Basarb, J. A., Price, M. A., Wang, Z. and Moore, S. S. (2006) *J. Anim. Sci*. **85**: 145.

Okine, E. K., Basarab, J., Baron, V. and Price, M. A. (2001) In 'Abstracts of Presentations and Posters', Agric. Inst. Canada (AIC), 2001 Conference, University of Guelph, 15: CSAS01-21.

Pelchen, A. and Peters, K. J. (1998) *Small Rum. Res*. **27**:137.

Steinfeld, H., Gerber, P., Wassenaar, T., Castel, V., Rosales, M. and de Haan, C. (2006) Livestock's long shadow – Environmental issues and options. FAO Rome, 408p.

BENEFITS OF GENETIC SUPERIORITY IN RESIDUAL FEED INTAKE IN A LARGE COMMERCIAL FEEDLOT

R.M. Herd[1], S. Piper[2], J.M. Thompson[2], P.F. Arthur[3], B. McCorkell[4] and K.C.P. Dibley[5]

Cooperative Research Centre for Beef Genetic Technologies
[1] NSW Department of Primary Industries, Beef Industry Centre, Armidale, NSW, 2351
[2] Division of Animal Science, University of New England, Armidale, NSW, 2351
[3] NSW Department of Primary Industries, Elizabeth Macarthur Agricultural Institute, Camden, NSW, 2570
[4] NSW Department of Primary Industries, Tamworth Agricultural Institute, Tamworth, NSW, 2340
[5] NSW Department of Primary Industries, Agricultural Research Institute, Trangie, NSW, 2823

SUMMARY

Growth, feed intake, feed efficiency and carcass traits on 208 yearling Angus steers differing in genetic merit for residual feed intake (RFI) were measured in a large commercial feedlot. At feedlot entry the steers were drafted into three groups: high efficiency (HE; midparent RFI-EBV ≤-0.3kg/day, medium efficiency (ME: midparent RFI-EBV >-0.3 to 0.14 kg/day), and low efficiency (LE; midparent RFI-EBV ≥0.16 kg/day), and feedlot performance evaluated over 251 days. Whilst individual animal weight and carcass data were collected, daily feed consumption was recorded on a group basis. The HE steers grew as fast or faster than either the ME and LE steers (1.11 ±0.02 (sd) *v* 1.06 ±0.01 and 1.07 ±0.02kg/day). At slaughter the HE and ME groups were heavier than the LE group (713 ±5 and 714 ±5 *v* 701 ±5 kg). The HE steers consumed less feed than the ME and LE steers over the 251 days (10.4 *v* 11.8 and 11.1kg/day). Compared to the LE steers, the HE steers had a 10% lower FCR and 0.98kg/day lower RFI. Carcass weight and dressing percentage were lowest in the LE steers, and eye muscle area was highest in the ME steers. Subcutaneous rib fat depth on the carcass was lower in the HE steers than for the ME and LE steers (15.6 ±0.6 *v* 17.6 ±0.6 and 20.7 ±0.6mm). AUSMEAT and USDA marbling scores were highest in the ME group and not different for the HE and LE groups. The experiment showed that genetic superiority in RFI reduced feed consumed over 251 days of feeding in a large commercial feedlot with no compromise in weight gain, carcass weight, dressing percentage or marbling grade.

INTRODUCTION

Genetic improvement in feed efficiency has the potential to reduce feed costs in a feedlot operation whilst still maintaining production levels. Residual feed intake (RFI) is a measure of feed efficiency and is the difference between actual feed intake by an animal and its expected feed intake based on its requirements to maintain weight and for its growth (Arthur *et al*. 2001). This experiment aimed to validate the benefits from feeding steers bred from parents known to be of high, medium and low genetic merit for RFI in a large commercial feedlot.

MATERIALS AND METHODS

Cattle breeding and RFI groups. The Angus steers were bred at the NSW DPI Agricultural Research Centre, Trangie, NSW. Trial Net (or residual) Feed Intake (NFI) estimated breeding values (RFI-EBV) which are produced by BREEDPLAN were available for their sires and dams.. The sires and dams were sorted on their RFI-EBV and mated in 2005 to produce offspring that were genetically-divergent for feed efficiency. The steer calves were managed together with their dams till weaning and continued to be managed together until feedlot entry. Of the 271 steers weaned, 1 died, 2 were excluded because of poor leg structure, another 6 excluded because of

uncertain parentage, and 42 failed to attain the specified induction weight. These 42 steers contained roughly equal proportions of high, medium and low efficiency candidate animals. The remaining 216 steers were drafted into three groups: being of high efficiency (HE; midparent RFI-EBV ≤-0.3kg/day; N=73), medium efficiency (ME: midparent RFI-EBV >-0.3 to 0.14 kg/day; N=73), and low efficiency (LE; midparent RFI-EBV ≥0.16kg/day; N=70), and sold to the cooperating feedlot. The HE steers were the progeny of 14 sires, the ME steers were the progeny of 14 sires (8 in common with the HE group), and the LE steers were progeny of 9 sires (1 in common with the ME group; no sire had progeny in all 3 groups). Sire RFI-EBV ranged from -0.92 to 1.24kg/day, with a mean accuracy of 67% (range 44 to 87%). Dam RFI-EBV ranged from -1.13 to 0.87kg/day, with a mean accuracy of 63% (range 49 to 77%).

Feedlot management and slaughter. Steers were inducted into the feedlot in October 2007 at an average age of 447 ±17 (sd) days and placed into three separate pens. Splitting of efficiency groups to provide replication was not possible. They received a starter ration for the first 16 days and then 2 intermediate ration formulations for another 16 days, before being placed on the finisher ration. In total they were fed for a 251 days. Feed delivered into each pen was electronically recorded. During the experiment four steers were pulled and treated for health reasons (2 HE; 2 LE), and 4 steers (3 HE: 1 ME) were sold early to meet a market order. These 8 steers were not included in the data analysis. The steers were weighed at induction, after 113 days and after they were killed on day 252. Average daily gain in weight (ADG) for each steer was calculated as change in weight divided by the number of days. After stunning and bleeding the body was weighed and 14kg added for blood loss (the usual weight of blood loss from steers of this weight recorded in the abattoir) to determine the final weight of each steer. Carcasses were split, and weighed separately before being chilled overnight. The left-side was quartered between the 7th and 8th ribs, and rib fat depth and area of the eye muscle, AUSMEAT marble score (1 (nil) to 9 (abundant)) and USDA marble score (100 to 900 by units of 10).

Data analysis The weight of feed delivered into each pen each day was adjusted to being equivalent to a ration with an energy density of 12MJ ME/kg DM, because of the differences in energy density between the three rations initially fed and the finisher ration. Weight of feed was then divided by the number of animals in the pen to calculate feed consumed on a per head basis. The way feeding was managed and recorded in the feedlot was known to increase apparent day-to-day intake per head. Analysing the data as means for consecutive three-day blocks reduced day-to-day variation whilst still allowing underlying trends in feed consumption to be apparent.

Without individual-animal feed intake data it was not possible to calculate individual animal feed conversion ratio (FCR) or RFI. For each pen, FCR was calculated by dividing pen DFI by ADG of the steers in the pen. To calculate RFI for each pen, the equations of SCA (1990) were used to predict DFI by each pen for each period based on the average (mid-period) weight for steers in the pen and ADG over the period. The predicted DFI was deducted from the observed DFI for each pen to calculate RFI. Differences between means for growth and carcass traits for the 3 efficiency groups were analysed in a General Linear Model (GLM) with pen as the only fixed effect. These traits were also regressed against the midparent RFI-EBV, and if found to be significantly different from zero was taken as evidence of a genetic association. Average age at induction of steers in the ME group was 452 ±21 days compared with 445 ±16 and 444 ±13 days for steers in either the HE or LE groups. Age was included as a covariate in the GLM and regression models and means for the three groups are presented as LS-means. Trends in DFI over time and differences between pens were modeled by fitting curves in the form of splines to the data for each pen and analysed using ASREML (Gilmour *et al*. 1999). The final model with a spline curve (with 8 knot points) fitted for each group accounted for 18.3% of the variation in DFI

for the 3 groups. The standard error about the predicted DFI for each group for each day was multiplied by 2 to give an approximate 95% confidence interval about each spline.

RESULTS AND DISCUSSION

The HE steers showed advantages in lower feed intake, no compromise in growth rate, and improved feed conversion over the first 113 days on feed that were sustained over the full 251 days on feed. There was no difference in induction weight between the HE and LE groups, although the ME group were heavier (Table 1). The HE steers grew as fast or faster than either the ME and LE groups over the first 113 days and over 251 days. By day 113 and at slaughter the HE and ME groups were heavier than the LE group. Regression coefficients for induction and final weights, and ADG over 251 days, with midparent RFI-EBV were not significantly different from zero providing evidence for lack of strong associations with genetic variation in RFI. The regression coefficient for weight at, and ADG to 113 days showed a small advantage in daily gain associated with lower midparent RFI-EBV. The regression models for these growth traits against age and midparent RFI-EBV had R^2 ranging from 10%, for weight at induction, to1.3% for ADG over 251 days, showing that even where statistically-significant, genetic variation in RFI only explained a small part of the variation in growth performance.

Table 1. Means (±se) for feedlot performance and carcass traits for Angus steers in high, medium and low efficiency groups, and regression coefficients with midparent RFI-EBV.

	Efficiency group			Regression
	High	Medium	Low	coefficient
Number of animals	68	72	68	
Midparent RFI EBV (kg/day)	-0.52 ± 0.02^{a}	-0.09 ± 0.02^{b}	0.62 ± 0.02^{c}	
Weight at induction (kg)	435 ± 4^{a}	448 ± 3^{b}	432 ± 4^{a}	-1.1 ± 4.2
ADG days 1-113 (kg/day)	1.31 ± 0.02^{a}	$1.28 \pm 0.02^{a,b}$	1.22 ± 0.02^{b}	-0.07 ± 0.03*
Weight at day 113 (kg)	583 ± 4^{a}	593 ± 4^{a}	570 ± 4^{b}	$-9.2 \pm 4.8^{\dagger}$
ADG days 1-251 (kg/day)	1.11 ± 0.02^{a}	1.06 ± 0.01^{b}	$1.07 \pm 0.02^{a,b}$	-0.03 ± 0.02
Final weight (kg)	713 ± 5^{a}	714 ± 5^{a}	701 ± 5^{b}	-8.2 ± 5.6
FI days 1-113 (kg/day)‡	10.7	12.1	11.6	
FI days 1-251 (kg/day) ‡	10.4	11.8	11.1	
FCR days 1-113 (kg/kg)‡	8.1	9.4	9.5	
FCR days 1-251 (kg/kg) ‡	9.4	11.1	10.4	
RFI days 1-113 (kg/day)‡	-0.93	0.40	0.55	
RFI days 1-251 (kg/day) ‡	-0.81	0.63	0.17	
Hot carcass weight (kg)	417 ± 3^{a}	420 ± 3^{a}	406 ± 4^{b}	-8.0 ± 3.7*
Dressing percentage (%)	58.5 ± 0.2^{a}	58.9 ± 0.2^{a}	58.0 ± 0.2^{b}	-0.44 ± 0.20*
Rib fat depth on carcass (mm)	15.6 ± 0.6^{a}	17.6 ± 0.6^{b}	20.7 ± 0.6^{c}	4.7 ± 0.7*
Eye muscle area(cm^2)	76.1 ± 0.4^{a}	78.6 ± 0.4^{b}	76.1 ± 0.4^{a}	-0.26 ± 0.46
AUSMEAT marble score	3.0 ± 0.1^{a}	3.6 ± 0.1^{b}	3.0 ± 0.1^{a}	-0.07 ± 0.13
USDA marble score	477 ± 11^{a}	569 ± 11^{b}	463 ± 12^{a}	-19 ± 15

Means within rows with different superscripts differ significantly ($P<0.05$).
*denotes regression coefficient significantly different from zero at $P<0.05$; †at $P<0.1$.
‡Could not be statistically compared as individual animal data were not available.

Modeling DFI over time showed the patterns of feed consumption were similar for the 3 groups. There was a significant pen effect ($P<0.001$) but no significant interaction ($P>0.05$) between pens and time, that is all the splines had a similar shape over time. Over the first 113 days and over the full 251 days, the pen of HE steers consumed less feed than the ME and LE steers (Table 1). Compared to the LE steers, the HE steers had a 14% lower (better) FCR over the first

113 days and a 10% lower FCR over the full 251 days in the feedlot, so should have been more profitable to feed. The RFI of the HE steers was 1.5kg/day lower (better) than that of the LE steers over the first 113 days, and 0.98kg/day lower over the full 251 day period.

Carcass weight and dressing percentage were lowest in the LE steers, eye muscle area was highest in the ME steers, and the 3 traits had negative regression coefficients with midparent RFI-EBV (Table 1). Subcutaneous rib fat depth on the carcass was lower in the HE steers than for the ME and LE steers and had a positive regression coefficient with midparent RFI-EBV. AUSMEAT and USDA marbling scores were highest in the ME group and not different for the HE and LE groups, and neither marbling trait was associated with midparent RFI-EBV. The regression models for these carcass traits against age and midparent RFI-EBV had R^2 of 19% for rib fat depth, and from 11% to 0.7% for the other traits, showing that, with the exception of rib fat depth, genetic variation in RFI explained only a small part of the variation in these carcass traits.

CONCLUSIONS

This experiment demonstrated that genetic superiority in RFI had a favourable impact on the commercial performance of Angus steers by reducing feed consumed with no adverse effect on final turnoff weight. Each HE steer consumed on average 2.60t of feed compared to 2.87t by the ME and LE steers, that is, saved the feedlot 0.27t or $53 (at $200/tonne) of feed with no compromise in weight gain. The HE steers finished with less subcutaneous fat measured at the 7/8 rib which may have an impact on meeting market specifications. However marbling scores were not influenced by RFI, and dressing percentage was higher in the HE steers, which together would be expected to result in a greater yield of retail beef with no reduction in marbling grade. The feed efficiency benefit was sustained for 251 days and showed that genetic improvement of feed efficiency will reduce feed costs in a large commercial feedlot.

ACKNOWLEDGEMENTS

The authors thank the commercial feedlot for their collaboration. Shelley Piper was the recipient of a CRC for Beef Genetic Technologies Summer Scholarship.

REFERENCES

Arthur, P.F., Archer, J.A., Johnston, D.J., Herd, R.M., Richardson, E.C. and Parnell, P.F. (2001) *J. Anim. Sci.* **79**:2805.

Gilmour, A.R., Cullis, B.R., Welham, S.J. and Thompson, R. (1999) "ASREML Reference Manual" Biometric Bulletin No. 3, NSW Agriculture, Orange.

SCA (1990) "Feeding Standards for Australian Livestock. Ruminants" Standing Committee on Agriculture, Ruminants Subcommittee. CSIRO, Australia.

META-ANALYSIS OF CROSS-BRED PROGENY DATA FOR AUSTRALIAN TERMINAL SIRE SHEEP

R.G. Banks[1], D. J. Brown[2] and S.R. Field[1]

[1]MLA, c/- Animal Science, University of New England, Armidale 2351
[2]Animal Genetics and Breeding Unit*, University of New England, Armidale, NSW 2351

SUMMARY

Meta-analysis of data from a number of progeny tests, involving crossbred progeny of terminal sire rams with known ASBVs, shows that in general LAMBPLAN ASBVs predict progeny performance as expected. There is however a considerable range around the mean regressions for all traits examined, and in the case of the regression of cross-bred progeny carcase weight on sire post-weaning weight ASBV, some anomalous results requiring further investigation. IN general, lamb producers can be confident in use of LAMBPLAN information in sourcing flock rams.

INTRODUCTION

LAMBPLAN has been providing genetic evaluation services to the Australian lamb industry since 1988 (Banks 1990), and now analyses include data on approximately 105,000 new animals per year from 415 terminal sire breed flocks. The core traits analysed in LAMBPLAN remain growth rate, ultrasound fat depth and eye muscle depth measured on animals in the ram breeding sector, although there has been significant development in terms of stages of growth that can be analysed, as well as in the models for evaluation. In addition, there has been substantial genetic change in all 3 traits over the period 1989-present (Swan *et al.* 2009). The primary aim of LAMBPLAN remains to assist industry identify animals with superior merit for breeding fast-growing slaughter lambs with valuable carcases. This paper examines the relationship between the core LAMBPLAN genetic evaluation traits and traits expressed in cross-bred commercial progeny, by analysing sire-progeny regressions in a number of progeny tests conducted since 1990.

MATERIALS AND METHODS

Data. Data from 45 separate progeny tests were available, comprising 25,134 progeny of 257 sires. These data were from a) Terminal Sire Central Progeny Test trials, conducted during the early 1990s, b) a small number of PIRD (MLA producer initiated research and development) projects and datasets collected by individual stud breeders during the period 1995-present, and c) Information Nucleus sites of the Sheep CRC. Progeny number per progeny test group ranged from 29 to 856, and data recorded varied but always included live and carcase weight at slaughter, a measure of fat or tissue depth on the carcase and eye muscle depth. Traits examined in the cross-bred progeny were birth weight (bwt), post-weaning weight (pwt), carcase weight (cwt), carcase fat depth (cfat) and carcase eye muscle depth (cemd). Breeding values of sires used in analysis were for birth weight (bwt), post-weaning weight (pwt), post-weaning fat (pfat) and post-weaning eye muscle depth (pemd). The fat and muscle depth traits of both sires and progeny were adjusted for body/carcase weight, and procedures for estimating breeding values have been described (Brown *et al*. 2007). Carcase fat measures were the GR Tissue Depth (tissue depth at the 12/13th rib, 110 mm out from the backbone).

* AGBU is a joint venture of NSW Department of Primary Industries and the University of New England

Model of analysis. Individual progeny deviations from contemporary group mean were calculated, fitting age, birth and rearing status, sex, dam age and for the fat and muscle measures, carcase weight fitted as a covariate. Contemporary groups were defined by breed, flock, year of birth, date of measurement and management group. Regressions of progeny deviation on sire breeding value were estimated for each site, year and trait combination, using sire ASBVs from the most recent LAMBPLAN analysis (March 2009). In addition, a combined analysis was conducted across all years and sites, fitting the same fixed effects.

RESULTS AND DISCUSSION

Regression Coefficients. Table 1 summarises the regression coefficients calculated from the combined dataset and the ranges of values obtained within individual datasets.

Table 1. Summary of the regression coefficients (se) across and within datasets

Sire Breeding Value	Progeny Trait	Regression Coefficients			
		All data, all sires	All data, only sires > 50% accuracy	Minimum value across datasets*	Maximum value across datasets*
Birth weight	Birth weight	0.45 (0.04)	0.52 (0.04)	0.06 (0.30)	0.99 (0.15)
Post-weaning live weight	Post-weaning live weight	0.37 (0.01)	0.36 (0.01)	- 0.24 (0.31)	1.09 (0.14)
Post-weaning live weight	Carcase weight	0.05 (0.01)	0.03 (0.01)	- 0.19 (0.18)	0.70 (0.09)
Post-weaning C fat**	Carcase GR Tissue depth**	1.96 (0.09)	1.94 (0.10)	- 0.96 (0.76)	3.22 (0.51)
Post-weaning eye muscle depth**	Carcase eye muscle depth	0.35 (0.02)	0.35 (0.03)	0.00 (0.21)	0.77 (0.11)

*only coefficients estimated from datasets with at least 15 progeny measured are included in this range
**trait is adjusted for constant weight (either live or carcase as appropriate)

With one exception, these sets of regression coefficients are broadly in agreement with expectations;

- Birth weight: cross-bred progeny birth weight increases by 0.45kg for each 1 kg increase in sire breeding value for birth weight. This is very close to expectation.
- Post-weaning live weight: cross-bred progeny post-weaning weight increases by 0.37 kg for each 1 kg increase in sire breeding value for post-weaning weight. This is slightly lower than expectation, and may reflect increased variance in dam contribution to cross-bred progeny weight, or that this contribution is less accurately accounted for in cross-bred progeny, where often little is known about the dams. This may also indicate that the direct genetic effects are larger in the crosbred progeny than observed in the purebred ram breeding flocks, which if coupled with a lower variance in the cross-bred progeny could generate a lower b value.
- Carcase weight: cross-bred progeny carcase weight increases by 0.05 kg for each 1 kg increase in sire breeding value for post-weaning weight. A simple expectation for this relationship would be an 0.25 kg increase, reflecting half for the sire contribution to progeny and half reflecting the approximately 50% dressing weight. This result is also surprising in view of the stronger relationship for post-weaning weight, which is expected to be well correlated with carcase weight.

- Carcase fat depth: the regression here is close to 2mm additional tissue depth in the cross-bred progeny carcase for every 1mm increase in sire's breeding value for C fat depth. There is an approximately 3:1 relationship between GR tissue depth and C site fat depth in the animal itself. This suggests that there is a scale difference between cross-bred progeny and their sires' measures, since otherwise the expected value of the regression would be 1.5:1.
- Carcase eye muscle depth: cross-bred progeny eye muscle depth increases by 0.33mm for each 1 mm increase in sires' breeding value for post-weaning eye muscle depth. Thus the observation is lower than expectation (0.5:1), with unaccounted variation in dam muscling being a possible contributory factor.

In broad terms, these results provide a basis for confidence that sires' breeding values predict crossbred progeny outcomes. However, two questions remain: why is the regression for carcase weight lower than expected, and how much variation is observed in these regressions, or in other words, how confident can a user be in the predictions?

Regression of carcase weight on sire post-weaning weight. Assuming a high genetic correlation between post-weaning weight and carcase weight (these are measured at similar ages), we would expect the sire:offspring regression between these two traits to be around 0.25. Here we find an estimate across all data sets of 0.05, with a range from -0.19 to 0.70. At the same time, the sire:progeny regressions for post-weaning weight, the live trait, is 0.37, slightly lower than expected.

Two possible explanations for this anomaly are:

- changes in the data recorded and/or the model for analysis of post-weaning weight mean that there has been a change in what is being predicted over time
- the genetic correlation between live weight and carcase weight is lower than anticipated.

The first hypothesis reflects the following changes: when LAMBPLAN began the main weight trait recorded in pure-bred studs as at around 10-12 months, rather than the 7-10 months now predominant, and closer to the age at which cross-bred lambs are slaughtered; and the analytical model for weight traits now separates the direct, maternal and common environment contributions. Together, these effects would lead to a time trend in the regression coefficients. Figure 1 shows the pattern of regression estimates against the year of observation.

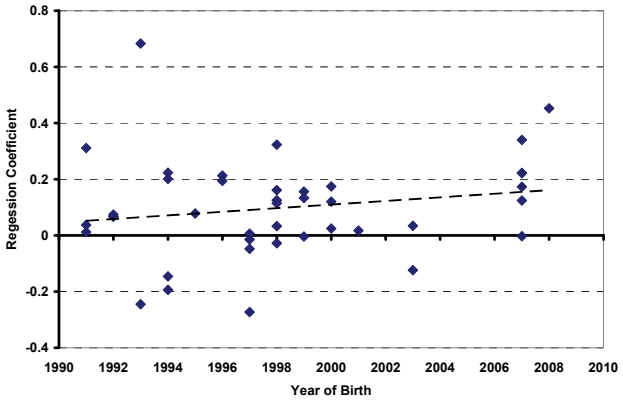


Figure. 1 Estimated carcase weight regression coefficient vs Year of Birth of lambs in each dataset

The regression has a positive slope, but is not significant. This does not suggest any clear change in the regression due to changes in the data. Trends in the regression coefficients across time were small but not significantly positive for carcase fat and carcase eye muscle depth.

The second hypothesis is that the correlation between live and carcase weight in cross-bred lambs is lower than might be expected. This could either be due to a biological relationship, such as there being considerable variation in skin weight and/or gut fill within datasets, or possibly statistical artefacts due to constrained slaughter groups. This latter possibility warrants further investigation.

A further aspect of these results with significance for extension messages is the large variation amongst estimates across sites and years. This is apparent for both the weight and composition traits, and can be seen clearly in Figure 1 for instance for year 2007. The data points in this year are all from Information Nucleus sites, with the same set of sires in each case, yet the regression coefficients for carcase weight on sire post-weaning weight range from 0 to 0.35. Similar patterns (not shown) are apparent for carcase fat depth and carcase eye muscle depth. In the case of Information Nucleus sites, the regression estimates are based on 100-200 progeny of at least 20 sires. This range in the observed regressions estimates again points to the need for further investigation of the carcase data.

CONCLUSIONS

This meta-analysis has shown that;

- for live weight and for carcase composition traits, regressions of cross-bred progeny performance on the corresponding sire trait LAMBPLAN breeding values are similar to expectations across a range of datasets collected between 1990 and 2008
- the regression for carcase weight on sire post-weaning weight in these data is lower than anticipated, and as yet unexplained
- there is considerable variation amongst the estimates across sites and years for all traits, which may be partly due simply to sampling effects, but also suggests that more detailed examination of factors affecting carcase weight is warranted.

Taken as a whole, these results suggest that lamb producers can confidently use LAMBPLAN breeding values to choose sires that will breed lambs with superior growth and carcase characteristics. The regression of progeny carcase weight on sire post-weaning weight predicts the impact of that selection on processor and retailer returns, and the values for that regression estimated here suggest that further investigation of both the data and the analytical model should be conducted.

ACKNOWLEDGMENTS

Sheep CRC is acknowledged for allowing use of data from Information Nucleus sites. This research was supported by Meat and Livestock Australia.

REFERENCES

Banks, R. G. (1990) Proc. Aust. Assoc. Anim. Breeding Genet. **8**:237

Brown, D. J., Huisman, A. E., Swan, A.A., Graser, H-U., Woolaston, R.R., Ball. A.J., Atkins, K.D. and Banks, R. G. (2007) Proc. Assoc. Advmt. Anim. Breed. Genet. **17**:187

Swan, A.A., Brown, D.J., and Banks, R.G. (2008) *Proc. Assoc. Advmt. Anim. Breed. Genet.* **18**:326.

EVALUATION OF THE ANGUS BREEDPLAN IMF% EBV IN 100d-FED ANGUS x HEREFORD STEER PROGENY

S.A. Barwick[1], D.J. Johnston[1], M.L. Wolcott[1], J.F. Wilkins[2], W.A. McKiernan[3]

Cooperative Research Centre for Beef Genetic Technologies, Armidale, NSW 2351
[1]Animal Genetics and Breeding Unit*, University of New England, Armidale, NSW 2351
[2]New South Wales Department of Primary Industries, Wagga Wagga, NSW 2650
[3]New South Wales Department of Primary Industries, Orange, NSW 2800

SUMMARY

The Angus BREEDPLAN IMF% (intramuscular fat %) EBV was evaluated in 320 100d-fed Angus x Hereford steer progeny of 25 sires. Results showed using the EBV to select among sires would significantly increase marbling scores and IMF%, and that this was unaffected by growth regime to feedlot entry or by whether slaughter was at a constant age or weight. For the production system studied, Angus breed average IMF% EBV level was associated with MSA-scale marbling scores of 1.6 and 1.5, and IMF% levels of 4.6 and 3.9 in 'fast' and 'slow' grown steers (0.71 and 0.48 kg/d, respectively). The ability of the sire IMF% EBV to explain differences among individual steers was low but in line with expectation. The same expectation limits how effective any EBV can be in sorting individual animals of a breed for a management or marketing purpose.

INTRODUCTION

There is an ongoing need for the performance of EBVs to be 'proven' in industry data, as this is important to all adoption of genetic evaluation and for achieving faster rates of genetic gain in the beef industry (Upton *et al*. 2007). Because production systems and components of genetic evaluation (e.g. data recording accuracy, traits recorded, parameter estimates used) can differ among breeds, the need extends across breeds, traits and time. Here, the Angus BREEDPLAN IMF% (intramuscular fat %) EBV is examined for its ability to predict marbling and IMF% differences in 100d-fed Angus x Hereford steer progeny, a production system where Meat Standards Australia (MSA) marbling scores are often only 1 to 2 (McKiernan *et al.* 2005). We relate Angus IMF% EBV level to phenotype level to provide breed benchmark information for valuing marbling improvement in breeding objectives (Barwick and Henzell 2003); and we briefly comment on the percentage of phenotypic variation that an EBV can be expected to explain.

MATERIALS AND METHODS

Data. Records for marbling on an MSA scale (MSA-MS; scored 0 to 6, but with 0.1 divisions within each score) and a USDA scale (USDA-MS), and chemical IMF% (IMF) were available for up to 320 steer progeny of 25 Angus sires from an experiment in south-western New South Wales described by McKiernan *et al*. (2005). Only Angus-sired progeny from the experiment were used. Two records for each of MSA-MS and USDA-MS, and three for IMF, were considered outliers for being more than 3 s.d. from their contemporary group mean, and were excluded from analyses.

Experiment design. Steers by Angus sires of a range of genetic potential were bred by AI from Hereford cows over 5 year-seasons from 2000 to 2002. Steers were grown from weaning to feedlot entry (~400 kg) under 'fast' or 'slow' growth regimes. Unadjusted mean growth rates for Angus-sired steers of the two growth regimes were 0.71 and 0.48 kg/d. There were 2 management groups

* AGBU is a joint venture of NSW Department of Primary Industries and University of New England

for each growth regime. Steers were 100d-fed in 5 slaughter groups, using 1 or 2 feedlot pens per group. Marbling scoring was carried out by MSA graders, with 1 or 2 graders per slaughter group.

Angus sire EBVs. BREEDPLAN IMF% EBVs are computed from breed database data and include adjustment to a 300 kg carcass (Graser *et al*. 2005). The BREEDPLAN IMF% EBVs for Angus sires here were calculated in November 2008. The present steer data did not contribute to the calculated EBVs. The average level of IMF% EBV for all 2007-born animals in the Angus database (the current young age group) was taken to be a relevant estimate of the current Angus breed average level of IMF% EBV. This breed average level of IMF% EBV was 0.8.

Statistical analyses. Analyses were carried out with PROC MIXED in SAS (SAS Inst., Cary, NC). MSA-MS, USDA-MS and IMF were each analysed at a constant slaughter age (within growth regime) and at a constant slaughter weight (within growth regime), by including either an age or liveweight covariate deviated from growth regime means. Initial models also included slaughter group, growth regime, management group within growth regime, pen within slaughter group, grader within slaughter group (only for MSA-MS and USDA-MS) and first-order interactions, with sire as a random effect. Growth regime and slaughter group jointly defined year-season. Heritabilities were estimated from final sire models, obtained by systematically omitting non-significant ($P > 0.05$) effects. These models included slaughter group and growth regime, and grader within slaughter group for MSA-MS and USDA-MS.

Regressions on sire IMF% EBV were evaluated by adding sire IMF% EBV, its interaction with growth regime, and the sire IMF% EBV quadratic effect to the above models, with sire excluded. Non-significant effects were again omitted. The MSA-MS, USDA-MS and IMF associated with an IMF% EBV level of 0.8 were predicted from these final models. Residuals from final models, excluding sire IMF% EBV and sire, were also derived for analysis both as individual steer phenotypes and as sire progeny means. Sire IMF% EBV was the only effect fitted in these further analyses, which were used to assess the percentage of phenotypic variation explained.

Table 1. Summary of data for 100d-fed Angus x Hereford steers, and for Angus sire IMF% EBVs

Variable	N	Mean	s.d.	Minimum	Maximum
Age at slaughter (d)	317	668	92	497	843
Liveweight (kg)	320	644	52	498	802
Carcass weight (kg)	320	359	30	267	451
Carcass P8 fat depth (mm)	320	18.8	5.0	9	36
MSA-scale marbling score	318	1.4	0.4	0.2	2.8
USDA-scale marbling score	318	357	50	230	560
Chemical IMF%	263	4.1	1.5	1.1	9.8
Sire IMF% EBV	320	0.50	0.91	-1.19	1.76
Sire IMF% EBV accuracy (%)	320	83.4	9.0	64	97

RESULTS AND DISCUSSION

A summary of the data is in Table 1. Other results are in Table 2. Heritabilities for MSA-MS, USDA-MS and IMF (0.38 to 0.48) were in the range usually observed for these traits; and variances for the same traits defined at different slaughter end-points were very similar.

Table 2. Sire IMF% EBV regression coefficients[1,2] (± s.e.) for MSA- and USDA-scale marbling scores and chemical IMF% at a constant age[3] (A) or weight[4] (W) in 100d-fed Angus x Hereford steers; and sire model estimates of variances and heritabilities.

Trait	MSA-MS		USDA-MS		Chemical IMF%	
Regression	A	W	A	W	A	W
Sire IMF% EBV	0.120*** (±0.022)	0.109*** (±0.022)	15.80*** (±2.78)	14.33*** (±2.73)	0.399*** (±0.085)	0.363*** (±0.083)
σ^2_P	0.129	0.123	2129	2016	1.68	1.59
σ^2_G	0.062	0.056	904	766	0.80	0.74
h^2	0.48	0.46	0.42	0.38	0.47	0.46

[1] *** indicates $P < 0.001$

[2]Pooled regressions; individual regressions for steer growth regimes did not differ ($P > 0.05$)

[3]At 600 and 751d for 'fast' and 'slow' steer growth regimes, respectively

[4]At 634 and 657 kg for 'fast' and 'slow' steer growth regimes, respectively

Breed benchmarks. Model predictions showed Angus sires of breed average IMF% EBV (ie. 0.8) were associated with age-constant steer progeny MSA-MS levels of 1.6 and 1.5 score, USDA-MS levels of 377 and 367 score, and IMF levels of 4.6 and 3.9 % when grown to feedlot entry at 0.71 kg/d and 0.48 kg/d, respectively. These predictions are specific to the 100d feedlot-fed production system studied, including to the joining of sires to Hereford cows.

Regressions on sire IMF% EBV. Pooled regressions on sire IMF% EBV were significant ($P < 0.001$) for each of MSA-MS, USDA-MS and IMF, whether these traits were defined at a constant age or constant weight (Table 2). This confirms the Angus IMF% EBV as a criterion for genetically improving marbling and IMF for the production system examined. It shows that the EBV is expected to be effective irrespective of the growth regime of steers and of whether slaughter is at a constant age or weight. There were no significant ($P > 0.05$) quadratic effects or differences in sire IMF% EBV regressions between growth regimes for any trait.

The regression coefficients for sire EBV on progeny performance were not significantly different from the usual expectation of 0.5 for IMF traits (Table 2), the traits that were expected to be closest to the EBV target trait (Johnston *et al*. 1999). Note that the 0.5 expectation applies only when the progeny trait is exactly the same as that for which the EBV was derived.

Ability to explain differences in sire progeny means. The percentage of variation among sire progeny means that was explained by sire IMF% EBV, at a constant age or weight, respectively, was 47.0 and 38.5 % for MSA-MS, 53.5 and 48.3 % for USDA-MS, and 41.4 and 29.2 % for IMF. These percentages are influenced by the accuracies of the IMF% EBVs that were available for sires, and which were relatively high (Table 1). The sizeable percentages of variation explained nonetheless are support for the use of the Angus IMF% EBV in sire selection.

Ability to explain individual steer differences. Sire IMF% EBV explained only a low percentage of the phenotypic variation among individual steers (e.g. 8.0 % for age-constant MSA-MS; Figure 1). Since the maximum percentage that a sire EBV, by itself, can explain is ¼ h^2 (x 100), or here 12%, this result is not surprising. The maximum percentage that can be explained occurs when the EBV has perfect accuracy and is for exactly the same trait as that predicted. When there are EBVs for individual steers (which might usually have lower accuracies than the EBVs of sires), the maximum percentage the EBV can explain is h^2 (x 100). Again, this maximum is expected only when the EBV has perfect accuracy and is for exactly the trait predicted. These maxima also

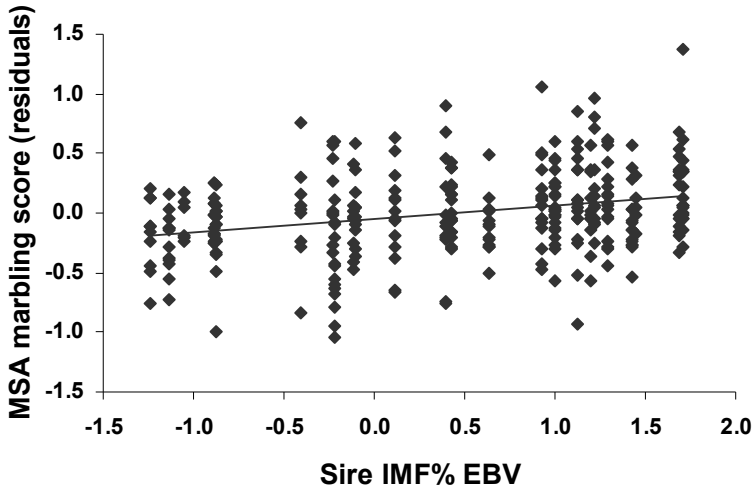


Figure 1. Residual phenotypic variation in individual steer age-constant MSA marbling score, plotted against sire IMF% EBV. The regression explains 8.0 % of the variation.

broadly indicate the limits that exist on the use that can be made of any EBV for sorting individual steers of a breed for a management or marketing purpose.

CONCLUSIONS

- Angus IMF% EBVs can be used to select among sires to increase marbling and IMF% in 100d-fed Angus x Hereford steer progeny. This result is unaffected by the growth regime of steers from weaning to feedlot entry.
- Angus breed average marbling score and IMF% benchmarks were able to be assessed for the production system examined, which included the joining of sires to Hereford cows.
- There was a lower ability of the sire IMF% EBV to explain differences among individual steers, but this was in line with expectation. The same expectation limits how effective any EBV can be in sorting individuals of a breed for a management or marketing purpose.

ACKNOWLEDGMENTS

We are pleased to acknowledge the support of AgReserves Australia Ltd, Meat and Livestock Australia, and authors' organisations; the skilled assistance of staff of the Meat Science Laboratory (Armidale), Jeffrey House, Greg Meaker and other colleagues; and the cooperation of Cargill Beef Australia Ltd, staff of the properties 'Bringagee' and 'Kooba', and the Angus Society of Australia.

REFERENCES

Barwick, S. A. and Henzell, A. L. (2003) *Proc. Assoc. Advmt. Anim. Breed. Genet.* **15**:347.
Graser, H-U., Tier, B., Johnston, D. J. and Barwick, S. A. (2005) *Aust. J. Exp. Agric.* **45**:913.
Johnston, D. J., Reverter, A., Thompson, J. M. and Perry, D. (1999) *Proc. Assoc. Advmt. Anim. Breed. Genet.* **13**:345.
McKiernan, W. A., Wilkins, J. F., Barwick, S. A., Tudor, G. D., McIntyre, B. L., Graham, J. F., Deland, M. P. B. and Davies, L. (2005) *Aust. J. Exp. Agric.* **45**:959.
Upton W. H., Nicol, D. C. and Freer, R. E. (2007) *Proc. Assoc. Advmt. Anim. Breed. Genet.* **17**:158.

GENETIC PARAMETERS FOR FEATHER WEIGHTS OF BREEDING OSTRICHES

Z. Brand[1,2] and S.W.P. Cloete[1,3]

[1] Department of Animal Sciences, University of Stellenbosch, Private Bag X1, Matieland 7602, SA
[2] Institute for Animal Production: Oudtshoorn, PO Box 313, Oudtshoorn 6620, SA
[3] Institute for Animal Production: Elsenburg, Private Bag X1, Elsenburg 7607, SA

SUMMARY

After the initial decline in value, ostrich feathers have become an integral part of the remuneration of ostrich producers. Categorised (according to the location where it was harvested) mature square-root transformed feather weights depended on gender (males having a 20% higher total feather weight than females) and genotype (South African Black ostriches having a 34% higher total feather weight than Zimbabwean Blues). Animal age interacted with gender, the difference between males and females diminishing from 10 years of age, with no significant gender difference in birds of 11+ years. Heritability estimates of categorised feather weights were low to moderate, and high at 0.30 ± 0.06 for total feather weight. The corresponding value for animal permanent environment amounted to 0.08 ± 0.04. Genetic correlations between categorised feather traits were moderate (0.54 ± 0.16) to very high (0.90 ± 0.10). The present results indicate that feather quantity can be improved by selection in ostriches.

INTRODUCTION

Ostrich farming started between 1857 and 1864 as a commercial enterprise in South Africa (Smith 1963) and has been an important contributor to the agricultural economy of South Africa for the past 150 years. Initially, ostrich feathers were the main commercial product and were highly prized by the European fashion industry. During the First World War between 1914 and 1918 the South African ostrich feather industry collapsed (Smith 1963). During the following years the ostrich industry recovered, but the skin and meat became the more dominant sources of income. Research on the improvement of the quality or quantity of ostrich feathers was thus discontinued. Swart *et al.* (1984) assessed the impact of feather quality traits on price determination in the marketplace, but no scientific selection programmes were implemented as a result. The objective of this study was thus to estimate environmental and genetic parameters influencing quantitative ostrich feather traits.

MATERIAL AND METHODS

The experimental population used for the study (2001-2006) was the commercial, pair-bred ostrich flock at the Oudtshoorn Research Farm in the Klein Karoo region of South Africa. The origin of the flock and the general management procedures implemented has been described previously (Bunter & Cloete 2004). During 2003, Zimbabwean Blue (ZB) breeders were introduced to the flock (Cloete *et al.* 2008). The flock consisted of 188 breeding pairs and the age of breeder birds in the flock ranged between 2 and 11+ years. The annual breeding season usually lasted for about eight months (from June to January of the following year) followed by a four month rest period. During the annual resting period, feathers were harvested, sorted, categorized and weighed, ensuring repeated animal-year records for the mature breeding birds in the flock at that stage. The feather categories considered were the 6 feather types harvested from the ostrich, namely, floss (one row of soft downy feathers under the wing), short hard body feathers (feathers in the centre of the dorsal surface of the wing just before long hard feathers), long hard body feathers (second and third row of shorter feathers on the outer edge of the wing), tail feathers, white plumes (first row of prominent plumes at the edge of the wing), and short body floss

(feathers under the wing and on the front and behind of the thigh). Separate weights for the categories of feathers were summed to obtain a total feather weight (TFW) for each bird. All feather weights were extremely variable, with coefficients of variation ranging from 32.3 to 58.9%. The data were therefore subjected to a square root transformation to stabilise the variance.

The data were subjected to single-trait genetic analyses, using ASREML software (Gilmour *et al.* 1999). Fixed effects that were considered included the two genotypes (SAB or ZB), gender (male or female), and breeder age (2 – 11+ years). Repeated records pertaining to each animal were accommodated by fitting direct additive animal (related to the data by the numerator relationship matrix) and animal permanent environment (related to the data by an identity matrix) as random effects. Subsequently, a six-trait animal model was fitted, involving the different feather categories as traits. TFW was excluded from this analysis, as it is a function of the weight of the other categories of feathers.

RESULTS AND DISCUSSION

Table 1 shows considerable variation for each of the 7 square root transformed feather weight traits, with coefficients of variation ranging from 17.3% to 31.3%. The geometric mean for TFW amounted to 615g, with a coefficient of variation (CV) of 17.3%. No comparable estimates for feather weights of ostriches were found in the literature.

Table 1. Descriptive statistics for square root transformed feather traits of breeding ostriches for the 2001 – 2006 production years

Traits	Number of records	Raw mean ± SD	CV (%)	Kurtosis	Skewness
Floss (g)	1429	4.31 ± 1.35	31.3	-0.36	0.08
Short hard bodies (g)	1673	11.5 ± 2.5	21.7	0.09	-0.12
Long hard body (g)	1734	9.39 ± 2.39	25.5	-0.05	0.01
Tail (g)	1676	9.08 ± 2.23	24.6	0.00	-0.08
White plumes (g)	1729	13.9 ± 2.4	17.3	0.46	-0.57
Short body floss (g)	1632	11.1 ± 2.4	21.6	0.30	-0.20
Total (g)	1782	24.8 ± 4.3	17.3	1.18	-0.41

SD, standard deviation; CV%, coefficient of variation

With the exception of floss and tail feathers, the weights of all feather types were ($P<0.05$) heavier for males than for females (Table 2). This resulted in a 20% higher TFW for males relative to females (respective geometric means being 600 vs. 501g respectively). The weights for all the different feather traits were higher ($P<0.05$) for SAB breeders compared to ZB breeders, long hard body plumes being the only exception. Overall, TFW of SAB breeders exceeded that of ZB breeders by 34%. Geometric means for TFW in the two genotypes amounted to 635g for the SAB genotype, compared to 475g in the ZB genotype. Gender and animal age were found to interact ($P<0.05$) for the respective categories of feather weights. Males produced a greater quantity of feathers up to an age of 10 years, where after the gender difference was reduced to a tendency ($P<0.10$).

The results of single-trait analyses for the categorised feather weights are provided in Table 3. The heritability estimate (h^2) for long hard body feathers were moderate at 0.20 ± 0.05, while the h^2 of the rest of the feather weights were low to moderate, ranging between 0.11 ± 0.05 and 0.16 ± 0.05. All these h^2 estimates were significant (at least double the corresponding SE). The h^2 estimate of TFW was fairly high at 0.30 ± 0.06.

Table 2. Least-squares means (± SE), depicting the effects of gender and genotype on the weight of different categories of square root transformed feather traits of breeding ostriches for the 2001 – 2006 production years ([a,b] depict differences between gender groups or genotypes; P<0.05)

Traits	Gender		Genotype	
	Males	Females	SAB	ZB
Floss (g)	4.19 ± 0.26	3.90 ± 0.15	4.40 ± 0.9[a]	3.80 ± 0.14 [b]
Short hard bodies (g)	11.8 ± 0.41[a]	10.4 ± 0.2[b]	11.8 ± 0.1[a]	10.7 ± 0.2[b]
Long hard body (g)	9.70 ± 0.43[a]	8.69 ± 0.25[b]	9.37 ± 0.17	9.10 ± 0.21
Tail (g)	8.69 ± 0.43	8.31 ± 0.26	9.26 ± 0.16[a]	7.78 ± 0.22[b]
White plumes (g)	13.6 ± 0.5[a]	12.6 ± 0.3[b]	14.1 ± 0.2[a]	12.0 ± 0.2[b]
Short body floss (g)	11.2 ± 0.4[a]	10.1 ± 0.2[b]	11.4 ± 0.2[a]	9,8 ± 0.2[b]
Total (g)	24.5 ± 0.7[a]	22.4 ± 0.5[b]	25.2 ± 0.3[a]	21.8 ± 0.4[b]

These results indicate that genetic improvement can be achieved in ostrich feather weights. The only previous indication that feather quantity may respond to selection come from Louw and Swart (1982). The latter authors reported that the number of wing quills had a heritability of 0.24. Animal permanent environment (pe^2) accounted for a proportion of between 0.08 and 0.22 of the phenotypic variation associated with the respective categories of feather weights.

Genetic parameters from the six-trait analysis are presented in Table 4. Estimates of h^2 for the weight of all categories of feathers were generally somewhat higher compared to those estimated from single-trait analyses, at the expense of animal permanent environment. It is reasonable to expect that the other categorised feather weights in a multi-trait analysis would contribute to a better partitioning of h^2 and pe^2 effects. Genetic correlations (r_g) among all categorised feather weights were moderate to very high and ranged from 0.54 ± 0.16 to 0.90 ± 0.10. Permanent environmental correlations were mostly lower than genetic correlations, while environmental correlations resembled genetic correlations in sign, but were considerably lower in magnitude.

Table 3. Estimated variance components and ratios (± SE) of the different square root transformed feather weight traits as derived from single-trait analyses of mature breeding ostriches for the 2001 – 2006 production years

Traits	σ^2_a	σ^2_{pe}	σ^2_e	h^2	pe^2
Floss (g)	0.166	0.331	1.029	0.11 ± 0.05	0.22 ± 0.05
Short hard bodies (g)	0.581	0.475	3.049	0.14 ± 0.05	0.12 ± 0.04
Long hard body (g)	0.858	0.503	3.042	0.20 ± 0.05	0.11 ± 0.04
Tail (g)	0.669	0.841	2.845	0.15 ± 0.05	0.19 ± 0.05
White plumes (g)	0.726	0.847	3.007	0.16 ± 0.06	0.19 ± 0.05
Short body floss (g)	0.686	0.691	2.877	0.16 ± 0.05	0.16 ± 0.05
Total (g)	4.012	1.081	8.181	0.30 ± 0.06	0.08 ± 0.04

σ^2_a, direct additive variance component; σ^2_{pe}, female permanent environmental variance component; σ^2_e, environmental (residual) variance component; h^2, direct heritability; pe^2, animal permanent environment

CONCLUSION

Results from this study confirm that feather quantity in ostriches is heritable, and would respond to directed selection if desired. Moderate to high genetic correlations indicate that

selection for specific categories of feathers are likely to result in correlated responses in other categories. However, below unity genetic correlations suggest that feather weights on different body parts may not always be the same genetic trait. Further investigations are needed to assess the genetics of ostrich feather quality traits.

Table 4. Estimates (± SE) of heritability (h^2), animal permanent environment (pe^2) (in bold letters) as well as genetic, permanent environmental and environmental correlations (in normal print) among square root transformed feather weight traits, as derived from a six-trait analysis

Traits	Traits	r_g	r_{pe}	r_e
Floss (g)	Floss (g)	**0.16 ± 0.05**	**0.12 ± 0.05**	
	Short hard bodies (g)	0.73 ± 0.14	0.19 ± 0.26	0.35 ± 0.03
	Long hard body (g)	0.54 ± 0.16	-0.04 ± 0.54	0.28 ± 0.03
	Tail (g)	0.73 ± 0.14	0.03 ± 0.25	0.18 ± 0.03
	White plumes (g)	0.74 ± 0.14	-0.15 ± 0.40	0.24 ± 0.03
	Short body floss (g)	0.54 ± 0.18	0.40 ± 0.23	0.24 ± 0.03
Short hard bodies (g)	Short hard bodies (g)	**0.16 ± 0.04**	**0.10 ± 0.04**	
	Long hard body (g)	0.73 ± 0.12	0.74 ± 0.50	0.13 ± 0.03
	Tail (g)	0.74 ± 0.13	0.12 ± 0.25	0.21 ± 0.03
	White plumes (g)	0.73 ± 0.12	0.11 ± 0.37	0.24 ± 0.03
	Short body floss (g)	0.87 ± 0.10	0.67 ± 0.18	0.31 ± 0.03
Long hard body (g)	Long body floss (g)	**0.21 ± 0.05**	**0.03 ± 0.04**	
	Tail (g)	0.82 ± 0.10	0.08 ± 0.46	0.20 ± 0.03
	White plumes (g)	0.66 ± 0.12	0.99 ± 0.62	0.20 ± 0.03
	Short body floss (g)	0.81 ± 0.11	0.23 ± 0.43	0.25 ± 0.03
Tail (g)	Tail (g)	**0.20 ± 0.05**	**0.14 ± 0.06**	
	White plumes (g)	0.90 ± 0.10	-0.30 ± 0.41	0.07 ± 0.03
	Short hard bodies (g)	0.74 ± 0.14	0.27 ± 0.21	0.16 ± 0.03
White plumes (g)	White plumes (g)	**0.22 ± 0.05**	**0.07 ± 0.04**	
	Short body floss (g)	0.65 ± 0.14	0.35 ± 0.29	0.14 ± 0.03
Short body floss (g)	Short body floss (g)	**0.18 ± 0.05**	**0.13 ± 0.05**	

r_g, genetic correlation; r_{pe}, permanent environmental correlation, r_e, residual correlation

ACKNOWLEDGEMENTS

We are grateful for the support of those responsible for the recording of the data used in study as well as for the husbandry of the resource flock.

REFERENCES

Bunter, K.L. and Cloete, S.W.P. (2004) *Livest. Prod. Sci.* **91**: 9

Cloete, S.W.P., Brand, M.M., Hoffman, L.C. and Muller, M. (2008) *S. Afr. J. Anim. Sci*. **39**:65.

Gilmour, A.R., Cullis, B.R., Welham, S.J. and Thompson, R. (1999) ASREML – Reference manual. NSW Agriculture Biometric Bulletin No. 3 NSW Agriculture, Orange Agriculture Institute, Forest Road, Orange 2800, NSW, Australia.

Louw, J.H. and Swart, D. (1982) *S. Afr. J. Sci*. **78**:455.

Swart, D., Heydenrych, H.J. and Poggenpoel, D.G. (1984) *S. Afr. J. Anim. Sci.* **14**:45.

GENETICS OF LAMB SURVIVAL: A STUDY OF MERINO RESOURCE FLOCKS IN SOUTH AUSTRALIA

F.D. Brien, M.L. Hebart, K.S. Jaensch, D.H. Smith and R.J. Grimson

South Australian Research and Development Institute, Roseworthy, SA, 5371

SUMMARY

Estimates of genetic parameters and variance components were made for lamb survival and correlated traits on data from the SA Merino Resource Flock (1988-1997) and the SA Selection Demonstration Flocks (1996-2006). Very low estimates of direct heritability were obtained for lamb survival, with values of 0.071, 0.044, 0.043, 0.032 and 0.032 for surviving birth, up to 3, 7, 40 (marking) and 91 days (weaning) after birth, respectively, indicating only slow progress would be likely from genetic improvement. Birth weight had low positive phenotypic correlations with lamb survival (values from 0.112 to 0.224), with the genetic correlations being negative to near zero (values from –0.137 to –0.025), indicating little would be gained in lamb survival from genetic manipulation of birth weight. The phenotypic correlations between lamb survival and birth coat score were virtually zero in all cases, however the genetic correlations were low and consistently positive (values from 0.071 to 0.192), being highest for survival to 3 days of age.

INTRODUCTION

Lamb survival rates above 90% for singles and 80% for twins are relatively uncommon in Australia, with such rates remaining a major factor contributing to reproductive inefficiency in the national sheep flock (Hinch 2008, *pers.com*). Heritability estimates for lamb survival suggest that making genetic gains would be slow (Safari *et al.* 2005). In 2007 the Sheep CRC commissioned studies on experimental Merino flocks where the data could provide parameter estimates of much higher precision than those currently available. This paper reports on one of these studies.

MATERIALS AND METHODS

Data was obtained from two projects, the SA Merino Resource Flock (SAMRF) and the SA Selection Demonstration Flocks (SDF), both conducted at Turretfield Research Centre.

Established in 1988, the foundation flock for the SAMRF project consisted of 2000 South Australian Merino strain ewes representative of the Bungaree and Collinsville family groups. Annually, 48 rams were selected from four studs representing the family groups and single sire mated to approximately 40 randomly allocated ewes each. Lambs were born in April-May of each year. For more details, see Gifford *et al.* (1990).

In 1996, to establish the SDF project, ewes were sourced from the SAMRF to establish four flocks of 200 ewes each, representing three major selection approaches and a randomly selected control. In 1999, a Meat Merino line was added. All SDF lambs were born in June-July of each year. For more details, see Kemper *et al.* (2006).

Data Collection. Pregnant ewes were allocated to lambing paddocks of two-hectares each. Twice-daily lambing rounds were conducted, with lambs identified with their dams within a maximum of 18 hours of birth. Lambs were weighed, ear tagged and scored for birth coat (BCS) (Ponzoni *et al.* 1997). Any dead lambs were recorded. Lambs were marked and mulesed at an average age of 40 days and then weaned from their dams at an average age of 91 days, with the identity of all surviving lambs recorded at those times. Before marking, daily checks were conducted, with less frequent checks made after marking. All deaths were recorded.

Statistical Analysis. The data were analysed with the ASREML software (Gilmour *et al.* 2006), using an animal model, treating lamb survival as a trait of the lamb. Further analyses, treating lamb survival as a trait of the dam have been performed, but are not reported in this paper. Litter size consisted of 2 classes, single and multiple. There were too few higher order births to warrant treating these separately to twins. Year of lamb birth, flock nested within year, litter size (singles, multiples), age of dam (2,3,4,5,6&7+ years), sex and date of birth were fitted as fixed effects in the statistical model and the additive genetic, maternal, and dam permanent environmental variances were fitted as random effects. The data included information from 23,873 individuals, 605 sires, and 7,526 dams and included all animals that were dead at birth. Five measures of lamb survival were assembled from the data, (i) survived birth (Birth, as recorded at lamb tagging), (ii) survived to 3 days after birth (3 days), (iii) survived to 7 days after birth (7 days), (iv) survived to marking (Mark), and (v) survived to weaning (Wean). All these measures are reported as a proportion of the total number of lambs born (alive and dead).

RESULTS

Environmental Effects. The survival rate for all flocks to weaning was 79%. Of the 21% that died, 40% were dead at birth (by lamb tagging), 74% were dead within 3 days and 80% were dead within 7 days of birth. The survival to weaning of singles was greater than multiples (86% vs. 73%, Table 1). Female lambs had higher survival rates than males, regardless of dam age or birth type (Table 2). The survival of lambs from 2 year old and 6+ year old dams was the poorest.

Table 1. Survival rates of singles vs. multiples between birth and weaning

Birth Type	Number	Birth	3 days	7 days	Marking	Weaning
Singles	11,792	0.94	0.89	0.88	0.87	0.86
Multiples	12,081	0.89	0.82	0.80	0.75	0.73
Combined	23,873	0.91	0.85	0.83	0.80	0.79

Table 2. The effects of sex (female – F, male – M) and age of dam on lamb survival

	Singles					Multiples				
	Birth	3 days	7 days	Mark	Wean	Birth	3 days	7 days	Mark	Wean
					Sex effect					
F	0.94	0.90	0.90	0.89	0.89	0.90	0.82	0.81	0.74	0.73
M	0.92	0.88	0.87	0.85	0.85	0.90	0.80	0.78	0.73	0.72
					Age of dam effect					
2	0.93	0.88	0.87	0.86	0.86	0.87	0.78	0.76	0.67	0.66
3	0.94	0.91	0.91	0.90	0.89	0.91	0.84	0.82	0.76	0.76
4	0.94	0.92	0.91	0.90	0.90	0.91	0.84	0.83	0.78	0.77
5	0.94	0.90	0.90	0.88	0.88	0.91	0.84	0.83	0.78	0.76
6	0.92	0.88	0.88	0.89	0.89	0.89	0.78	0.76	0.72	0.71
7 +	0.91	0.85	0.84	0.81	0.82	0.89	0.75	0.76	0.69	0.68

Genetic Parameters. Estimates of genetic parameters for lamb survival are shown in Table 3. The heritability estimates were very low. Direct heritability declined from the time period of surviving birth to those for the periods including up to 3 and 7 days after birth (0.071 to 0.044 and 0.043, respectively) with the estimates only slightly less for the periods up to marking and weaning. Correlations of lamb survival with birth weight and birth coat score are given in Table 4.

Table 3. Heritability estimates for lamb survival. Standard errors are shown in brackets

Lamb Survival Trait	Direct h^2	Maternal h^2
Birth	0.071 (0.008)	0.016 (0.006)
3 Days	0.044 (0.008)	0.024 (0.006)
7 Days	0.043 (0.008)	0.019 (0.007)
Mark	0.032 (0.007)	0.020 (0.007)
Wean	0.032 (0.007)	0.020 (0.007)

Table 4. Phenotypic (r_p) and genetic (r_g) correlations of lamb survival with birth weight (Birth Wt) and birth coat score (BCS). Standard errors are shown in brackets

Lamb Survival Trait	Birth Wt		BCS	
	r_p	r_g	r_p	r_g
Birth	0.112 (0.008)	-0.137 (0.139)	-0.037 (0.009)	0.071 (0.120)
3 days	0.183 (0.008)	-0.045 (0.147)	0.014 (0.008)	0.192 (0.120)
7 days	0.184 (0.008)	-0.112 (0.112)	0.013 (0.008)	0.130 (0.110)
Mark	0.224 (0.008)	-0.111 (0.161)	0.013 (0.008)	0.127 (0.119)
Wean	0.213 (0.008)	-0.025 (0.165)	0.014 (0.008)	0.109 (0.120)

All correlations between lamb survival and birth weight were low, with phenotypic correlations being positive and genetic correlations being smaller and negative. Similarly, all correlations of lamb survival with birth coat score were low; phenotypically, the correlations were virtually zero in all cases. However, the genetic correlations were consistently positive, with the highest being the correlation with lamb survival to 3 days of age.

DISCUSSION

The data set for lamb survival traits used in our study is among the largest yet to be reported and provides precise estimates of genetic parameters. We confirm earlier reports (e.g. Obst and Day, 1968) that the great majority of lamb loss is in the first 72 hours after birth.

Direct selection for lamb survival. This study confirms earlier reports of very low heritability estimates obtained for lamb survival (Safari *et al.* 2005) and is in close agreement with a recent study on a data set of similar size (Hatcher *et al.* 2009). Response to selection is however not only a function of heritability, but also available variation, selection intensity and generation interval. Similar to Fogarty *et al.* (2006), we have calculated the relative response to selection for lamb survival compared to other traits (Table 5), but have treated it as a threshold character with an underlying scale distributed normally and the deviation of the threshold from the mean being the selection differential (Falconer 1981).

Table 5. Relative response per generation for different traits

Trait	Heritability	CV%	Relative Response
[1]Clean Fleece Weight	0.42	16.3	100
[2]Lamb Survival (to weaning)	0.032	49.4	14

([1]Values from Fogarty *et al.* 2006. [2]Values from this study)

Despite the very low heritability, with high variation, predicted genetic improvement for lamb survival still manages to be 14% of the relative gains possible of that for clean fleece weight, a highly heritable trait. Notwithstanding, the rate of gain predicted is still slow and behoves researchers to find ways to improve it.

Correlated traits and indirect selection. Based on the negligible genetic correlations between birth weight and lamb survival, there appears to be little to be gained in lamb survival from genetic manipulation of birth weight. Our finding that there was only a very slight advantage in survival for progeny of lambs with hairier birth coats (especially to 3 days of age) is similar to the results of a study on the SAMRF project data only (Ponzoni *et al.* 1997). However, data from a range of environments and lambing times is needed before industry recommendations are made about the usefulness of birth coat score as a selection criterion for improving lamb survival.

Finally, with reproductive wastage in sheep flocks attracting increased attention in recent years due to animal welfare and economic considerations, sheep breeders are looking for viable options to improve lamb survival, including genetic ones despite the predictions of slow genetic progress. Increasing potential genetic gain via improving accuracy by identifying a range of useful indirect selection criteria for lamb survival (Brien *et al.* 2009) and by progeny testing of industry sires are approaches currently under investigation by the Sheep CRC. We also note that breeding values for lamb survival are available from Sheep Improvement Limited in New Zealand (Newman 2003) and that useful genetic gains in lamb survival have been made in a South African flock selected for improved maternal ability (Cloete *et al.* 2009).

ACKNOWLEDGEMENTS

Funding was provided by the Sheep CRC. The SA Merino Resource Flock and SA Selection Demonstration Flocks received major funding support from Australian Wool Innovation and its predecessors, SARDI and the former South Australian Department of Agriculture. We acknowledge Raul Ponzoni as the founder of the Selection Demonstration Flocks. Technical advice from Alex Safari is also gratefully acknowledged.

REFERENCES

Brien, F.D., Hebart, M.L., Hocking-Edwards, J.E., Greeff, J.C., Hart, K.W., Refshauge, G., Gaunt, G., Behrendt, R., Thomson, K., Hinch, G.N., Geenty, K.G. and van der Werf, J.H.J. (2009) *Proc. Assoc. Advmt. Anim. Breed. Genet.* **18***:*108

Cloete, S.W.P., Misztal, I. and Olivier, J.J. (2009) *Proc. Assoc. Advmt. Anim. Breed. Genet.* **18**:104

Falconer, D.S. (1981) Introduction to Quantitative Genetics. Second Edition. (Published by Longman, London and New York).

Fogarty, N.M., Safari, E., Gilmour, A.R., Ingham, V.M., Atkins, K.D., Mortimer, S.I., Swan, A.A., Brien, F.D. and van der Werf, J.H.J. (2006) *Proc.2006 Aust.Sheep.Ind. Conf., Orange, Australia* pp 36.

Gifford, D.R., Ponzoni, R.W., Walkley, J.R.W., Hynd, P.I. and Ancell, P.M.C. (1990) *Int.J.Sheep.Wool.Sci.* **40**:114.

Gilmour, A.R., Gogel, B.J., Cullis, B.R. and Thompson, R. (2006) 'ASReml User Guide Release 2.0.' (VSN International Ltd: Hemel Hempstead).

Hatcher, S., Safari, A. and Atkins, K.D. (2009) *J.Anim.Sci. (Submitted for publication).*

Kemper, K.E., Hebart, M.L., Brien, F.D., Jaensch, K.S., Smith, D.H. and Grimson, R.J. (2006) *Proc. 8th Wld. Congr. Genet. Appl. Livest. Prod., Brazil* CD-Rom Communication 05-05.

Newman, S-A.N. (2003) *Proc. N.Z. Soc. An. Prod.* **63:**194.

Obst, J.M. and Day, N.R. (1968) *Proc.Aust.Soc.Anim.Prod.* **7:**239.

Ponzoni, R.W., Grimson, R.J., Jaensch, K.S., Smith, D.H. and Hynd, P.I. (1997) *Int.J.Sheep.Wool.Sci.* **45**:12.

Safari, E., Fogarty, N.M. and Gilmour, A.R. (2005) *Livest.Prod.Sci.* **92**:271.

MEAT QUALITY IN MERINO RAM HOGGETS

S. F. Walkom[1], F.D. Brien[2], M.L. Hebart[2], J.C. Greeff[3], D.L. Hopkins[4] and W.S. Pitchford[1]

[1]University of Adelaide, Roseworthy, SA, 5371
[2]South Australian Research and Development Institute, Roseworthy, SA, 5371
[2]Department of Agriculture and Food Western Australia, South Perth, WA, 6151
[3]New South Wales Department of Primary Industries, Cowra, NSW, 2794

SUMMARY

Genetic, environmental and management effects on meat pH and colour and their relationship with other production traits were estimated from pooled data based on carcase measurements from 5870 Merino hogget rams from New South Wales, South Australia and Western Australia. Principal component analyses failed to generate a trait which would be superior to the direct use of meat pH and colour as criteria in selection against dark cutting meat. No genetic or environmental links were found between dark cutting and production traits that could provide a causative effect or potential for indirect selection. In addition to parameter estimates, the contribution of various sources of variation for meat pH and colour have been quantified and the implications of the findings are discussed.

INTRODUCTION

Australia's sheep industry differs from most other countries in that the majority of sheep are either purebred or derived from the Merino, a specialist wool breed. Currently, 33% of lambs slaughtered for human consumption are second cross (25% Merino), 42% first cross (50% Merino) and 22% are purebred Merino. The Merino has a greater prevalence of high pH meat compared to more traditional meat breeds (Hopkins and Fogarty 1998; Gardner *et al.* 1999), so high pH is an important meat quality issue in Australia, as it is a predisposing factor in dark cutting meat.

There are two major hypotheses for the increased prevalence of dark cutting meat in Merinos. The first is associated with the Merino's greater response to stressors. It suggests this heightened response occurs from the farm to slaughter and especially during lairage, resulting in greater glycogen depletion, a lower level of lactic acid produced and a higher pH (Gardner *et al.* 1999; Warner *et al.* 2006). The second theory is that breeding for a specialist wool producing animal has unintentionally resulted in a decline in meat based traits and that this has lead to an increase in dark cutting meat (Warner *et al*. 2006).

The aim of this study was to obtain a better understanding of the high pH condition within the Merino breed. The first objective was to test if the pH and colour traits could be combined into a single trait that provides a stronger association with the high pH condition. The second objective was to quantify sources of variation within meat quality traits; how much of this variation can be accounted for and therefore what are the driving forces behind the high pH condition.

METHODS

Recently, Greeff *et al.* (2008) reported on the largest meat quality data set analysed to date in Merino flocks within Australia. This same data set was utilised in the current study. The data source comprised 5870 animals from three research locations in southern Australia. The Merino hogget rams slaughtered were approximately 18 months of age. A total of 543 sires and 4284 dams were used across 3 research flocks which each had their own selection guidelines (Safari *et al.* 2007). No sires were used across flocks. The research flocks were based at Trangie (New South Wales), Turretfield (South Australia) and Katanning (Western Australia).

Ultimate pH was measured 24hrs post mortem on the loin at the 12th rib. In South Australia this measurement was taken at 48hrs as there was no electrical stimulation. Meat colour was measured on the cut surface of the loin muscle at the 12th rib after 30 minutes of exposure to air allowing the meat to 'bloom'. Measurements were taken with a Minolta Colorimeter (Konica Minolta Sensing Inc., Osaka, Japan) using the Commission Internationale deÉclairage laboratory colour system; CIELAB. Brightness (L*) is a measure of white light reflected from the meat surface whilst redness (a*) and yellowness (b*) refer to the ratios of red to blue light and yellow to green light respectively (Warner *et al.* 2006).

Principal Component (PC) Analysis. PC analysis was used with the aim of producing an independent descriptor trait of the relationship between pH and meat colour. This method uses the correlation matrices between variables to deduce the latent variables describing the relationships in a more economical way. The new components are then described by the loadings of the feeder traits, in this case pH and meat colour (Jobson 1992). The loading values describe the proportion that each trait accounts for and the direction they are acting within the PC. The benefit of this process is observed in highly correlated traits which cluster together and are affected by common factors which can therefore be described by a single trait. For a more in-depth discussion of PC analysis see Jobson (1992) and Forkman *et al.* (1995). The data were analysed using ASREML (Gilmour *et al.* 2006) with location, year and breeding line within location and year fitted as random effects. Effects for type of birth and rearing, age of dam and date of birth accounted for zero variance in meat colour and were removed from the model. Hot carcass weight and GR fat at the 12 rib were also fitted as covariates but accounted for zero variance and were removed for both pH and colour.

RESULTS

At all 3 locations, a high proportion of animals produced high pH meat characterised by a pH of 5.8 or higher (NSW 63%; SA 90%; WA 74%).

Principal Component Analysis. Initially PCs were formed from four traits; pH, brightness (L*), redness (a*) and yellowness (b*). The first PC accounted for 67% of the variation but the secondary colour traits a* and b* had a large influence on the latter PCs despite being of relatively low importance in actually quantifying the problem of high pH meat. Meat brightness was the key driving factor in tonal changes within the meat colour chips used in beef meat quality analysis (Walkom, unpublished). Thus, it was decided to form a PC solely on pH and brightness. The PC analysis was possible because pH and brightness were moderately correlated ($r_P = -0.41$). The first PC was significant (latent root 1.41) and accounted for 70% of the variation in the two traits. The PC in this case described meat colour with positive values associated with brighter, low pH meat. By chance, the colour threshold measurement of 34 (Hopkins and Fogarty 1998) corresponded to a PC value of zero.

Fixed Effects. The PC was influenced by location, year and genetic line nested within years. This was also the case for pH and brightness although the influence of line on the PC appears to be mainly driven by the brightness value. The model was only able to account for 36-45% of the variation within pH, brightness and the PC with the genetic variation accounting for less than 20% of the total variation (Figure 1), leaving approximately 60% of the variation unaccounted for. Variation within location, year and breeding lines did not coincide with any trends in prime lamb production traits.

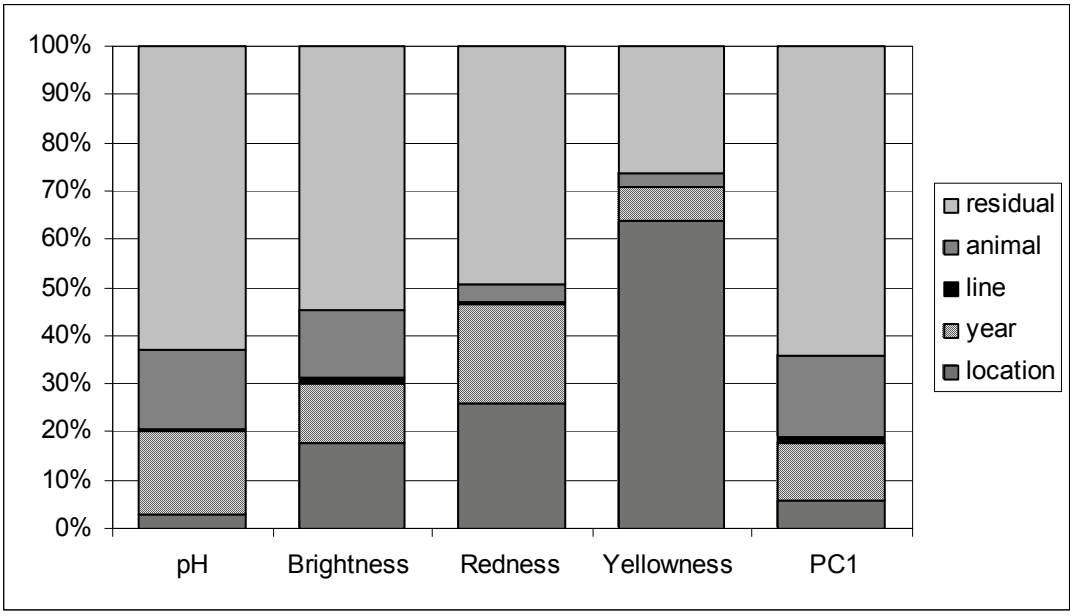


Figure 1. Analysis of sources of variation on meat quality. The proportion accounted for is shown on the vertical axis in percentage from. The traits analysed were, ultimate pH, the meat colour traits brightness (L*), redness (a*) and yellowness (b*) and the principal component combining pH and brightness (PC1).

DISCUSSION

Principal component. Whilst the PC was of value to the analysis and provided an ideal description of the relationship between the original traits, it was not as useful as a stand alone trait. Selection on the PC threshold within industry would be difficult due to its dependence on the size of the data source and level of error within the source. The heritabilities of pH and brightness were low to moderate (0.20 and 0.21 respectively, Greeff *et al.* (2008) with the heritability of the PC no better at 0.22. The moderate heritability suggests potential for improvement by selection directly for pH or colour. A genetic correlation of -0.57 between pH and brightness (Greeff *et al.* 2008) will assist with improving the meats appearance along with the selection against high pH meat. With a selection intensity of 1 and a generation interval of 3 years and single trait selection against dark cutting for 10 years, ultimate pH ($h^2 = 0.21$, $\sigma_P = 0.38$) could be improved by 0.26 units (from 6.05 to 5.79) or brightness ($h^2 = 0.20$, $\sigma_P = 3.35$) could be improved by 2.2 units (from 33.9 to 36.1). These calculations suggest the majority of animals could be within the acceptable industry threshold after focussed selection for approximately 10 years. Due to the need to include other traits within the breeding system and that meat quality traits are measured on the carcass, selection against high pH meat would take far longer than the direct selection calculation suggests.

Factors. The statistical model only accounted for 37% (pH) and 45% (brightness) of total observed variation within the flocks and the level of genetic variation was small in both key traits, so approximately 60% of the variation was unaccounted for. This residual variation would include non-additive genetic effects, technical error within machine measurements as well as other biological variation such as behaviour associated with stress response and the physiological

differences between animals associated with glycogen storage and breakdown; these factors were not measured in this study. A higher proportion of high pH meat is expected in Merino hogget rams although there are some reports of high levels (68%) in wethers (Ferguson *et al.* 2008).

The major factors accounting for variation were location and year effects. There are a number of variables that are confounded with location and year effects that can not be explained within the data available; the largest of these being genotypic strain and nutrition. The genetic background of the flocks at each site varied greatly although the influence of line on high pH levels is unknown.

Variation in animal nutrition between sites and years was quite large, however long term nutrition (as defined by fat depth) did not affect high pH within the flocks. This is supported by Sañudo *et al.* (2000) where fat covering of commercial lambs had no significant influence on pH or brightness and is probably expected for ruminants. The period prior to haulage and slaughter may provide the greatest opportunity to prevent high pH in Merinos, as the animals can not mobilize fat fast enough during these times to restore glycogen stores. The month prior to slaughter appears to be the most critical phase of the animals' nutritional regime for ensuring meat quality, as Ferguson *et al.* (2008) found that Merino wethers fed a concentrate diet of 11.4 ME/kg over a 34 day period prior to slaughter only had a 9% incidence of dark cutting compared to 68% for animals feed a roughage diet of 8 ME/kg. Animals at Trangie were fed pellets before slaughter and Katanning animals were slaughtered off green pasture in an attempt to lift muscle glycogen levels however the influence of this is unknown.

Genetic improvement through selection against high pH in Merinos provides some hope to improve meat quality due to a moderate heritability. However, other studies (Ferguson *et al.* 2008) indicate a greater potential for improvement by maintaining high muscle glycogen concentration through better nutrition prior to lairage and slaughter.

ACKNOWLEDGEMENTS

Funding for this study was provided by the Commonwealth Government through the Australian Sheep Industry Cooperative Research Centre and by Meat and Livestock Australia. We thank the many other scientists, technical and support staff from the NSW Department of Primary Industries, the South Australian Research and Development Institute and the Department of Agriculture and Food Western Australia, who have contributed to flock management and data collection over several years. Contributions from sheep breeders and Australian Wool Innovation Ltd for their support in the establishment of the research flocks is also gratefully acknowledged.

REFERENCES

Ferguson, D.M., Daly, B., Gardner, G.E. and Tume, R.K. (2008) *Meat Sci.* **78**:202.

Forkman, B., Furuhaug, I.L. and Jensen, P. (1995). *Appl. Anim. Behav. Sci.* **45**:31.

Gardner, G.E., Kennedy, L., Milton, J.T.B. and Pethick, D.W. (1999) *Aus. J. Ag. Res.* **50**:175.

Gilmour, A.R., Gogel, B.J., Cullis, B.R. and Thompson, R. (2006) 'ASReml User Guide Release 2.0.' (VSN International Ltd: Hemel Hempstead).

Greeff, J., Safari, E., Fogarty, N.M., Hopkins, D.L., Brien, F.D., Atkins, K.D., Mortimer, S.I. and van der Werf, J.H.J. (2008) *J. Anim. Breed. Genet.***125**:205.

Hopkins, D.L. and Fogarty, N.M. (1998) *Meat Sci.* **49**:477.

Jobson, J.D. (1992) 'Applied Multivariate Data Analysis VII.' (Springer-Verlag: New York).

Safari, E., Fogarty, N.M., Gilmour, A.R., Atkins, K.D., Mortimer, S.I., Swan, A.A., Brien, F.D., Greeff, J.C. and van der Werf, J.H.J. (2007) *Aus. J. Ag. Res.* **58**:169.

Sañudo, C., Alfonso, M., Sánchez, A., Delfa, R. and Teixeira, A. (2000) *Meat Sci.* **56**:89.

Warner, R.D., Dunshea, F.R., Ponnampalam, E.N., Ferguson, D., Gardner, G.E., Martin, K.M., Salvatore, L., Hopkins, D.L. and Pethick, D.W. (2006). *Int. J. Sheep and Wool Sci.* **54**:48.

GENETICS OF WOOL COLOUR IN THE SOUTH AUSTRALIAN SELECTION DEMONSTRATION FLOCKS

M.L. Hebart and F.D. Brien

South Australian Research and Development Institute, Roseworthy, SA, 5371

SUMMARY

A wide range of objectively measured and subjectively assessed characters were recorded in the South Australian Selection Demonstration Flocks between 1997 and 2005. These data provided an opportunity to study the genetics of wool colour and relationships with other traits, as part of a wider study by the Sheep CRC. In 2000, 2001 and 2002 clean wool colour was objectively measured by the Australian Wool Testing Authority on wool samples from 1,116 hogget rams. Clean colour was highly heritable (0.45). For assessed greasy colour the heritability was even higher (0.53), with a favourable negative genetic correlation (-0.38) between the two traits. There was a strong (0.57) positive genetic correlation between mean fibre diameter and clean wool colour and similarly a moderate negative genetic correlation with greasy wool colour. There was also a positive genetic correlation between clean colour and greasy fleece weight (0.34), dust penetration (0.35), and the standard deviation of fibre diameter (0.33). This indicates that selecting animals with heavier fleece weights tends to lead to progeny with wool that is more yellow in clean colour and that selecting animals with lower fibre diameter tends to lead to progeny that have whiter, more lustrous wool. The implications of these findings are discussed.

INTRODUCTION

A current challenge for the Australian wool industry is to produce a product that can compete in the light weight, trans-seasonal, next-to-skin capable market. This requires whiter, more photostable wool. Currently wool requires the application of oxidative bleaches and fluorescent whiteners to bring it to an acceptable white colour that is comparable with its competitor, cotton, but unfortunately these treatments have a detrimental effect on photostability, placing wool at a distinct disadvantage (Millington *et al.* 2008). Unbleached cotton measures approximately 7 tristimulus units (Y-Z), where lower tristimulus units equal whiter colour. It has been estimated that between 1 – 3% of the Australian wool clip currently measures 7 tristimulus units (Y-Z) or lower, with 70 – 80% greater than 8.5. Consequently, there is economic merit in improving clean wool colour (Millington, unpublished; Millington *et al.* 2008).

Genetic improvement through direct or indirect selection is one method of improving clean wool colour. Currently greasy wool colour, a component of style is estimated as a subjective measure of wool colour rather than objectively measured wool colour. There are many genetic parameter estimates of greasy wool colour (most recently reviewed by Mortimer 2007). However the relationship between greasy and clean wool colour is not clear, with the most recent estimates indicating a poor genetic correlation between the two traits. The aim of this study, as part of a wider Sheep CRC project (see also Smith and Purvis 2009) was to estimate the relationship between clean and greasy wool colour and their relationship with other important wool traits.

MATERIALS AND METHODS

SA Selection Demonstration Flocks Project (SDFs). In 1996, SA Merino ewes were sourced from a previous trial conducted at Turretfield Research Centre (the SA Merino Resource Flock, Gifford *et al.* 1990) to establish four flocks of 200 ewes each, representing three major selection

approaches and a randomly selected control. In 1999, a Meat Merino line was added. All SDF lambs were born in June-July of each year. For more details, see Kemper *et al.* (2006).

A wide range of objectively measured and subjectively assessed characters were recorded on the SDFs between 1997 and 2005. For this paper the objectively measured traits analysed included greasy fleece weight (GFW), clean scoured yield (CSY), clean fleece weight (CFW), mean fibre diameter (MFD), standard deviation of fibre diameter (SDFD), coefficient of variation of fibre diameter (CVFD), staple length (SL), staple strength (SS), mean fibre curvature (MFC), body weight (BWT) and dust penetration (DUST) recorded at 16 months of age on both sexes.

In 2000, 2001 and 2002 clean wool colour (CCOL) was objectively measured on 1,116 hogget rams in tristimulus units (Y-Z), where lower values equal whiter colour (Table 1). The subjectively measured traits reported were also measured at 16 months and include handle (HAND, 1 harsh – 5 very soft) and greasy colour (GCOL, 1 yellow – 4 lustrous white). Table 2 indicates the numbers of animals with each of these traits recorded.

Table 1. Sire representation among clean wool colour (CCOL) records

	Total records[#]	CCOL records	Total sires represented	Progeny group size	No. sires with >9 progeny & Ccol
Hoggets	9436	1116	124	1-168	81

[#] fully pedigreed

The initial data analyses fitted a univariate animal model to determine the main effects for the subsequent multivariate analyses using ASReml (Gilmour *et al.* 2006). The model included the fixed effects of flock (SDF, 1-5), drop (1997-2005), age of dam (2 – 6 years of age), type of birth and rearing (born as a single and raised as a single, born as a multiple and raised as a single, born as a multiple and raised as a multiple) and age to first shearing fitted as a covariate. For the subjectively measured traits that were measured by more than one classer, classer was added as a fixed effect to the model. The final models used for the multivariate analyses were determined by iteratively removing any non-significant terms from the univariate model. Variance and covariance components were used to estimate heritability and phenotypic and genetic correlations. While CCOL was normally distributed, GCOL was slightly right-skewed. However, no data transformation was undertaken.

Table 2. Summary statistics for (16mth) hogget traits in the Selection Demonstration Flocks

Trait	Mean	SD	Min.	Max.	Count
GFW (kg)	6.6	1.2	2.1	11.5	7,066
CSY (%)	71.9	5.5	50.8	85.0	7,103
CFW (kg)	4.8	1.0	1.6	8.8	7,055
MFD (µm)	20.1	2.0	14.1	29.2	7,103
SDFD (µm)	4.5	0.7	2.7	8.3	7,103
CVD (%)	22.1	2.9	13.5	34.8	7,103
SL (mm)	100.0	15.6	56.7	157.1	7,095
SS (N/kTex)	30.7	11.9	3.9	98.1	7,092
MFC (°/mm)	83.7	10.7	52.6	126.1	7,103
CCOL (Y-Z)	9.6	0.7	7.5	14.4	1,116
BWT (kg)	59.2	11.4	19.5	96.4	7,068
HAND (1-5)	3.0	0.9	1.0	5.0	4,162
DUST (1-5)	5.1	1.1	0.5	9.5	5,601
GCOL (1-5)	2.5	0.8	1.0	5.0	4,162

RESULTS AND DISCUSSION

Fixed Effects. The SDFs were significantly different for both greasy and clean wool colour (Table 3), with wool from the control flock being significantly more yellow than the other flocks for both traits. There was a significant effect of drop on both traits with those animals born in the year 2000 producing more lustrous, white wool than those born in 2001 and 2002. Interestingly animals born and reared as multiples tended to have more yellow greasy wool than those reared as singles (and born as either singles or multiples). This result was not reflected in the objectively measured wool colour (CCOL) with no significant difference between birth and rearing types.

Table 3. Significance of fixed effects on yearling and adult clean and greasy wool colour

Trait	Dam age	Birth-Rear type	Sex	Flock	Age	Classer	Drop
CCOL (Y-Z)	n.s.	n.s.	-	***	*	-	***
GCOL (1-5)	n.s.	***	**	***	***	**	***

*** P<0.001, ** P<0.01, * P<0.05, ns not significant

Table 4. Hogget (16mth) heritability (h^2) and phenotypic (r_p) and genetic (r_g) correlations among clean scoured wool colour (CCOL) and assessed greasy wool colour (GCOL) and other production traits with s.e's in parentheses

Trait	h^2	CCOL		GCOL	
		r_p	r_g	r_p	r_g
GFW (kg)	0.46 (0.03)	0.21 (0.03)	0.34 (0.09)	-0.15 (0.02)	-0.36 (0.04)
CSY (%)	0.52 (0.03)	-0.07 (0.03)	-0.13 (0.10)	0.29 (0.01)	0.56 (0.03)
CFW (kg)	0.38 (0.03)	0.10 (0.03)	0.20 (0.10)	0.01 (0.02)	-0.03 (0.04)
MFD (µm)	0.63 (0.03)	0.32 (0.03)	0.57 (0.08)	-0.17 (0.02)	-0.36 (0.03)
SDFD (µm)	0.56 (0.02)	0.20 (0.03)	0.33 (0.09)	-0.26 (0.02)	-0.47 (0.03)
CVD (%)	0.50 (0.03)	0.02 (0.03)	-0.03 (0.10)	-0.17 (0.02)	-0.30 (0.03)
SL (mm)	0.49 (0.03)	-0.04 (0.03)	0.05 (0.10)	0.07 (0.02)	0.05 (0.04)
SS (N/kTex)	0.27 (0.02)	0.14 (0.03)	0.21 (0.11)	0.03 (0.02)	0.07 (0.05)
MFC (º/mm)	0.62 (0.02)	-0.23 (0.03)	-0.35 (0.09)	-0.01 (0.02)	0.01 (0.03)
CCOL (Y-Z)	0.45 (0.08)			-0.15 (0.03)	-0.38 (0.07)
BWT (kg)	0.48 (0.03)	0.01 (0.03)	0.02 (0.10)	-0.01 (0.02)	-0.06 (0.04)
HAND (1-5)	0.46 (0.01)	-0.12 (0.03)	-0.31 (0.08)	0.48 (0.01)	0.79 (0.02)
DUST (1-5)	0.25 (0.03)	0.12 (0.03)	0.35 (0.12)	0.02 (0.02)	-0.10 (0.06)
GCOL (1-5)	0.53 (0.01)	-0.15 (0.03)	-0.38 (0.07)		

Genetic parameters. The estimates of phenotypic variance and heritability and genetic correlations for most of the hogget production traits in table 4 are in agreement with those already published on this dataset (Safari *et al.* 2007, 2008). Clean colour was highly heritable (0.45) indicating that genetic improvement for clean colour can be made through selection. For assessed greasy colour the heritability was even higher (0.53), with a favourable negative genetic correlation (-0.38) between the two traits. These estimates are similar to those of James *et al.* (1990) who reported heritability estimates of 0.42 for both CCOL and GCOL and a high favourable genetic correlation of 0.87 (GCOL was scored in the opposite direction to our study) between the two traits in Collinsville Merino sheep. However in fine wool Merinos, Smith and Purvis (2009) estimated low genetic correlations between the two colour measures of 0.13 and 0.l8 in yearling and adult sheep respectively, where GCOL was also scored in reverse to our study.

Interestingly, there was a strong (0.57) positive genetic correlation between mean fibre diameter and clean wool colour and similarly a moderate negative genetic correlation with greasy wool colour. There was also a positive genetic correlation between clean colour and GFW (0.34), DUST (0.35), and SDFD (0.33). This indicates that animals with heavier fleece weights would also have hogget progeny with wool that has a yellower clean colour. Two other genetic correlations of note between clean colour and key productions traits, although only slight, are those with CFW (0.20) and SS (0.21). These values are slightly higher than those estimated by Smith and Purvis (2009) in fine wool Merinos at both yearling and adult time points.

In conclusion, our results suggest that either direct selection of hogget animals for improved clean wool colour or indirect selection based on greasy wool colour will lead to genetic improvement in clean wool colour in hogget animals. The consequences of such selection for hogget production traits likely include a slight reduction in hogget fleece weight, a lower fibre diameter and a slight reduction in staple strength. Whilst these genetic parameters for hogget performance are an important step to the implementation of selection for improved wool colour in commercial breeding programs, so too are genetic parameters linking hogget and adult production and the development of economic values (Smith and Purvis 2009). Finally, current breeding programs placing strong emphasis on improving fleece weight (but not fibre diameter), unless counter-balanced by other selection efforts, may unintentionally lead to increases in wool colour. This may not be an issue for traditional end uses for wool, but would limit the availability of suitable wools for supplying a developing market in light weight, trans-seasonal, next-to-skin wear.

ACKNOWLEDGEMENTS

Funding was provided by the Sheep CRC. The SA Merino Resource Flock received major funding support from Australian Wool Innovation and its predecessors and the South Australian Research and Development Institute. We acknowledge Raul Ponzoni as the founder of the Selection Demonstration Flocks and the contribution of a number of current and former SARDI staff members, particularly Kaylene Jaensch, Darryl Smith, Richard Grimson, Anne Ramsay and Peter James.

REFERENCES

Gifford, D.R., Ponzoni, R.W., Walkley, J.R.W., Hynd, P.I. and Ancell, P.M.C. (1990). *Int. J. Sheep. Wool. Sci.* **40**:114.

Gilmour, A.R., Gogel, B.J., Cullis, B.R., and Thompson, R. (2006). "ASReml User Guide Release 2.0" VSN International Ltd, Hemel Hempstead

James, P., Ponzoni, R.W., Walkley, J.R.W., and Whiteley, K.J. (1990) *Aust. J. Agric. Res.* **41:**583.

Millington K.R., Dyer J.M., Fleet M.R., Mahar T.J., Smith J. and Swan P. (2008). The 86th Textile Institute World Conference: Fashion and Textiles: Heading for New Horizons. 18-21 November, Hong Kong.

Kemper, K.E., Hebart, M.L., Brien, F.D., Jaensch, K.S., Smith, D.H. and Grimson, R.J. (2006). *Proc. 8th Wld. Congr. Genet. Appl. Livest. Prod., Brazil* CD-Rom Communication 05-05.

Mortimer S.I. (2007). *Int. J. Sheep Wool Sci.* **55**, Article 6.

Safari, E., Fogarty, N.M., Gilmour, A.R., Atkins, K.D., Mortimer, S.I., Swan, A.A., Brien, F.D., Greeff, J.C. and van der Werf, J.H.J. (2007) *Aust. J. Agric. Res.* **58**, 177.

Safari, E., Fogarty, N.M., Gilmour, A.R., Atkins, K.D., Mortimer, S.I., Swan, A.A., Brien, F.D., Greeff, J.C. and van der Werf, J.H.J. (2008). *J. Anim. Breed. Genet.* **124:**65.

Smith, J.L. and Purvis, I.W. (2009). *Proc. Assoc. Advmt. Anim. Breed. Genet*. **18**:390.

VARIATION IN SOW HEALTH AFFECTS THE INFORMATION PROVIDED BY LACTATION FEED INTAKE DATA

K.L. Bunter[1], C.R.G. Lewis[1] and B.G. Luxford[2]

[1]Animal Genetics and Breeding Unit*, University of New England, Armidale, NSW 2351
[2]QAF Meat Industries Pty Ltd, Corowa, NSW 2646

SUMMARY

Using data from two maternal lines of pigs (N~2200), medication events were used as proxy indicators of sow health, to examine changes in associations between lactation feed intake and other traits that occur with changes to sow health status. Estimates of heritability for total born, average piglet birth weight, litter gain until day 10 (LG10), average sow feed intake during lactation (LFI), total sow feed intake during the first three days of lactation, lactation length and the underlying liability for a shortened lactation (SL) or surviving to farrow in parity 2 (FP2) were 0.14±0.01, 0.33±0.03, 0.09±0.04, 0.18±0.04, 0.06±0.04, 0.06±0.03, 0.15±0.09 and 0.06±0.07. Genetic (r_a) and phenotypic (r_p) correlations indicate that high lactation feed intake was favourably associated with SL and FP2 (r_a: -0.78±0.19 and 0.42±0.41; r_p: -0.49±0.01 and 0.31±0.01). Compared to estimates obtained using only subsets of data from sows that met lactation length targets or unmedicated sows, heritabilities for LFI were higher in the medicated data set and phenotypic correlations with LG10, SL or FP2 were of increasingly larger magnitude across these data subsets. Sow health status affects the information content of lactation feed intake data, but larger studies will be required to confirm if significant changes also occur in genetic parameters because of health status. Knowledge of health status could be important for other studies which examine associations between feed intake, recorded in any physiological state, and other production traits.

INTRODUCTION

Studies have shown that sows with higher lactation feed intake rear heavier piglets, have better body condition at weaning, and are more likely to farrow in the subsequent parity (Hermesch and Jones 2007; Bergmsa *et al*. 2008), but knowledge of the genetic associations between lactation feed intake and sow performance is relatively limited. In addition, where lactation feed intake data are available the health status is typically unknown. It is therefore unclear how much phenotypic or genetic associations with feed intake are influenced by prevailing acute health issues, which affect feed intake, compared to variation in feed intake in normal healthy sows. In this study we use medication events as proxy indicators of sow health to create data subsets for lactation records, to examine changes in parameters that occur with changes to sow health status.

MATERIALS AND METHODS

Approximately 2200 sows from two maternal lines (Primegro™ Genetics, Large White and Landrace based) were recorded for their first gestation and farrowing outcomes between January 2007 and June 2008. Data in this study included total number of piglets born (TB), average piglet birth weight (ABW) of live born piglets, along with litter gain from day 1 (after cross fostering) until day 10 (LG10), the average daily lactation feed intake of the sow (LFI) up to 35 days, total feed intake in the first three days of lactation (FID3), and the lactation length of the sow (LACL). Events such as a shortened primiparous lactation (SL) and farrowing in parity 2 (FP2) were treated as binary traits (0/1, where event=1) for uncensored sows. Continuous data were subsequently

* AGBU is a joint venture of the NSW Department of Primary Industries and the University of New England

edited based on trait distributions. Proc UNIVARIATE (SAS, 2003, SAS Institute Inc. Cary, N.C.) was used to identify outliers, whereby trait records that deviated by more than 3× the interquartile range from the mean value were deleted. After editing, there were data from 2164 animals representing 197 sires and 1221 dams. The pedigree was extended back to include all animals born since 2003 (N=44950).

Subsets of the complete data (ALL) for feed intake records (LFI, LACL, FID3) were then developed based on LACL and specific health information; data for other traits remained constant. The targeted lactation length and weaning age in this data was 28-32 days, with a maximum weaning age of 35 days. Subset 1 contained sows with LACL>21 (N=1778); subset 2 contained sows with no evidence of medication events during lactation (N=1030); and subset 3 contained sows that received any non-feed medication within 5 days prior to farrowing, until 35 days post-farrowing (N=1008). As with other species, primiparous sows frequently have more periparturient and post-farrowing difficulties than multi-parous sows, and several medication categories were grouped together over this time period. The heritability estimate for a medication event in this data was zero.

Models for analyses were developed using ASReml software (Gilmour *et al*. 2006). Univariate analyses were used to identify significant fixed effects and to obtain initial estimates of genetic parameters under either an animal (continuous traits) or sire (binary traits) model. Parameter estimates for binary traits were estimated on the underlying scale using a logit link function. Systematic effects for all traits included year/month of farrowing (levels: 19) and sow line (2 levels). In contrast to Bunter *et al*. (2009) lactation length was excluded as a covariate for lactation intake in this study, so that lactation length could be examined as a trait. Correlations between specific traits were estimated in a series of bivariate analyses. Bivariate analyses of trait combinations involving any binary trait were all performed using a sire model for both traits.

RESULTS AND DISCUSSION

Characteristics of the data. Focusing on less common traits, feed intake in the first three days of lactation was more variable (CV=37%) than LFI (CV=22%) despite consistent 3× daily feed delivery schedules in the postfarrowing period (Table 1). Unmedicated sows ate significantly more feed per day in the first three days after farrowing and over the complete lactation than medicated sows, but lactation lengths were similar (Table 2). This is consistent with subjective observation that animals with health issues have reduced intake (Weary *et al*. 2009) but can also indicate that low feed intake was a cue for initiating treatment. Variability in lactation traits was also 17 to 25% higher in the medicated data set relative to the unmedicated data set. The incidence of shortened lactations in primiparous sows was 13%; many of these affected sows were not given an opportunity to re-mate. Seventy three percent of farrowed sows had a second parity. However, this is an underestimate as some sows were transferred after their first parity, and were generally confirmed pregnant at transfer.

Heritability estimates. The estimate of heritability for LFI (0.18±0.04) was similar to the estimate of 0.16±0.04 from Bunter *et al*. (2009) where LACL was used as a covariate for LFI, and comparable to other studies (Bergsma *et al.* 2008; Hermesch 2007). In this study, low estimates of heritability were also obtained for FID3 (0.06±0.04) and LACL (0.06±0.03), whereas lactation length was not heritable in the data of Bergsma *et al*. (2008) or Hermesch (pers. comm.). Heritabilities for LFI or FID3 were higher in the subset of medicated sows; genetic variation for appetite variation may be better expressed during lactation under more challenging environmental conditions (Table 2, subset 2 vs 3).

The frequency of progression to the second parity and shortened lactations were lowly to

moderately heritable (0.06±0.07 and 0.15±0.09) when estimated on the underlying liability scales, although standard errors of estimates were large. For comparison, heritability estimates for TB and sow survival to parity 2 were 0.13±0.02 and 0.05±0.02 in the study of Bergsma *et al.* (2008), the latter estimated on the observed scale from a population with an incidence of 85.1%. Genetic parameters for TB and ABW were consistent with averages from the review of Rothschild and Bidanel (1998).

Table 1. Characteristics of the data after editing for outliers, along with parameter estimates

Trait	Abbreviation	N	Mean (SD)	Min-Max	h^2	σ_p
Total born (N)	TB	2163	11.6 (3.20)	2-21	0.14±0.01	2.82
Average piglet birth weight (kg)	ABW	2099	1.41 (0.24)	0.63-2.42	0.33±0.03	0.24
Lactation feed intake (kg/day)	LFI	2031	4.99 (1.10)	0.50-9.00	0.18±0.04	0.94
Total feed intake in first 3 days (kg)	FID3	2020	11.4 (4.20)	0-27	0.06±0.04	3.73
Lactation length (days)	LACL	2038	27.3 (7.47)	0-35	0.06±0.03	7.35
Litter gain to 10 days (kg)	LG10	1885	10.0 (6.08)	-12.5 to 32.5	0.09±0.04	5.98
Shortened lactation (0/1, event=1)	SL	2047	13%	0 and 1	0.15±0.09	1.85
Farrowed in parity 2 (0/1, event=1)	FP2	2040	73%	0 and 1	0.06±0.07	1.83

Table 2. Characteristics and heritability estimates from data subsets, along with genetic (1st row) and phenotypic (2nd row) correlations ($se_{subscript}$) (all parameter estimates and se ×100)

Trait	LFI				FID3				LACL			
Subset	ALL	1	2	3	ALL	1	2	3	ALL	1	2	3
N	2031	1778	1029	1002	2020	1778	1021	999	2038	1778	1030	1008
$Mean_{SD}$	$4.99_{1.1}$	$5.19_{0.9}$	$5.16_{1.0}$	$4.81_{1.2}$	$11.4_{4.2}$	$11.6_{4.1}$	$12.3_{3.9}$	$10.4_{4.3}$	$27.3_{7.5}$	$29.9_{2.9}$	$27.9_{7.1}$	$26.6_{7.8}$
h^2	**18_4**	**9_4**	6_6	**23_8**	6_4	4_4	0_5	**19_8**	**6_3**	**11_4**	9_6	5_6
σ_p	0.94	0.72	0.74	0.80	3.73	3.63	3.49	3.90	7.35	2.66	6.97	7.70
					Correlations							
TB	-14_{13}	9_{19}	22_{22}	-26_{14}	-6_{20}	3_{24}	NE	-31_{17}	-46_{20}	-27_{16}	-21_{21}	-66_{37}
	3_2	**8_2**	4_3	3_3	0_2	1_2	-1_3	1_3	-3_2	2_2	-4_3	-2_3
ABW	**33_{12}**	**42_{18}**	**66_{24}**	19_{14}	-24_{20}	-46_{29}	NE	-17_{16}	**49_{19}**	**44_{16}**	**50_{22}**	43_{30}
	2_2	-1_2	2_3	1_3	**-10_2**	**-11_2**	**-8_3**	**-12_3**	**7_2**	5_3	**6_3**	**8_3**
LG10	**55_{18}**	4_{32}	18_{33}	**66_{24}**	16_{35}	-21_{45}	NE	21_{30}	**138_{34}**	**55_{25}**	**79_{27}**	NR
	34_2	**23_2**	**33_3**	**30_3**	**6_2**	**6_2**	2_3	**6_3**	**46_2**	**22_2**	**47_3**	**36_3**
SL	**-78_{19}**	-72_{36}	**-65_{31}**	-90_{23}	-52_{51}	-35_{72}	NE	-25_{42}	NE	NE	NE	NE
	-49_1	**-33_4**	**-38_2**	**-52_2**	**-11_1**	-1_4	**8_1**	**-10_2**	NE	NE	NE	NE
FP2	42_{41}	69_{73}	-55_{70}	**116_{47}**	39_{72}	30_{97}	NE	78_{61}	73_{40}	88_{55}	72_{69}	**122_{56}**
	31_1	**15_2**	**25_2**	**27_2**	**10_1**	**7_2**	**6_2**	**10_2**	**42_1**	**9_2**	**37_2**	**35_2**

Trait abbreviations are in Table 1. NE= Not estimated; NR: Not reported (se×100>100). Estimates are in bold if significantly (P<0.05) different to zero.

Correlations between traits. LFI had positive genetic correlations (r_a) with ABW, LG10 and FP2 (range 0.33 to 0.55) and the direction of correlations was generally similar across data subsets. Phenotypic correlations (r_p) between LFI and LG10 or FP2 were also moderate (around 0.30) supporting the positive associations typically observed between these traits. However, similar to Bergsma *et al*. (2008) a negative genetic correlation was estimated between LFI and FP2 in data from unmedicated sows. This outcome might arise under a positive relationship between feed intake and elevated maintenance requirements. In contrast, the estimate (r_a) over all data was positive although not significantly different to zero. Correlations (r_a and r_p) between LFI and SL

were moderately negative indicating sows that had high intake were less likely to have shortened lactations. The magnitude of correlations between LFI and SL or FP2 was largest in the medicated data set. The pattern of correlations between FID3 and LG10 or FP2 were similar to that for LFI, but lower magnitude. However, the phenotypic correlations between FID3 and SL were opposite in sign in data subsets 2 vs 3. Correlations between FID3 and ABW were negative indicating that sows farrowing heavier piglets ate less initially, despite genetic potential to consume more feed over the complete lactation. Correlations between LACL and ABW, LG10 and FP2 were consistently positive. High ABW and LG10 may promote initiation and continuation of lactation, by virtue of increased suckling stimulus of more robust piglets, and a longer lactation can assist with rebreeding performance (eg. see Tholen *et al*. 1996). Some estimates of genetic correlations involving LG10 or FP2 exceeded the upper boundary of the parameter space, predominantly in the medicated data subset.

Phenotypic correlations between LFI or LACL with LG10, SL or FP2 were of smaller magnitude in the data subset containing sows with lactation lengths longer than 21 days. This implies that the stronger phenotypic correlations observed in subsets 2 and 3 arise partly from partial lactation data for the 13% of sows which did not achieve the targeted lactation length. There were generally no significant correlations between TB and lactation traits, with the exception that for sows with normal lactation lengths (subset 1) the association between TB and LFI was low and positive. Bivariate analyses involving LACL and SL generally failed to converge, since SL was described as subsets of LACL.

CONCLUSIONS

Relative to unmedicated sows, LFI and FID3 of medicated sows was lower, more variable and more heritable, and correlations between feed intake traits and the prevalence of shortened lactations or survival to the second parity were of larger magnitude. Therefore, high lactation feed intake is relatively more important to sow and piglet survival in the presence of sow health issues, which was not a heritable trait in itself in this data. Standard errors of genetic correlations were relatively large such that only phenotypic correlations significantly differed between these data subsets. Recording individual health status should be considered important for studies involving feed intake; a larger study will provide more accurate estimates of genetic correlations under different health circumstances.

ACKNOWLEDGMENTS

This research was funded by the Pork CRC under projects 2D-101-0506 and 2E-104-0506. The authors are grateful for diligent data collection by Matthew Tull and staff at QAF Meat Industries.

REFERENCES

Bergsma, R., Kanis, E., Verstegen, M.W.A. and Knol, E.F. (2008) *J. Anim. Sci*. **86**:1067.

Bunter, K.L., Luxford, B.G., Smits, R., and Hermesch, S. (2009) *Proc. Assoc. Advmt. Anim. Breed. Genet.* **18**:203.

Hermesch, S. and Jones, R.M. (2007) *Proc. Aust. Pig Sci. Assoc.* **11**:196.

Hermesch, S. (2007) *Proc. Assoc. Advmt. Anim. Breed. Genet.* **17**:61.

Rothschild, M.F. and Bidanel, J.P. (1998) In "The Genetics of the Pig", p. 313, Eds. M.F. Rothschild and A. Ruvinsky, CAB International.

Gilmour, A.R., Cullis, B.R., Welham, S.J. and Thompson, R. (2006) "ASReml User Guide", Release 2.0 VSN International Ltd, Hemel Hempstead, HP1 1 ES, UK.

Tholen, E., Bunter, K.L., Hermesch, S. and Graser, H.-U. (1996) *Aust. J. Agric. Res*. **47**:1261.

Weary, D.M., Huzzey, J.M. and Keyserlingk, M.A.G. von (2009) *J. Anim. Sci.* **87**:770.

COMPLEXITIES AND STRATEGIES TO IDENTIFY THE CAUSATIVE MUTATION RESPONSIBLE FOR SINGLE LOCUS INHERITED DISEASES IN LIVESTOCK.

J. A. L. Cavanagh, I. Tammen and H. W. Raadsma.

Reprogen, Faculty of Veterinary Science, The University of Sydney, Camden, NSW 2570

SUMMARY

With the advent of high density SNP genotyping chips for livestock species, it is possible to quickly identify genomic regions of interest for well characterized inherited diseases given appropriate sample material. However, the search for causative mutations may be severely hampered if there is no functional candidate gene in the region, if the candidate genes are large, or if the disease causing mutation is not located in coding regions. Further complications are likely if a reference genome sequence is not available or incomplete. A quick scan of the coding region of a small number of candidate genes using traditional Sanger sequencing techniques can yield rapid results in amino acid changing mutations – either on genomic DNA or cDNA. If no coding mutations are found, a *de novo* search for disease causing mutations in genomic regions of interest is required. The use of next generation sequencing technologies in combination with targeted sequence capture is a recent development which so far eluded systematic searches studying the genetic architecture for simple monogenic and possibly complex traits.

INTRODUCTION

There is a worldwide interest in the identification of genes and causative mutations in those genes that cause single locus inherited diseases in livestock and companion animals. Whole genome sequencing has progressed rapidly for a number of livestock species and companion animals. From the whole genome sequencing effort, high density SNP genotyping arrays (chips) have been produced with unparalleled capacity to screen entire genomes with relative ease.

Inherited diseases are often disseminated widely with the use of popular sires. The aim is to characterize the mutation causing the defect and develop a direct DNA-test available for use by livestock industries to reduce the frequency of the disease allele and therefore the incidence of affected animals in the population. Direct and immediate benefits of a DNA-test are the reduction and minimization of losses caused by the defect and improvement of animal welfare. Indirect benefits pertain to Australia's reputation as a producer of high quality livestock with minimal incidence of major disease and genetic defects. In addition, the study of genetic defects may have medical significance as a large animal model for a homologous human disease.

There is an expectation that these genes and corresponding causative mutations would be relatively straightforward to map given the simple mode of inheritance, clear penetrance and a well defined phenotype. This is not always the case, although over the last few years, we have successfully identified causative mutations for Dexter cattle chondrodysplasia, Neuronal Ceroid Lipofucinosis (NCL) in Devon cattle and Merino sheep, and zinc deficiency in Angus cattle. We are currently implementing new strategies as discussed in this paper to search for causative mutations for NCL in South Hampshire sheep, Fawn Calf Syndrome (FCS) in Angus cattle and several other emerging defects in livestock. The challenge for all such diseases has not been the mapping or genome localization but rather the fine search for causative mutations underlying the phenotype. The advent of new large scale sequencing and screening technologies is likely to revolutionize the rate at which causative mutations are identified despite some pitfalls.

STRATEGY AND POTENTIAL PITFALLS

In the past, large scale microsatellite screening projects were employed to identify an area of the genome in linkage with the specific disease or phenotype of interest if no obvious candidate gene was identified. This was a time consuming and costly approach but relatively powerful given the specific location of a causative underlying genetic variant and a clear phenotype. Furthermore, in the case of recessive disorders, regions of homozygosity were expected within LD around the causative gene, thus allowing large scale comparison of affected, obligate carrier, and control animals. However, fine mapping to a single locus region remained problematic and in the absence of a suitable positional candidate gene, a challenge to map down to a single mutation. If functional candidate genes have been identified, homozygosity mapping can proceed in the targeted regions containing such candidate genes (Tammen *et al.* 1999). In one such case, the identification of the gene causing chondrodyplasia in Dexter cattle (Cavanagh *et al.* 2007), 11 functional candidate genes were selected and microsatellite markers targeted to these regions. Fortunately in this case, by way of comparative analysis across species, one of these candidate genes was later shown to be the target gene responsible, but in many cases the area of linkage identified is large (e.g. >3MB), spanning many genes and the causative mutation is not found. With the advent of SNP chips, identification of a region of interest has been greatly simplified. By sending away a sufficient number of informative DNA samples to a service laboratory (e.g Illumina bovine 54,000 SNP chip), a target region can be readily identified using a homozygosity mapping approach. A region of homozygosity amongst affected animals is sought whilst carrier and control animals are heterozygous (or not homozygous for the allele linked to the disease phenotype). The region of homozygosity can potentially be reduced by increasing the number of samples analysed and therefore mapping historical recombination events. The SNP chip strategy was employed in several studies (Charlier *et al.* 2008) as well as for mapping the homozygous region of FCS in Angus cattle to a region of 3.5Mb, which was further reduced with additional targeted SNP typing (unpublished data).

On identifying a genome region of interest a well annotated genome sequence either in the species of interest or via comparative approaches allows for identification of any functional candidate genes. Those diseases which are characterized biologically as phenotypes, enable similar diseases to be identified in human and mice, often with known causative genes. Additionally, the use of mice knockout models for each of the potential target genes, may allow for phenotype comparisons relevant to each gene. However, there are still many anonymous genes with unknown function, and it may not be possible to identify a likely candidate gene. One avenue to narrow down the list of relevant genes underlying a target region is through transcriptome analysis (gene expression studies) and looking for expression differences amongst affected and control animals. The usual constraints of timing of sampling and tissue specific expression prevent this approach from being definitive.

Traditional Sanger sequencing methods of positional candidate genes, can be used in amplified PCR products covering the coding region of a gene(s) of interest and sequenced for putative causative mutations. Again, this is time consuming and costly. This method was successful for several projects in our laboratory including Dexter cattle chondrodysplasia in which 2 causative mutations were identified in a single gene (Cavanagh *et al. 2007*), and single coding mutations in NCL in Devon cattle (Houweling *et al.* 2006), NCL in Merino sheep (Tammen *et al.* 2006), and Zinc deficiency in cattle (Tammen *et al.* 2002). All of these mutations were found in exons (coding and non-coding). With strong evidence for a positional candidate and high likelihood of an exonic mutation, this is still a preferred method of choice to map a mutation relatively quickly. However in the absence of a strong candidate gene, and a large region of homozygosity harbouring multiple genes, sequencing of all genes in a 3MB region has proven prohibitive until recently. With the advent of next generation sequencing technologies such as Roche 454, Illumina Solexa,

and AB SOLiD, a large volume (up to multiple Gigabases/sample) of sequencing data is obtainable. Each of these technologies have different strengths and weaknesses as discussed by Harismendy *et al.* (2009). There are several strategies such as whole genome re-sequencing, end sequencing of mate pair libraries, and sequence capture using arrays are of particular interest. To specifically target a region, a targeted DNA capture and sequence service, such as NimbleGen (http://www.nimblegen.com/products/seqcap/index.html), Febix (http://www.febit.com/go/en/services/hybselect/), Agilent (http://www.opengenomics.com/SureSelect) and LC Sciences (http://www.lcsciences.com/products/genomics/targeted_sequencing/targeted_sequencing.html) is required. In general the procedure involves fragmenting the DNA into small pieces and hybridizing it to a custom array containing probes to match the sequence of interest. The DNA of interest is then eluted and directly sequenced using a next generation sequencing platform. The major limitation is where no reference sequence is available to design capture probes. In this case a preliminary BAC sequencing step is required covering the target to generate denovo reference sequence.

DE NOVO MUTATION SCREENING/FINDING CAUSATIVE MUTATIONS

By implementing targeted sequence capture and next generation sequencing technology, the DNA sequence of control versus affected animals can be compared for differences. Not only does this reduce the emphasis on identification and characterization of positional candidate genes before hand, it also provides non-coding sequence information within the region of interest. In the case of FCS in Angus cattle and NCL in South Hampshire sheep, coding regions of each respective candidate gene were screened for mutations with a negative result. Now, with the use of targeted sequence capture and next generation sequencing, large amounts of DNA can be sequenced which will lead inevitably to the identification of many possible mutations, which in turn may lead to false positive direct DNA-tests. By carefully selecting animals to sequence, such as groups of full/half sibs (control/affected pairs), chance differences can be reduced. However, there will still be many mutations in perfect linkage disequilibrium with the disease and it may not yield direct information on the likely causative mutation. Furthermore, many elements in the genome may act as regulatory factors, copy number variants, repeat elements and changes in transcription factor binding sites or micro-RNAs. Such mutations are subtle and may not lead to an obvious causative role.

TOWARDS DEFINITIVE PROOF ON CAUSATIVE MUTATION

Simple single nucleotide polymorphisms (SNPs) can be causative mutations but difficult to substantiate if they are not directly related to predicted changes in the amino acid sequence. Mutations affecting coding sequence often mean that the RNA is subject to non-sense mediated RNA decay (Frischmeyer and Dietz 1999), resulting in little or nill of the mutant RNA being translated into protein. This is not the case for regulatory mutations that can affect the expression levels of the RNA and subsequent translated protein. Once a putative causative mutation has been identified, it is extremely difficult to prove without using allele substitution which may be done in appropriate cell lines, but is not generally done *in vivo*. There are many mouse knockout models which support a specific gene function, but these are generally not sensitive to define all possible mutations which could lead to the same phenotype and are therefore a blunt instrument. Targeted induction of the mutation in normal animals remains difficult but provides a powerful level of proof to a gold standard. Furthermore long range trans-acting regulatory elements, although rare, may also affect a phenotype and it will be difficult to identify these if not captured within the target sequencing strategy.

IMPLICATIONS FOR COMPLEX POLYGENIC TRAITS

Although the search for underlying causative mutations or QTN in polygenic traits is of interest, such searches will be an order of magnitude more difficult. Firstly with any population the strong LD which exists between markers surrounding the QTN will be extensive with many ambiguous SNP as possible QTN. Furthermore in the absence of definite proof of a gene let alone a specific mutation in a complex pathway, the power to attribute phenotypic changes to genes of relatively small effect is problematic. This is further hampered where compensatory or alternative pathways may act on the same phenotype expression such as in disease /immune response. As seen in studies with human height, a quantitative polygenic trait with a heritability of >80%, simply increasing the density of SNP to ultra high densities is not sufficient to account for all the additive genetic variance. Large scale sequencing of individual genomes will become a reality within the near future, but will only add to the complexities of dealing with far more explanatory variables than observations and most studies will fail to definitively map all causative mutations.

CONCLUSIONS

With the advent of new technologies, high density SNP screening, whole genome re-sequencing and targeted sequence capture, the characterisation of inherited diseases in livestock will become easier. Studies can proceed directly from whole genome SNP genotyping to targeted SNP genotyping panels if required, then to targeted sequence capture and sequence thus avoiding the steps of PCR amplification and Sanger sequencing. Diseases that have not yet had a causative mutation identified can be processed through the targeted sequence capture process. This method is also extremely efficient for characterising new or emerging inherited diseases in livestock. However, proof to a gold standard that the predicted causative mutation identified remains extremely difficult. Once predicted causative mutations are validated they can be used as a direct DNA-test to identify carrier animals and eliminate the deleterious defect in livestock populations and may be assembled as a single test for routine applications.

REFERENCES

Cavanagh, J.A., Tammen, I., Windsor, P.A., Bateman, J.F., Savarirayan, R., Nicholas, F.W. and Raadsma, H.W. (2007) *Mamm Genome*. **18**:808

Charlier, C., Coppieters, W., Rollin, F., Desmecht, D., Agerholm, J.S., Cambisano, N., Carta, E., Dardano, S., Dive, M., Fasquelle, C., Frennet, J.C., Hanset, R., Hubin, X., Jorgensen, C., Karim, L., Kent, M., Harvey, K., Pearce, B.R., Simon, P., Tama, N., Nie, H., Vandeputte, S., Lien, S., Longeri, M., Fredholm, M., Harvey, R.J. and Georges, M. (2008) *Nat Genet*. **40**:449

Frischmeyer, P.A., and Dietz, H.C. (1999) *Hum Mol Genet* **8**:1893

Harismendy, O., Ng, P.C., Strausberg, R.L., Wang, X., Stockwell, T.B., Beeson, K.Y., Schork, N.J., Murray, S.S., Topol, E.J., Levy, S., and Frazer, K.A. (2009) *Genome Biol*. **10**:R32. [Epub ahead of print]

Houweling, P.J., Cavanagh, J.A., Palmer, D.N., Frugier, T., Mitchell, N.L., Windsor, P.A., Raadsma, H.W., and Tammen, I. (2006) *Biochim Biophys Acta*. **1762**:890

Tammen, I., Cavanagh, J.A.L., Harper, P.A.W., Cook, R.W., Raadsma, H.W., and Nicholas, F.W. (1999). *Archives of Animal Breeding.* **42**:163

Tammen, I., Cook, R.W., Gitschier, J., Nicholas, F.W., and Raadsma, H.W. (2002) *Proceedings of the 28th International Conference on Animal Genetics*. C059 p70

Tammen, I., Houweling, P.J., Frugier, T., Mitchell, N.L., Kay, G.W., Cavanagh, J.A., Cook, R.W., Raadsma, H.W., and Palmer, D.N. (2006) *Biochim Biophys Acta*. **1762**:898

IGF1 GENOTYPES AFFECT GROWTH NOT TENDERNESS IN CATTLE

Lei Yao Chang, W.S. Pitchford, and C.D.K. Bottema

Cooperative Research Centre for Beef Genetic Technologies
Animal Science, The University of Adelaide, Roseworthy, S.A. 5371

SUMMARY

368 backcross progeny from crosses between two *Bos taurus* breeds (Limousin and Jersey) used to study the effects of IGF1 genotypes on meat weight, tenderness and muscle hypertrophy. The results revealed that single nucleotide polymorphisms (SNPs) in the *IGF1* gene were associated with carcass weight and meat weight but not tenderness as measured by shear force. Interestingly, the *IGF1* SNPs were not associated with meat percentage or fibre diameter, suggesting the gene affects growth in general rather than muscle hypertrophy specifically.

INTRODUCTION

Insulin-like growth factor I (*IGF1*) is known to play an important role in various aspects of muscle growth and development (Bunter *et al*. 2005; Davis and Simmen 2006). Due to the effect of *IGF1* on the hypertrophy of muscle cells, muscle fibre diameter can be affected by *IGF1* (Musaro *et al.* 2001). Since increasing muscle fibre diameter may decrease tenderness (Herring *et al.* 2009), it can be postulated that *IGF1* may also affect tenderness by increasing the size of the muscle fibres (Koohmaraie *et al.* 1995). The aim of this study was to investigate relationship between DNA polymorphisms in the *IGF1* gene and tenderness and muscle development.

MATERIALS AND METHODS

The experimental herd design, phenotypes and genotypes were used from JS Davies Gene Mapping Cattle Project (Esmailizadeh *et al*. 2008). The Australian mapping herd (with 368 backcross progeny) was derived from crosses between two extreme *Bos taurus* breeds (Limousin and Jersey). Two single nucleotide polymorphisms (SNPs) in the *IGF1* gene were genotyped using the Illumina system and high resolution melt, namely SNP1 (C/T) 313 bp before exon 1 in the 5' flanking region and SNP2 (C/T) 7 bp from the exon 4 of splice junction. Tenderness was quantified as a measure of Warner-Bratzler (WB) shear force on two muscles: *M. longissimus dorsi* muscle (LD) and *M. semitendinosus muscles* (ST). To improve the accuracy of the tenderness phenotype, the shear force values from four time points (that is, 4 different days of ageing) were adjusted for using a mixed model. The fixed effects fitted in the mixed model were cohort (combination of sex and year), breed, sire, *myostatin* F94L genotype (AA, AC, CC), ageing time (1, 5, 12 and 26 days), muscle (*M. longissimus dorsi* muscle, LD and *M. semitendinosus muscles*, ST) and their interactions. Random effects fitted in the mixed model included animal, animal.muscle and animal.ageing time. The BLUPs for animal.muscle were used as the 'adjusted' values for tenderness.

The phenotypes for the same animals of LD weight, ST weight, meat percentage, total meat weight, HSCW (carcass weight) and fibre diameter (from the ST muscles) were also used in the study. The effect of *IGF1* on tenderness was analysed with genotypes for *IGF1*-SNP1 and SNP2 and the interaction between these two SNPs of *IGF1* as fixed factors in the analysis model. For analysing the effect of the *IGF1* gene on the other traits, cohort (combination of sex and year), breed, sire, *myostatin* F94L genotype, genotypes for *IGF*1-SNP1 and SNP2 and the interaction between these two SNPs of *IGF1* were fitted as fixed factors in the model. The *myostatin* F94L had a large effect on body composition (Esmailizadeh *et al*. 2008) and therefore, was included in the model. The "C" allele frequency of the *IGF1*-SNP1 was 41% and for SNP2 was 80%. These

alleles were in Hardy-Weinberg equilibrium. All analyses were conducted with Genstat 8.1 (Lawes Agricultural Trust 2005). Significance was defined as P<0.05.

RESULTS AND DISCUSSION

The *IGF1* gene (SNP1, SNP2 or the interaction between the two SNPs) did not show any effect on tenderness as measured by shear force for either the LD or ST muscles (Table 1.). In addition, fibre diameter was not affected by *IGF1* in the ST muscle. Likewise, the weights of the LD and ST muscles were not affected by SNP1 and/or SNP2 of the *IGF1* gene. This suggests that the *IGF1* gene may not cause muscle hypertrophy in either the LD or ST muscles. This is not consistent with the previous research results that have shown a relationship between the level of IGF1 and hypertrophy of muscles (Musaro et al., 2001). However, the effect of the interaction between SNP1 and SNP2 of *IGF1* gene on ST muscle weight is nearly significant (P=0.06).

Table 1. Test of significant of *IGF1* SNP genotypes on carcass traits.

	No. of observations	Mean	Standard Deviation	*SNP1*	*SNP2*	*SNP1.SNP2*
Adjusted_WB_{ST} [a]	366	4.758	0.385	0.909	0.820	0.410
Adjusted_WB_{LD} [b]	366	4.228	0.676	0.987	0.487	0.369
ST weight [c]	349	2.49	0.837	0.679	0.356	0.060†
LD weight [d]	347	6.28	1.507	0.112	0.299	0.445
ST fibre diameter [e]	276	66.04	12.84	0.412	0.505	0.175
meat % [f]	330	68.62	2.99	0.306	0.686	0.111
Total meat weight	329	230.3	48.5	0.032*	0.284	0.406
carcass weight [g]	356	334.7	61.7	0.013*	0.082†	0.192

* (P<0.05) † P<0.10

[a] Adjusted_WB_{ST} means Warner-Bratzler (WB) shear force (adjusted by mixed model) on *M. semitendinosus muscles* (ST)

[b] Adjusted_WB_{LD} means Warner-Bratzler (WB) shear force (adjusted by mixed model) on *M. longissimus dorsi* muscle (LD)

[c] ST weight means the weight of *M. semitendinosus* muscle (ST)

[d] LD weight means the weight of *M. longissimus dorsi* muscle (LD)

[e] ST fibre diameter measured with *M. semitendinosus* muscle (ST)

[f] meat % refers to meat percentage

[g] carcass weight refers to hot standard carcass weight

On the other hand, the *IGF1*-SNP1 was associated with total meat weight (P=0.032) and hot standard carcass weight (HSCW) (P=0.013). Cattle with the TT and CT genotypes for *IGF1*-SNP1 had more meat than the cattle with CC genotype (Figures 1 and 2). The *IGF1*-SNP1 showed a significant dominance effect on meat weight and carcass weight (Table 2). The estimated allelic substitution effect was 7.08 ± 3.09kg on meat weight and 11.76 ± 4.05kg on hot standard carcass weight (Table 2).

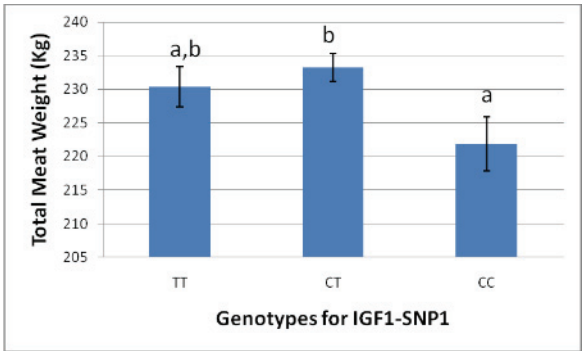


Figure 1. Effect of *IGF1*-SNP1 on meat weight. Different letters indicate significant differences between groups.

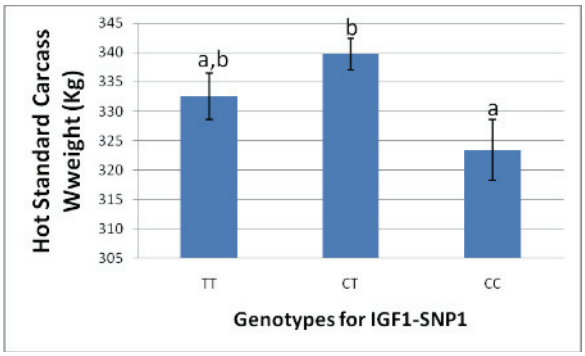


Figure 2. Effect of *IGF1*-SNP1 on carcass weight. Different letters indicate significant differences between groups.

Table 2. Additive and dominance effects (+ standard errors) of *IGF1* SNPs on significant carcass traits.

	*IGF1*_SNP1 additive	*IGF1*_SNP1 dominance	*IGF1*_SNP2 additive	*IGF1*_SNP2 dominance
meat weight	-4.24 (2.73)	7.08 (3.09)*	4.94 (4.33)	7.57 (4.76)
carcass weight	-4.51 (3.56)	11.76 (4.05) **	10.13 (5.27) †	12.89 (5.93) *

†(P<0.10); * (P<0.05), * * (P<0.01), *** (P<0.001)

The *IGF1*-SNP2 did not show a significant effect on hot standard carcass weight (P=0.082). However, the cattle with the TT genotype for the *IGF1*-SNP2 had significantly lower carcass weights than the cattle with the CC and CT genotypes (Figure 3). The dominance effect was found and its estimated allelic substitution effect was 12.89 ± 5.93 (Table 2). These results support the observations of Davis and Simmen (2006), who found that serum IGF1 levels were moderately to highly heritable and were correlated to pre- and post-weaning weight gain in cattle.

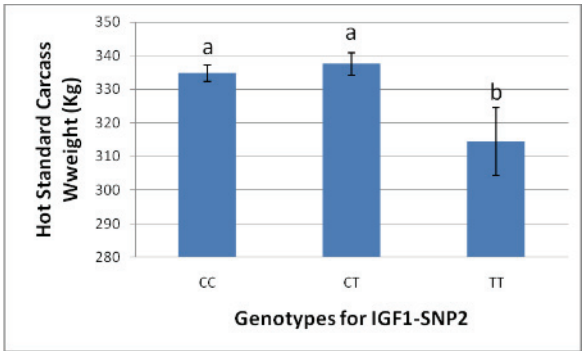


Figure 3. Effect of *IGF1*-SNP2 on carcass weight. Different letters indicate significant differences between groups.

Interestingly, meat percentage was not influenced by the *IGF1* gene. The results taken together suggest that the *IGF1* gene does not appear to specifically increase the size of the muscle fibres, but does affects growth overall. Hot standard carcass weight increased because the animals with the "C" allele of the *IGF*1-SNP2 and the "T" allele of the *IGF*1-SNP1 were larger overall, not because the animals had more muscle as a proportion of the carcass. Hence, the polymorphisms in the *IGF1* gene were only associated with growth but not with the size or weight of specific muscles. Given that DNA variants in the *IGF1* gene do not appear to be associated with muscle hypertrophy, it is not surprising the polymorphisms in the *IGF1* gene were also not associated with tenderness as measured by shear force. On the other hand, *IGF1* has been shown to change the proportions of muscle types, which may affect tenderness (Lynch *et al.* 2001 and Klont *et al.* 1998). Hence, the relationship between the *IGF1* gene, muscle fibre types and tenderness needs to be further investigated.

ACKNOWLEDGEMENTS

We would like to thank David Lines and Zbigniew Kruk for measuring fibre diameter and tenderness and the J.S. Davies Bequest for funding the cattle gene mapping project.

REFERENCES

Bunter, K., Hermesch, S., Luxford, B.G., Graser, H.U. and Crump, R.E. (2005) *Aust J Exp Agric* **45**:783.

Davis, M.E. and Simmen, R.C.M. (2006) *Journal of Animal Science* **84**:2299.

Esmailizadeh, A.K., Bottema, C.D.K., Sellick, G.S., Verbyla, A.P., Morris, C.A., Cullen, N.G. and Pitchford, W.S. (2008) *Journal of Animal Science* **86**:1038.

Musarò. A., McCullagh, K., Paul, A., Houghton, L., Dobrowolny, G., Molinaro, M., Barton, E.R., Sweeney, H.L. and Rosenthal, N. (2001) *Nature Genetics* **27**: 195.

Koohmaraie, M., Shackelford, S. D., Wheeler, T. L., Lonergan, S. M. and Doumit, M. E. (1995) *Journal of Animal Science* **73**:3596.

Klont, R. E. Brocks, L. and Eikelenboom, G. (1998) *Meat Science* **49**:S219

Lawes Agricultural Trust (2005) GenStat Release "8.1 (PC/Windows 2000) 09 April 2009 15:10:55 Copyright". (Rothamsted Experimental Station).

Lynch, G.. S., Cuffe, S. A., Plant, D. R. and Gregorevic, P. (2001) *Neuromuscular Disorders* **11**:260.

Herring, H. K., Cassens, R. G., SUESS, G. G., Brungardt, V. H. and Briskey, E. J. (2009) *Journal of Food Science* **32**:317.

IS METHANE PRODUCTION LIKELY TO BE A FUTURE MERINO SELECTION CRITERION?

D.J. Cottle[1], J.H.J. van der Werf[1,2] and R.G. Banks[3]

[1]School of Environmental and Rural Science, UNE, Armidale, NSW 2351
[2]Cooperative Research Center for Sheep Industry Innovation, Armidale, NSW 2351
[3]Meat and Livestock Australia, Armidale, NSW 2350

SUMMARY

Selection index theory was used to model the effects of including feed intake and methane (CH_4) production in the breeding objective and having both, either or none of these traits as selection criteria on genetic responses in Merino production traits. A range of economic values (EV) were assumed for CH_4 production based on expected future prices for CO_2-e ($/tonne). The implicit price of carbon required to achieve desired gains of a reduction in CH_4 of 1% p.a. were calculated. The sensitivity of desired gains to changes in the correlations between production traits and CH_4 production were modeled as these correlations are currently unknown. If the correlations between production traits and CH_4 are positive (as expected) then it is very unlikely that CH4 production would be used as a selection criterion as an implicit carbon price of over $400/tonne CO_2-e was needed to achieve the desired reductions in CH_4 production. However, if the correlations are unexpectedly negative, the carbon price needed to achieve such gains was more likely at ~$30/tonne CO_2-e. The correlations need to be determined from research trials for informed advice about breeding for CH_4 reduction to be given to Merino breeders.

INTRODUCTION

Rumen methanogenesis results in the loss of up to 12% of gross energy intake (Johnson *et al.* 1993). Methane, a greenhouse gas, is estimated to contribute about 24% of anthropogenic global warming, second only to carbon dioxide (Houghton 1997). Most of Australia's agricultural emissions come from enteric emissions and 15% from sheep enteric emissions (Australian Government (AG) 2008). Sheep daily CH_4 production (~20g/d) is highest when the energy density of the diet is about 10.5 MJ/kg DM and diet digestibility about 70% (Pelchen and Peters 1998). Most current technologies to control CH_4 emissions are not cost effective (Keogh and Cottle 2009). While CH_4 emissions are greater on improved pasture with higher stocking rates, additional farm profit exceeds the potential cost of additional emissions (Alcock and Hegarty 2006). Development of CH_4 mitigation strategies, without causing negative impacts on production, is a major challenge for ruminant nutritionists (McAllister *et al.* 1996).

An alternate approach is to achieve small cumulative decreases in CH_4 production through sheep selection. Between-sheep variations in CH_4 emission have been observed in respiration chambers (Blaxter and Clapperton 1965) and under grazing conditions (Lassey *et al.* 1997; Pinares-Patin *et al.* 2003; Ulyatt *et al.* 1999), where ~85% of the variation in daily CH_4 production from sheep grazing temperate pastures was due to variation between animals.

The Australian government has committed to reduce carbon pollution by 5% of 2000 levels by 2020 (AG 2008). It is not considered practical to immediately include agriculture in a carbon trading scheme, however it is likely to be included after 2015. A final decision will be made in 2013, contingent on there being reliable and cost-effective methods of emissions estimation and reporting. The likely cost of carbon permits is $10-40/tonne CO_2-e (AG 2008).

This paper addresses the questions - if CH_4 production and/or feed intake were more easily measured, are they likely to be used as selection criteria in future breeding programs, and what (incentive) price of carbon is needed to achieve targets of 0 or 1% reduction in CH_4 /year?

METHODS

The selection index program, MTIndex, was used to construct the Sheep Genetics Merino 14% MP (medium micron premium) index with the addition of feed intake (kg/year) and CH_4 (kg/year) (Tables 1 and 2). The additional traits were included in the breeding objective using genetic parameters from the literature (e.g. Ponzoni 1986, Lee *et al.* 2002, Safari *et al.* 2005). Correlations between CH_4 and production traits were not available and were assumed after discussion amongst colleagues (e.g. P. Amer, pers. comm.). The EV of CH_4 production (kg/day) was calculated as \$/tCO_2-e * 21/1000, where 21 is the internationally accepted global warming potential of CH_4 in CO_2 equivalents (AG 2008). The EV of NLW accounted for predicted lamb offspring CH_4 production. The index was run unconstrained or with a desired gain of -0.16 kg CH_4 /ewe /generation/i (~1% p.a. CH_4 reduction) to calculate the implicit carbon price needed for this reduction. Sensitivity of calculated trait genetic gains to the assumed CH_4 correlations and EV were studied by randomly sampling 100 times within the normal distributions of these values with an assumed standard deviation (0.1). The impacts of including feed intake and/or CH_4 production as selection criterion on gains in all traits were calculated.

Table 1. Traits in the 14% MP index - assumed genetic parameters and economic values

Trait*	Units	σ_p	Heritability	Repeatability	Economic Value
YCFW	%	18.27	0.29	0.40	\$0.29
YFD	μm	1.22	0.55	0.55	-\$3.22
FI	kgDM/year	75.00	0.13	0.35	-\$0.02
YWT	kg	5.56	0.35	0.50	\$0.26
ACFW	%	14.16	0.44	0.60	\$0.30
AFD	μm	1.41	0.60	0.60	-\$4.14
AFDCV	%	2.24	0.35	0.60	-\$0.83
AWT	kg	6.32	0.40	0.40	\$0.04
NLW	No.	0.65	0.06	0.35	\$36.49
CH_4	kg/year	1.46	0.25	0.40	-\$0.63

*YCFW – yearling clean fleece weight %, YFD – yearling fibre diameter, FI- feed intake, YWT – yearling body weight, ACFW – adult clean fleece weight %, AFDCV – adult fibre diameter coefficient of variation, NLW - number of lambs weaned

Table 2. 14% MP index- assumed correlations (phenotypic above, genetic below diagonal)

	YCFW	YFD	FI	YWT	ACFW	AFD	AFDCV	AWT	NLW	CH_4
YCFW		0.31	0.13	0.30	0.42	0.20	-0.01	0.26	0.07	0.10
YFD	0.30		0.07	0.20	0.17	0.62	-0.05	0.15	0.03	0.00
FI	0.14	0.39		0.30	0.11	0.09	-0.04	0.30	0.40	0.77
YWT	0.25	0.20	0.73		0.23	0.14	-0.06	0.61	0.10	0.62
ACFW	0.70	0.25	0.03	0.15		0.29	0.06	0.22	0.04	0.10
AFD	0.30	0.84	0.44	0.15	0.34		-0.11	0.17	0.03	0.04
AFDCV	0.15	-0.06	-0.23	-0.05	0.24	-0.14		-0.14	0.02	0.00
AWT	0.10	0.21	0.73	0.80	0.15	0.20	-0.14		0.10	0.62
NLW	-0.10	0.00	0.40	0.15	-0.10	0.00	0.00	0.15		0.26
CH_4	0.05	0.00	0.70	0.65	0.05	0.02	0.00	0.65	0.25	

RESULTS AND DISCUSSION

When positive correlations between CH_4 production and production traits were assumed, CH_4 production increased as a result of selection until carbon price was over \$150/t CO_2-e, even when both feed intake and CH_4 production were included as selection criteria.

To achieve a 1% p.a. CH_4 reduction the carbon price needed to be \$418/t CO_2-e if both feed intake and CH_4 production were used as selection criteria, \$421/t CO_2-e if only CH_4 was a criterion, \$570/t CO_2-e if only feed intake was a criterion and \$698/t CO_2-e if neither were used as additional criterion. A carbon price of \$140/t CO_2-e was needed for zero predicted change in CH_4.

Table 3. Predicted gains in traits or ratios (/10 years) assuming a carbon price of \$30/tCO_2-e, positive production- CH_4 correlations, with different combinations of selection criteria in addition to CFW, FD, FDCV, WT and NLW.

Selection criteria	ACFW	AFD	AWT	CH_4	Feed intake	CH_4 /kg AWT	CH_4 /kgACFW
None	0.41	-2.77	-0.72	0.31	-21.85	-5.5%	5.0%
CH_4	0.40	-2.77	-0.73	0.29	-22.09	-5.3%	5.1%
Feed intake (FI)	0.51	-2.81	-0.89	0.27	-23.91	-5.3%	7.5%
CH_4 and FI	0.68	-2.84	-1.04	0.33	-24.91	-6.4%	10.0%
*CH_4+FI (sire)	1.74	-2.84	-0.66	0.38	-22.05	-6.4%	25.6%

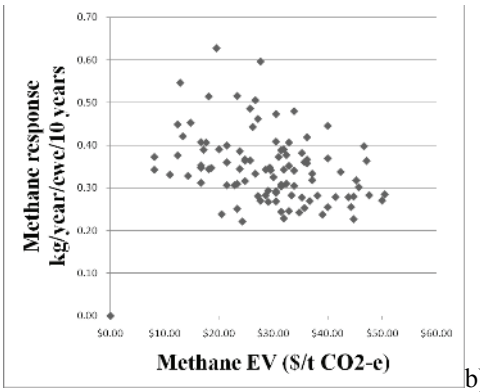
*Half sib progeny measured in addition to individual performance

CH_4 production (per year and per kg CFW) increased when feed intake and/or CH_4 production were included as criteria as they were assumed positively correlated to production traits and indirectly increased these traits. An annual fleece was worth ~\$36.00 compared to the carbon cost of annual CH_4 production (~7.3kg/head) of -\$4.60, using the assumed CH_4 EV in Table 1. When the production –CH_4 correlations in Table 2 were given the opposite, negative sign the predicted reductions in CH_4 production were, as expected, much greater (Table 4).

Table 4. Predicted gains in traits or ratios (/10 years) assuming a carbon price of \$30/ t CO_2-e, negative production- CH_4 correlations, with different combinations of selection criteria in addition to CFW, FD, FDCV, WT and NLW.

Selection criteria	aCFW	aFD	aWT	CH_4	Feed intake	CH_4 /kg wt	CH_4 /kgCFW
None	-0.20	-2.95	-1.85	-0.62	-28.02	5.6%	3.9%
CH_4	-0.12	-3.02	-1.86	-1.26	-36.60	14.6%	14.8%
Feed intake	0.63	-2.82	-0.19	-0.92	-18.93	12.3%	24.0%
Both	0.77	-2.74	-0.60	-0.14	-10.87	0.9%	17.1%
Both (sire)	1.82	-2.82	0.25	-0.81	-13.86	11.5%	38.0%

The sensitivity analyses of production trait – CH_4 correlations and CH_4 EV demonstrated that the predicted CH_4 responses were very dependent on these values (Figure 1). However, it is expected that these correlations are positive, as the feed intake- CH_4 correlation is highly positive, and production traits are positively correlated to feed intake, so the predicted responses of increases in CH_4 production (Figure 1a) are more likely than small increases or reductions in CH_4 (Figure 1b).

a) 

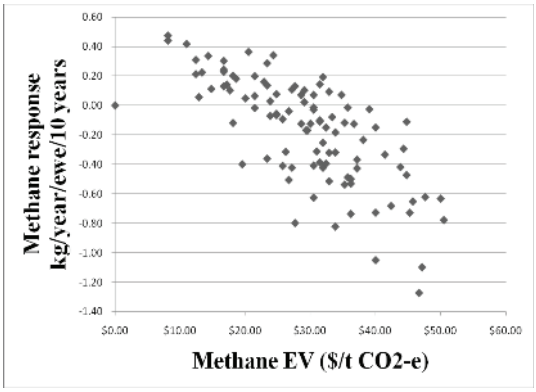
b) 

Figure 2. Sensitivity of CH_4 response to selection (/10 years) versus carbon price with CH_4 and feed intake as additional selection criteria where correlations were samples around a) positive, or b) negative values for the mean production-CH_4 correlation.

CONCLUSIONS

The results of these analyses suggest that CH_4 production per sheep has probably been increasing in Merino flocks using the 14% MP index, where CH_4 and feed intake have not been available for use as practical selection criteria. If they could be used, CH_4 production would probably still increase if carbon prices are around \$20-30/tCO_2-e because the price is not high enough to place enough selection pressure to reduce CH_4 production. Estimates of production trait- CH_4 correlations are needed to design optimum sheep breeding programs if agriculture is to be included in carbon trading schemes in the future.

REFERENCES

Alcock, D. and Hegarty, R.S. (2006) *Proc. 2nd Int. Conf. Greenhouse Gases and Anim. Agric*., Zurich, Switzerland **1293**:103.

Australian Government (2008) Carbon Pollution Reduction Scheme White paper. December http://www.climatechange.gov.au/whitepaper/summary/index.html

Blaxter, K.L. and Clapperton, J. L. (1965) *Br. J. Nut.* **19:**511.

Houghton, J. (1997) Global Warming: the Complete Briefing, 2nd ed.. Cambridge, UK

Johnson, D. E., Hill, T. M.,Ward, G. M., Johnson, K. A., Branine, M. E., Carmean, B.R. and Lodman, D.W. (1993) In *Atmospheric CH4: sources, sinks, and role in global change* (ed. M. A. K. Khalil), p. 219. Berlin, Germany.

McAllister, T.A., Okine, E.K., Mathison, G.W. and Cheng, K.J. (1996) Dietary, environmental and microbiological aspects of methane production in ruminants. *Can. J. Anim. Sci.* **76**:231.

Keogh, M. and Cottle, D.J. (2009) *Recent Adv. Anim. Nut.* (in press)

Lassey, K. R., Ulyatt, M. J., Martin, R. J., Walker, C. F. and Shelton, I. D. (1997) *Atmospheric Environ.* **31**:2905.

Lee, G.J., Atkins, K.D. and Swan A.A. (2002) *Livest. Prod. Sci.* **73**:185.

Pelchen, A. and Peters, K.J. (1998) *Small Ruminant Research* **27**:137.

Pinares-Patino, C. S., Ulyatt, M. J., Lassey, K. R., Barry, T.N.. and Holmes, C. W. (2003) *J. Agric. Sci.* **140**: 227.

Ponzoni, R.W. (1986) *J. Anim. Breed. Genet*. **103**:342.

Safari, E., Fogarty, N.M. and Gilmour, A.R. (2005) *Livest. Prod. Sci.* **92**:271.

Ulyatt, M. J., Baker, S. K., McCrabb, G. J. and Lassey, K. R. (1999) *Aust. J. Agric. Res.* **50**:1329.

POLYMORPHISMS DETECTED IN THE *TYROSINASE* AND *MATP* (*SLC45A2)* GENES DID NOT EXPLAIN COAT COLOUR DILUTION IN A SAMPLE OF ALPACA (*VICUGNA PACOS*)

R. Cransberg and K. A. Munyard

Alpaca Molecular Research Group, School of Biomedical Science, Curtin University of Technology, Perth, WA 6102

SUMMARY

The molecular basis of inheritance of alpaca fibre colour is poorly understood. However, colour dilution genes are anticipated to be causing a major effect on alpaca fibre colour. Candidate genes for dilutions included *tyrosinase* (*tyr*) and *membrane associated transport protein* (*matp*), both of which have been associated with coat colour dilution in other species. The coding regions of the *tyr* and *matp* genes were sequenced in 24 animals with various colour phenotypes. No polymorphism found in the coding region of *tyr* and *matp* exons 1, 3, 4, 5 and 7 could account for the dilutions in fibre colour observed.

INTRODUCTION

Alpaca fibre is highly valued because it is strong and suitable for use in high quality clothes, fabrics and carpets (Frank *et al.* 2006). White alpaca fibre is considered the most valuable because it is easy to dye, thus coloured fibres only sell for about 65% of the price of white fibres (Frank *et al.* 2006). However there are not enough white alpacas in Australia to meet market demand, with only approximately a third of alpacas being light enough to be classed as white (Paul 2006). This project examined the *tyr* and *matp* genes in alpacas, located at the *albino* and *underwhite* loci respectively. The aim was to identify polymorphisms that occur within these genes and ascertain if any of these can be consistently linked to a particular colour phenotype. We hypothesise that polymorphisms in these genes may result in a structural change, or a change in expression levels which will therefore alter the quantity of melanin produced, and result in a coat colour phenotype which is lighter or darker than that which would be expected of the wildtype individual.

MATERIALS AND METHODS

Samples. Blood (approximately 5mL) was taken from 24 alpacas of both sexes, with a variety of different colour phenotypes, from various herds across Western Australia, New South Wales and Victoria. The colour variants selected placed an emphasis on potentially diluted phenotypes such as pink skin (PS) white and PS light fawn and also included non-fading and fading black, black and tan, bay, chestnut, black skin (BS) brown, PS rosegrey, BS silvergrey, BS white and BS blue-eyed white.

DNA Extraction. DNA was extracted from alpaca blood using the salt precipitation method described by Miller et al. (1988) Where this method failed to yield sufficient quantities of DNA, or where the quantity of blood obtained was insufficient for the salt precipitation method, the DNeasy blood and tissue DNA extraction kit (Quigen) was used, according to the manufacturer's instructions.

Primer Design and Amplification. Alpaca *tyr* and *matp* sequences were retrieved from trace archive files deposited in GenBank (NCBI) where possible, using BLAST analysis against known cattle exon sequences. Primers were then designed in regions flanking each exon. Where the trace

archive sequence was incomplete (for *matp* exons 1, 2 and 6) a comparative genetics approach was adopted. Sequences from the cattle, dog, human and mouse genomes were aligned and primers were designed using sequence from consensus regions flanking each exon.

Genomic DNA (50-100ng) was used as template for PCR using 2 μM of forward primer, 2μM reverse primer, 1× Polymerisation buffer (Fisher biotec), 2mM $MgCl_2$ (Fisher biotec) and 0.75u BIOTAQ polymerase (Bioline) in a 10μL reaction. Optimized amplification conditions were: 95°C for 2 min; 30-40 cycles of 95°C for 20s, 57-58°C for 30s and 72°C for 1 min; then 72°C for 5 min. *tyr* exon 1 and *matp* exons 3, 4 and 5 were amplified using 30 cycles with 58°C annealing temperature; *tyr* exons 2, 4 and 5 were amplified using 30 cycles with 57°C annealing temperature; *tyr* exon 3 was amplified using 40 cycles at 58°C annealing temperature, *matp* exon 1 was amplified using 35 cycles at 57°C annealing temperature; and *matp* exon 7 was amplified using 35 cycles at 58°C annealing temperature.

DNA Sequencing. PCR products were amplified from genomic DNA in five separate 10μL reactions and pooled prior to sequencing. Sequencing was performed using the ABI Big Dye Terminator® system at either Murdoch University, Perth using a 48-capillary ABI 3730 DNA analyser or at Macrogen, Korea, using 96-capillary ABI 3730XL DNA analysers.

Analysis. Vector NTI software (Invitrogen) was used to visualise and analyse sequencing results and SpliceView (ORIEL) was used to aid in determining intron/exon boundaries for predicted proteins. Statistical analyses were performed using the JMP® 7 statistical package (SAS institute Inc. 2007). The Pearson χ^2 was used to investigate if an association existed between fibre colour and SNP's or nucleotides. Statistical significance was accepted at $p \leq 0.05$.

RESULTS

The five exons of the *tyr* gene and five of the seven *matp* exons were sequenced all 24 animals. While *tyr* exon 2 was largely sequenced only in one direction due to difficulties sequencing through an adjacent microsatellite, most exons were sequenced in both directions for the majority of alpacas. Ten single nucleotide polymorphisms (SNPs) were detected in the coding region of *tyr*.

Table 1. Single Nucleotide Polymorphisms found in the coding region of *tyr*

Exon	Nucleotide Position	Nucleotide Change	Amino Acid Change
Exon 1	78	C/A	NA
Exon 1	126	C/T	NA
Exon 1	162	C/T	NA
Exon 1	480	C/T	NA
Exon 1	513	C/T	NA
Exon 1	784	C/T	NA
Exon 2	851	G/T	Ala/Ser
Exon 5	1372	C/T	NA
Exon 5	1490	G/T	Arg/Leu
Exon 5	1498	T/C	Cys/Arg

Seven of these SNPs were synonymous polymorphisms and three were non-synonymous (Table 1). Two SNPs were detected in the five *matp* exons that were sequenced. One of these was

synonymous and the other was non-synonymous (Table 2). The C1526T mutation was found in seven heterozygotes and 17 homozygote C animals. There was no association between SNP and fibre colour (χ^2=0.14, p=1.0) for any of the SNPs found in *tyr* or *matp*.

An alpaca TYR protein sequence was predicted from the five exon sequences, using SpliceView (ORIEL) coupled with comparative genetics using the known cattle coding sequence. This protein was 530 amino acids long; 89.4% identical to the cattle and 88.1% identical to the human protein.

Table 2. Single Nucleotide Polymorphisms found in the coding region of *matp*

Exon	Nucleotide Position*	Nucleotide Change	Amino Acid Change
Exon 3	843	G/A	NA
Exon 7	1526	C/T	Thr/Met

* Nucleotide positions are speculative, using cattle sequence to fill the regions of exon 2 and 6

Matp exons 2 and 6 gave multiple amplicons in PCR after extensive optimisation of reaction conditions. The multiple amplicons produced by PCR, from two different animals, were cloned, and three of these clones were sequenced for each animal. However, BLAST searches confirmed that none of these clones contained sequence homologous to *matp* from any other species.

DISCUSSION

Tyrosinase. None of the three non-synonymous SNPs found in the *tyr* gene had any significant correlation with fibre and/or skin colour. The G851T polymorphism occurs around the middle of the protein, and is 53 amino acids away from the nearest glycosylation site or copper binding domain (Schweikardt *et al.* 2007, Wang and Hebert 2006). The likelihood of this change having little effect is further supported by the observation that serine is present in the same position of the protein in the human and rabbit, while the alanine, more common in alpacas (Table 2), is also present in cattle, dog and chicken (Schmutz *et al*. 2004).

The G1490T polymorphism causes a change in the last amino acid of the TYR transmembrane region, if the alpaca protein is similar to most known Tyr proteins (Schmutz *et al*. 2004). However, the alpaca *tyr* gene encodes for arginine in this position, differing from the genes of the dog, human, mouse and cattle, which all encode serine; the rabbit which encodes glycine; and the chicken which encodes isoleucine (Schmutz et al. 2004). This represents a large biological difference in amino acid, however, no SNP in the cytoplasmic tail of Tyr have been documented to have any effect on coat colour in the extensively studied human and mouse *Tyr* genes and there was no apparent correlation with a change in alpaca fibre colour either.

The T1498C SNP occurred only three codons after the one previously described. The fact that two non-synonymous SNPs were found close together in this region, suggests that the region is under little selective conservation. The cysteine residue present at this position in the wild-type is highly conserved in Tyr proteins of other species (Schmutz *et al.* 2004). However the position of this residue in the cytoplasmic tail separates it from other cysteine residues by the melanosomal membrane, therefore this residue does not form any disulfide bonds in the final protein structure of other species (Wang and Hebert 2006).

Membrane Associated Transport Protein. For the non-synonymous C1526T polymorphism identified in the *matp* gene no homozygote T animals were observed. This suggests that homozygote T at this position may be non-viable. As seven of the animals were heterozygous for

this SNP and the other 17 were homozygote C, the sample group is also clearly not in Hardy-Weinberg proportions. Given the broad range of animals in the sample group this would appear to be significant, however there is no apparent correlation from this SNP to any given colour phenotype. Based on the Newton et al. (2001) prediction of the human MATP protein, this polymorphism will lie in the 12th and final transmembranous region. This region is highly conserved among human, mouse, cattle and alpaca coding regions. Given that the polymorphism lies within a transmembranous region, the change from threonine to a non-polar, hydrophobic methionine may prove to have little effect. Indeed a similar threonine for proline substitution was reported by Inagaki *et al*. (2004) earlier in the same transmembranous region in humans, which was identified as a population variant, with no significant effect on hair colour.

CONCLUSION

A sample group of 24 alpacas yielded a total of ten SNPs in the coding region of the *tyr* gene, while two SNPs were found in five of the seven exons of *matp*. None of these SNPs correlated with a detectable change in colour within the sample group. However, our spectrophotometric studies have strongly suggested that colour dilution genes play a major role in alpaca pigmentation (data not shown). Key regulatory regions of the *tyr* and *matp* genes will have to be examined thoroughly before these genes are confirmed as not playing a significant role in alpaca pigmentation.

ACKNOWLEDGMENTS

This work was funded largely by the Rural Industry Research and Development Corporation (RIRDC).

REFERENCES

Frank, E. N., Hick, M. V. H., Gauna, C. D., Lamas, H. E., Renieri, C. and Antonini, M. (2006) *Small Ruminant Research,* **61:**113.

Inagaki, K, Suzuki, T, Shimizu, H, Ishii, N, Umezawa, Y, Tada, J, Kikuchi, N, Takata, M, Takamori, K, Kishibe, M, Tanaka, M, Miyamura, Y, Ito, S & Tomita, Y (2004) *The American Journal of Human Genetics,* **74:**466.

Miller, S. A., Dykes, D. D. and Polesky, H. F. (1988) *Nucleic Acids Research,* **16:**1215.

Newton, J. M. Cohen-Barak, O., Hagiwara, N., Gardner, J. M., Davisson, M. T., King, R. A. and Brilliant, M. H. (2001) *The American Journal of Human Genetics,* **69:** 981.

Paul, E (2006) Colour Review. *Australian Alpaca Association National Conference.* Adelaide.

Schmutz, S. M., Berryere, T. G., Ciobanu, D. C., Mileham, A. J., Schmidtz, B. H. and Fredholm, M. (2004) *Mammalian Genome,* **15:** 62.

Schweikardt, T., Olivares, C., Solano, F., Jaenicke, E., Garcia-Borron, J. C. and Decker, H. (2007) *Pigment Cell Research,* **20:** 394.

Wang, N. andHebert, D. N. (2006) *Pigment Cell Research,* **19:** 3.

QTL ANALYSES OF BEEF MUSCLE FIBRE TYPE

N. G. Cullen[1], C. A. Morris[1], P.M. Dobbie[1], D.L. Hyndman[2] and B.C. Thomson[3]

[1]AgResearch, Ruakura Research Centre, PB 3123, Hamilton, New Zealand
[2]AgResearch, Invermay Agricultural Centre, PB 50034, Mosgiel, New Zealand
[3]On Farm Research, PO Box 1142, Hastings, New Zealand

SUMMARY

Beef cattle have been selected in particular for increased growth rates and greater muscle development. The myostatin gene is responsible for the muscle hyperplasia in double muscled animals and it is also responsible in these animals for differences, during foetal development, in the relative proportions of muscle fibre types. A mutation in the myostatin gene in Limousin cattle (F94L) has been found to produce a milder form of double muscling. A whole genome scan identified a highly significant quantitative trait locus (QTL) at the proximal end of chromosome 2 in a Jersey-Limousin double backcross trial for percentage meat and percentage fat in the carcass, and also for proportions of fibre type IIb count and area in the *M. longissimus*. We propose that this same F94L mutation is responsible for the QTL associated with increased proportions of type IIb muscle fibres in the Jersey-Limousin crosses. A significant QTL for fibre type was also found on chromosome 13 ($P<0.05$).

INTRODUCTION

A collaborative study began in 1995 between AgResearch in New Zealand (NZ) and the University of Adelaide in Australia to search for DNA markers linked to production, carcass and beef meat quality traits (Morris *et al.* 2009). The present paper reports on muscle fibre type composition, which is one factor suggested as causing meat quality differences. Fibre-type composition has been reviewed in relation to meat quality by Klont *et al.* (1998). A sub-set of the animals born in NZ were analysed for muscle fibre type and we report here a QTL search performed to identify chromosomal regions with significant linkage to fibre type.

MATERIALS AND METHODS

Trial design. The trial design involved dams of two very different *Bos taurus* breeds, Jersey (J) and Limousin (L). In NZ, three first-cross JxL or LxJ bulls were mated with both J and L cows, to produce a total of about 400 heifer or steer back-cross progeny. The marker-search involved identifying in the calves sire-derived alleles whose presence was associated with performance in one or more traits ("phenotypes"). The primary traits of interest were carcass composition and measures of beef meat quality. Other simple traits during the growth phase were also recorded, such as live weights and ultrasound measurements. Animals in NZ grazed mainly on pasture. At slaughter, muscle samples were taken to measure meat quality during the aging process. For the first calf crop in NZ (1996 births), the phenotypes used in the present study were measures of fibre type distribution in the *M. longissimus thoracis et lumborum (M. longissimus)* at slaughter.

Sections from a slice of the *M. longissimus* were stained using myosin ATPase histochemistry (Martyn *et al.* 2004) and the counts and areas of each fibre type (I, IIa and IIb) were recorded for 192 animals. From the total samples analysed for each animal, the proportion of the total count and area from all the sections was calculated for each of the three muscle types.

Data analyses. The proportions of total count and area were first analysed by least squares to identify the appropriate model for a subsequent QTL scan. Factors tested were breed, birth type

(the ¾ Limousin were produced by embryo transfer and included many sets of twins), live weight prior to slaughter, sex and hot carcass weight. Sex was confounded with slaughter group as these animals were slaughtered in 18 groups at intervals of a week – animals being allocated to groups before slaughter commenced, with groups being balanced for sire, breed and live weight. Each slaughter group comprised only animals of the same sex with heifers being slaughtered on average earlier than steers. Significant effects were breed and sex, after slaughter group was nested within sex.

Table 1. Numbers of animals, means and residual standard deviations for the proportion of Count and Area for each of the 3 fibre types by breed and sex

			Count (proportion)			Area (proportion)		
			Type I	Type IIa	Type IIb	Type I	Type IIa	Type IIb
Breed	Sex	n	Mean	Mean	Mean	Mean	Mean	Mean
LJJJ	Heifer	54	0.244	0.430	0.326	0.197	0.418	0.385
	Steer	65	0.244	0.385	0.371	0.198	0.359	0.442
LJLL	Heifer	43	0.285	0.365	0.350	0.187	0.336	0.478
	Steer	30	0.281	0.325	0.394	0.182	0.294	0.523
All		192	0.259	0.384	0.357	0.193	0.360	0.447
		rsd*	0.042	0.053	0.057	0.037	0.063	0.073
		cv*	0.162	0.138	0.160	0.191	0.175	0.163

* rsd = residual standard deviation; cv = coefficient of variation

Initial Haley-Knott linkage (Knott *et al.* 1996) runs in SAS identified a large QTL at the proximal end of BTA2. Earlier published work (Sellick *et al.* 2007) had identified a mutation (F94L) in the myostatin gene derived from Limousin, which had a very large effect on the percentages of meat and of fat in a carcass. Records from 183 animals that had fibre type data and also carcass composition data from a butcher's dissection were run through a Haley-Knott routine with a model fitting breed, slaughter group within sex, and sire, to identify QTL. Marker positions were taken from the map of Ihara *et al* (2004). Permutation tests were conducted to determine thresholds for the significance of QTL. With the use of haplotypes (Sellick *et al.* 2007), the maternally inherited haplotype was also identified and we fitted both sire- and dam-derived haplotypes in an analysis of variance to determine the additive and dominance estimates for the myostatin mutation on fibre type distributions. (Some Limousin dams carried the wild-type allele, possibly reflecting the genetic background of some Limousin animals in NZ which were bred up by back-crossing). Similarly, the maternal haplotypes from mainly un-related J and L dams were used to test for linkage disequilibrium. To further test that this mutation was responsible for the differences described, the F94L genotype was fitted as a fixed effect in a re-run of the Haley-Knott scans, with the expectation that the highly significant QTL for fibre type would disappear.

RESULTS

The Haley-Knott graph for chromosome (BTA) 2 is shown in Figure 1 which also included the QTL for percentage of meat in the carcass. There were highly significant QTL identified for both Count and Area of type IIb fibres and lesser significant QTL for both type IIa and type I fibres. Given the fact that the three proportions sum to unity, this is not surprising but the residual correlations between pairs varied widely. For proportions of Total Count and of Total Area, type IIa and IIb have correlations of -0.71 and -0.85 respectively; type I correlations with types IIa and IIb are -0.33 and -0.43 for Count and -0.04 and -0.48 for Area.

A QTL, significant at the 5% level, was found on BTA13 (type I) for one family, and QTL at the suggestive level were found on BTA5 (types I and IIa), BTA7 (type IIb), BTA12 (types IIa and IIb), BTA13 (types I and IIb), BTA14 (type I), BTA20 (type I) and BTA22 (type I).

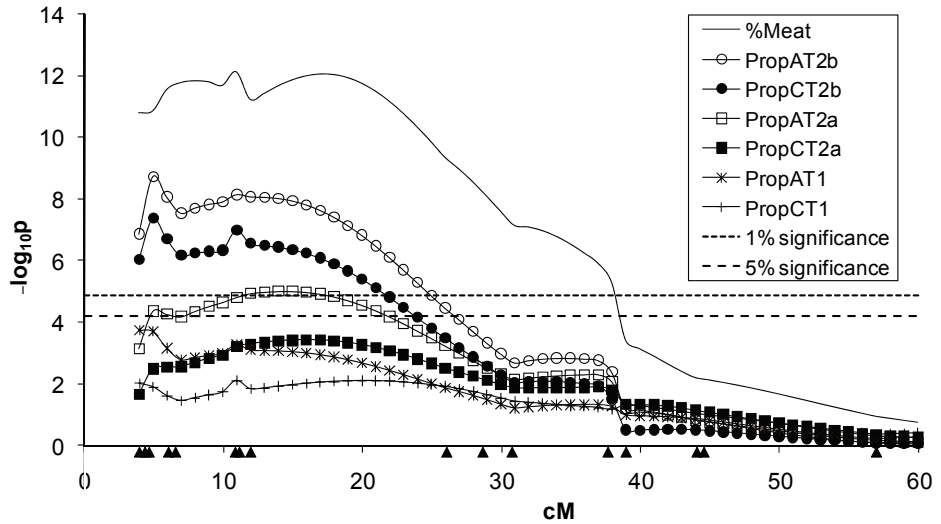


Figure 1. Haley-Knott plot for the first 60 cM of BTA2 with $-\log_{10}(p)$ values from across-family results. Trait abbreviations are Prop (Proportion), C or A (Count or Area) and T1 or T2a or T2b (Fibre Types I, IIa or IIb). Informative markers are shown (triangles) in relation to their distance (cM) along the chromosome, from left to right BY5, BULGE23, BTAFJ1, MSTN, BULGE20, ILSTS26, INRA40, TGLA431, TEXAN2, OARHH30, TGLA377, URB042, ILSTS30, OAFCB20, NEB, RM356.

The results from an ANOVA fitting additive effects (where the Limousin mutation was defined as positive) and dominance effects for the F94L genotype showed an additive effect but no dominance for 5 of the 6 fibre type traits. There was no effect on Count of type I fibres but a negative effect on Area of type I ($P < 0.01$). The Limousin allele had a large positive effect ($P < 0.0001$) on both measures for type IIb fibres and conversely a large negative effect on type IIa reflecting the high negative correlation between these phenotypes.

For the ANOVA fitting both the sire- and dam-derived haplotypes, effects for both parental alleles were again highly significant for all the type II measures. For type I, there was only a dam effect for Count ($P < 0.001$), whilst there were sire ($P < 0.01$) and dam effects ($P < 0.05$) for Area.

The re-runs of the Haley-Knott scans fitting, in addition, the myostatin F94L genotype reduced the test statistics to below the 5% significance threshold across families for 5 of the 6 fibre type traits on BTA2. The exception was type I Count which was significant at the 5% threshold and this was due to one sire family. This same family also still had a suggestive QTL for type I Area and IIb Count and meat percentage. In the complete data-set, this sire has more QTL for muscling and fatness than the other 5 sires – possible explanations are suggested in Morris *et al.* (2009).

DISCUSSION

Differentiation of muscle fibre type occurs during the foetal stage in mammals but it is possible that their relative proportions may be modified after birth by exercise. The J-backcross animals in this trial were hand-reared whilst the L-backcross animals were reared on their dams; both backcrosses grazed together from about 6 months of age in sex mobs until slaughter.

Highly double-muscled animals in breeds such as Belgian Blue and Piedmontese have been shown to have more type IIb fibres and less type I fibres, and this partly explains the overall increase in the whiteness of meat from such animals (Holmes and Ashmore 1972). The *M. longissimus* is a darker muscle than the *M. semitendinosus* and this is probably due to a higher proportion of the white type IIb fibres in the *M. semitendinosus*.

Our QTL trial with J and L breeds did not show QTL around myostatin for striploin (*M. longissimus*) percentage (Sellick *et al.* 2007) or striploin tenderness (Esmailizadeh *et al*. 2008) in Australia or NZ. However, in Australia, where tenderness was also measured on the *M. semitendinosus*, there were QTL in the myostatin region for both tenderness (Esmailizadeh 2006) and silverside weight as a percentage of carcass weight. This difference may be explained by differences in the proportion of type II fibres between the two muscles. Meat percentage in the whole carcass has a correlation of 0.35 with the proportion of type IIb fibres in the *M. longissimus* (and negative with the other two types); this correlation may be higher, for example, if calculated with the proportion of type IIb fibres in the *M. semitendinosus*.

The lack of any dominance in the effect of the F94L mutation on fibre types in the *M. longissimus* as opposed to a relatively large negative dominance effect on meat percentage and silverside percentage (Sellick *et al.* 2007) suggests that the three NZ sires were similar for other gene effects on fibre type.

Myosin genes also have a role in muscle development and some myosin genes appear under the other QTL reported here.

In conclusion, it appears that the F94L mutation found in Limousin cattle does have an effect on fibre type distributions in the *M. longissimus*, though its effect may be less than that of the inactivating mutations found in the Belgian Blue and Piedmontese breeds.

ACKNOWLEDGEMENTS

This work was funded by the New Zealand Foundation for Research, Science and Technology. We thank Gabby Sellick for her sequencing and genotyping of the myostatin F94L SNP.

REFERENCES

Esmailizadeh, A. (2006) PhD Thesis, University of Adelaide.

Esmailizadeh, A., Bottema, C.D.K., Sellick, G.S., Verbyla, A.P., Morris, C.A., Cullen, N.G. and Pitchford, W.S. (2008) *J. Anim. Sci.* **86**:1038.

Holmes, J.H.G. and Ashmore, C.R. (1972) *Growth* **36**:351.

Ihara, N., Tagasuga, A., Mizoshita, K., Takeda, H., Sugimoto, M., Mizoguchi, Y., Hirano, T., Itoh, T., Watanabe, T., Reed, K.M., Snelling, W.M., Kappes, S.M., Beattie, C.W., Bennett, G.L. and Sugimoto, Y. (2004) G*enome Res.* **14**:1987.

Klont, R.E., Brocks, L. and Eikelenboom, G. (1998) Muscle fibre type and meat quality. *Meat Sci.* **49** (Suppl. 1):S219.

Knott, S.A., Elsen, J.M. and Haley, C.S. (1996) *Theoret Appl. Genet.* **93**:71.

Martyn, J.K., Bass, J.J. and Oldham, J.M. (2004) *The Anat. Rec. Part A.* **281A**:1363.

Morris, C.A., Pitchford, W.S., Cullen, N.G., Esmailizadeh, A.K., Hickey, S.M., Hyndman, D.L., Dodds, K.G., Afolayan, R.A., Crawford, A.M. and Bottema, C.D.K. (2009) *Anim. Genet.* (in press).

Sellick, G.S., Pitchford, W.S., Morris, C.A., Cullen, N.G., Crawford, A.M., Raadsma, H.W. and Bottema, C.D.K. (2007) *Anim. Genet.* **38**:440.

LEPTIN GENE POLYMORPHISMS HAD NO EFFECTS ON OPEN DAYS AND CALVING INTERVAL

H. Yazdani[1], H. R. Rahmani[1], M. A. Edris[1] and E. Dirandeh[2]

[1] Department of Animal Science, Faculty of Agriculture, Isfahan University of Technology, Isfahan, Iran
[2] Department of Animal Science, Faculty College of Agriculture and Natural Resources, University of Tehran, Karaj, Iran

SUMMARY

Leptin plays an important role in the regulation of feed intake, energy metabolism, growth and reproduction of cattle. We used the polymerase chain reaction-restriction fragment length polymorphism (PCR-RFLP) technique to screen for DNA polymorphisms of the leptin gene in 255 cows of Iranian Holstein. Amplified region is located in exon three of leptin gene. The genomic bovine leptin sequences, which consist of three exons, were obtained from Gene Bank (Accession number U50365). Genotype frequencies were 0.588, 0.388 and 0.024 for AA, AB and BB respectively and allelic frequencies were 0.782 and 0.218 for A and B, respectively. We investigated effect of A59V polymorphism in the leptin gene on two reproduction traits. Significances of the genotype effects were tested using approximated F-statistic (SAS GLM procedure). This study showed that genotype had no effect on open days and calving interval (NS). Animals with the AB genotype had higher open days and calving interval.

INTRODUCTION

Leptin is a 16-kDa polypeptide hormone synthesized and secreted predominantly by Adipose tissue. It functions regulating body weight, food Intake, energy expenditure, reproduction and immune system functions. Leptin was first identified, as gene product found deficient in obese (ob/ob) mice. A single base mutation of the leptin gene at the codon 105, as observed in the ob/ob mouse involved C/T mutation and replacement of arginine by a premature stop codon and a subsequent production of an inactive form of leptin (Zhang *et al.* 1994). The genetically obese ob/ob mouse exhibits obesity, infertility, hyperglycemia, impaired thyroid function and hyperinsulinemia with insulin resistance (Dubuc 1976). Treatments of the ob/ob mice with recombinant leptin reduce feeding and body weights (Halaas *et al.* 1995). Leptin treatment of animals has been shown to cause a decrease in food intake, body weight loss, fat deposit weight loss and increase in energy metabolism, therefore leptin not only causes reduced food intake, but the potential body weight losses are enhanced due to an increased metabolic rate (Houseknecht *et al.* 1998). The gene encoding Leptin was mapped to Bovine chromosome 4 and it consists 3 exons and 2 introns of which only two exons are translated into protein. The coding region of the leptin gene (501 nucleotides in length) is contained in exon 2 and 3, which are separated by intron of approximately 2 kb. Leptin is considered as a candidate gene for milk performance related traits in cattle. Several polymorphism in this gene have been found (Liefers *et al*. 2002). In exon three, A59V polymorphism, causes as amino acid change from alanine to valine. These amino acids both belong to the group of aliphatic amino acids, but valine is more hydrophobic. Aim of this study is to analyze A59V Polymorphism in exon 3 of leptin gene in Iranian Holstein cattle.

MATERIALS AND METHODS

Blood was collected from 255 Holstein cattles of four different herd managements in Isfahan province. Simple statistics is presented in Table 1.Genomic DNA extracted from whole blood. Genotypes of A59V polymorphism were identified with PCR-RFLP technique. Amplified region

is located in exon three of Leptin gene. The genomic bovine leptin sequences, which consist of three exons, were obtained from GeneBank (Accession number U50365). The polymerase chain reaction was used to amplify the 331 bp DNA fragments from genomic DNA. The PCR reaction contained 100 ng of genomic DNA, 0.3 µ*M* of each primer, 1.5 m*M* $MgCl_2$, 200 µ*M* dNTP, 10m*M* Tris HCl, 50 m*M* KCl and 1 U *Taq*-polymerase in total volume of 20 µL. Sequences of primers that were used in PCR were reported previously by Haegeman et al. (2000). The sequence of the forward and reverse primers, respectively were: 5´-GGG AAG GGC AGA AAG ATA G-3´ and 5´-TGG CGA ACT GTT GAG GAT C-3´. Conditions for PCR were 94°C for 2 min, followed by 35 cycles of 94°C for 30 s, 55°C for 1 min, and 72°C for 30 s. Followed by final extension for at for 15 min 72°C. Digestion of PCR product of 331 bp with with 5 U of H*ph*I (Fermentas) in 20 µL of reaction volume at 37°c for 8 h and analysed on 8% nondenature polyacrylamyde gel. Allele A in the A59V polymorphism was the allele not digested by restriction enzyme, allele B was the restriction enzyme-digested PCR product. Digestion revealed 3 genotypes, AA (331 bp), AB (331, 311, and 20 bp), and BB (311 and 20 bp). The PCR reaction result in an artefact band of 600 bp which did not interfere it was not digested by the enzyme. The allele and genotype frequencies of A59V polymorphism were examined for deviation from Hardy-Weinberg equilibrium using $\chi 2$ test. Data were analyzed using PROC GLM of SAS (2000), using the following linear model:

$$Y_{ijk} = \mu + G_i + H_j + b_1(X_{ijk} - \overline{X}) + b_2(Z_{ijk} - \overline{Z}) + e_{ijk}$$

Where Y_{ijklmn} = traits measured on each animal

μ = Overall mean

G_i = Effect of Genotype

H_j = Effect of herd

X = Effect of dry period

b_1 = Linear regression for dry period trait

Z = Effect of Lactation period

b_2 = Linear regression for lactation period trait

e_{ijk} = Residual error

RESULTS AND DISCUSSION

Genotype and allele frequencies of the A59V polymorphism are listed in Table 2. Genotype frequencies in all herds were 0.588, 0.388 and 0.024 for AA, AB and BB, respectively and allelic frequencies were 0.782 and 0.218 for A and B, respectively. Allelic frequency analysis have shown that frequencies ranged from 0.759 to 0.824 for allele A and 0.176 to 0.241 for allele B in all herds. In 2th herd we did not found BB genotype. The genotype frequencies were distributed according to Hardy-Weinberg equilibrium proportions in every four herd but were not in all herds ($p < 0.05$). Open days was higher in heterozygous genotype. Calving interval did not differ among genotypes and was less in BB animals than in heterozygotes (Table 3). Woodside et al. (1998) showed that leptin influences the length of the anestrus period in rats suffering from severe negative energy balance due to food deprivation. As high producing cows also suffer from a negative energy balance due to lactation, leptin may influence the postpartum anestrus period in early-lactating cows.

Our findings for A59V polymorphism in bovine leptin gene are similar to those of Hanna Kulig (2005) who reported A and B allele frequencies of 0.760 and 0.240, respectively. Liefers *et al.* (2002) found a frequency of 0.747 for the A allele and of 0.254 for the B allele Result show that

allele B has lower frequency in all study. Nassiry et al (2008) reported that allele C in Sarabi, Taleshi, Sistani, Golpayegani, Brown Swiss and Holstein cattle with 68, 55, 69, 71, 55 and 57% value were the most frequent alleles. Observed heterozygosities were highest in Golpayegani (57.89%).

Two days of total feed restriction in 11–12-month-old heifers markedly reduced leptin mRNA in adipose tissue, as well as circulating concentrations of leptin IGF-I, and insulin, and reduced the frequency of LH pulses compared to controls (Amstalden et al., 2000). In contrast to the prepubertal heifer, short-term fasting (60 h) did not attenuate pulsatile LH release in the mature cow. Central administration of leptin increased plasma LH in fasted but not in control-fed cows (Amstalden et al., 2002). Short-term (72-h) fasting and fasting-mediated reductions in LH pulse frequency were attenuated by peripherally administered recombinant leptin (Nagatani et al., 2000).

Table 1. **Mean ± SD and CV for open days and calving interval**

Trait	N	Mean ± SD	Min	Max	CV
Open days	248	142.69 ± 4.05	40	393.5	44.62
Calving interval	248	418.35 ± 4.14	303.67	666.5	15.57

Table 2. Genotype and allele frequencies of the A59V polymorphism

Herd	N	Genotype			Allele	
		AA	AB	BB	A	B
1	58	0.552	0.414	0.034	0.759	0.241
2	54	0.648	0.352	0.000	0.824	0.176
3	40	0.650	0.325	0.025	0.813	0.188
4	103	0.553	0.417	0.029	0.762	0.238
Total	255	0.588	0.388	0.024	0.782	0.218

Table 3. Effect of the A59V polymorphism on reproductive traits

Trait	Genotype		
	AA	AB	BB
Calving interval	407.51±5.6[a]	416.05±6.8 [a]	394.3±26.3 [a]
Open days	128.15±6.1 [a]	135.72±7.3 [a]	119.71±28.6 [a]

a.b Least squares means within a row without a common superscript letter differ, $P < 0.05$.

CONCLUSIONS

This polymorphism could be further evaluated for marker assisted selection. Polymorphisms had no effected on open days and calving interval (NS). For finding the evolutionary relationships among closely populations, Leptin is a suitable and informative marker system.

ACKNOWLEDGMENTS

This research was supported by Department of Animal Science, Faculty of Agriculture, Isfahan University of Technology, Isfahan, Iran.

REFERENCES

Amstalden, M., Garcia, M. R., Stanko, R. L., Nizielski, S. E., Morrison, C. D., Keisler, D. H. and Dubuc, P. U. (1976). *Metabolism*. **25**:1567.

Haegeman, A., van Zeveren, A. and Peelman, L. J. (2000). *Anim. Genetics.* **31:**79.

Halaas, J. L., Gajiwala, K. S., Maffei, M., Cohen, S. L., Chait, B. T., Rabinowitz, D., Lallone, R. L., Burley, S.K. and Friedman, J. M. (1995). *Science*. **269**:543.

Houseknecht, K. L., Baile, C. A., RMatteri, R. L. and Spurlock, M. C. (1998). *J. Anim. Sci.* **76:** 1405.

Kulig, H. (2005). *Arch. Tierz Dummerstorf.* **48:**547.

Liefers, S. C., te Pas, M. F., Veerkamp, W. R. F. and van der Lende, T. (2002). *J. Dairy. Sci.* **85:** *1633.*

Nagatani, S., Zeng, Y., Keisler, D. H., Foster, D. L,, Jaffe, C. A. (2000). *Endocrinol.* **141**:3965

Nassiry, M. R., Eftekhari Shahroudi, F., Mousavi, A. H., Sadeghi, B. and Javadmanesh A. (2008). *African J. Biotech.* 7 (15), 2685

Williams, G. L. (2002). *Biol Reprod*. **66**:1555

Woodside, B., A. Abizaid, and S. Jafferali. 1998. *Am. J. Physiol*. **274**:1653.

SAS. (2000) User'r Guide Statistics. SAS Institute Inc., Cary, North Caroline.

Zhang,Y., Proenca, R., Maffei, M., Barone, M., Leopold, L. and Friedman, J. M. (1994). *Nature.* **372**: 425.

POTENTIAL TO DOUBLE THE NUMBER OF CATTLE TESTED FOR RESIDUAL FEED INTAKE

K. A. Donoghue[1], P. F. Arthur[2] and R. M. Herd[3]

Cooperative Research Centre for Beef Genetic Technologies
[1]NSW Department of Primary Industries, Agricultural Research Centre, Trangie, NSW 2823
[2]NSW Department of Primary Industries, Elizabeth Macarthur Agricultural Institute, Camden, NSW, 2570
[3]NSW Department of Primary Industries, Beef Industry Centre, Armidale, NSW 2351

SUMMARY

The effect of foregoing daily feed intake records when cattle were rotated weekly between pens with or without functional feed intake recorders on the estimate of residual feed intake over a 70-day test period was investigated. Over the test period, animals (n=48) were rotated on a week-on/week-off basis from recording feeders (week-on) to non-recording feeders (week-off). In addition, weekly weights were taken on all animals. There was no evidence of re-adaptation to feeders during recording week or significant differences in average daily gain between the recording and non-recording weeks. Comparisons between this test and previous tests where feed intake was continuously recorded indicated no significant differences in the amount of variation in feed intake explained by average daily gain and metabolic weight. These results indicate that a week-on/week-off strategy, when implemented as in this study, has potential to double the number of animals measured through a feed intake recording system without comprising data collected for RFI computation. Genetic improvement programs which include feed efficiency in their breeding objective will benefit from increased number of records available.

INTRODUCTION

Residual feed intake (RFI) is the amount of feed eaten, net of the requirements for maintenance and production. Inclusion of this trait in selection indices allows beef producers to make genetic improvement in feed efficiency. While previous studies have indicated that a test length of 35 days is adequate to measure feed intake, a longer test period (70 days) is required to accurately measure growth; thus the recommended test length to measure RFI is 70 days (Archer *et al.* 1997; Exton 2001). The number of animals that can undergo a continuously recorded 70-day feed intake test is often limited by the capacity of the feed intake system, and ways to increase the number of animals tested are keenly sought. One such possibility is the periodic recording of daily feed intake (DFI) through a rotation of animals on a week-on/week-off basis, such that animals have DFI records for half of the test period (35 days), but still have growth records for the full 70 days. This strategy would double the number of animals that could be measured through a feed intake recording system and would increase the amount of information available for genetic improvement programs for feed efficiency. The objective of this study was to examine the impact of foregoing DFI records when cattle were rotated on a week-on/week-off from feed intake recorders on the estimate of RFI over a 70-day test period.

MATERIALS AND METHODS

Data. Forty-eight Angus heifers underwent a postweaning RFI test at the Agricultural Research Centre, Trangie, NSW. These heifers were born in 2007 and were part of the Trangie RFI divergent selection lines. Details of the postweaning test procedure are given in Arthur *et al.* (2001a) and establishment of the divergent selection lines in Arthur *et al.* (2001b). The heifers

were randomly split into four pens with 12 heifers in each pen. Two of the pens each contained an automated feed intake recorder (described by Bindon 2001) which recorded individual feed intakes. The other two pens each contained an identical but disabled feed intake recorder. Heifers were rotated on a week-on/week-off basis between pens with recording feeders (week-on) and pens with non-recording feeders (week-off) such that a total of 35 days of individual feed intakes were available. Heifers were weighed weekly for the duration of the test.

After an adjustment period of 21 days, the heifers commenced a test of 70 days duration. A pelleted diet composed of 70% lucerne hay and 30% grain, with approximately 10.5MJ metabolisable energy (ME) per kg dry matter and 17% crude protein was fed. All heifers were given *ad libitum* access to this feed, and a daily allowance of 0.5kg/animal of wheaten straw (approximately 5.2MJ ME /kg dry matter) was provided.

Statistical analysis. Growth of the heifers was modelled by linear regression of weekly weight against time with the regression coefficients used as the average daily gain (ADG) for each animal (SAS Institute 2003). A separate regression was fitted for each animal. The weight of each animal at the mid-point of the test period raised to the power 0.75 (metabolic mid-weight) and ADG were used in a multiple regression with daily feed intake (DFI) as the dependent variable. RFI was equated to the residual error term in the model.

Tests of significance for ADG in recording and non-recording weeks were undertaken to examine whether heifers were experiencing different rates of gain in recording vs. non-recording weeks. For this test, individual ADG were averaged over recording and non-recording weeks, and a pair-wise t-test of significance was calculated. Tests of significance for DFI during days of recording weeks were undertaken to examine whether heifers were experiencing re-adaptation to the recording feeders; that is, eating significantly less feed when first re-introduced to the recording feeders. For these tests, DFI on days 1 to 7 of each recording week were averaged for every animal. Pair-wise t-tests of significance were calculated for DFI on day 1 vs. days 2-7; days 1-2 vs. days 3-7 ; and days 1-3 vs. days 4-7.

The number of heifers with low feed intake on each day throughout the test was also examined in order to observe whether re-adaptation to the recording feeders was occurring. An animal was defined as having a low feed intake on a particular day if their DFI was lower than 1.5% of their weekly liveweight. An adjusted R^2 value was obtained from the regression of ADG and metabolic mid-weight on DFI in this study, and was compared with values from recent tests in the amount of variation in DFI explained by ADG and MWT.

RESULTS AND DISCUSSION

Data. Descriptive statistics of the data are contained in Table 1. Significant variation in the heifers was observed for all traits, including ADG (range of 1.1 kg/day) and RFI (range of 4.5 kg/day).

Table 1. Descriptive statistics of the data

Trait	Mean	SD	Minimum	Maximum
Start weight, kg	304	36	230	417
End weight, kg	401	44	300	539
Metabolic mid-weight, kg	72	5.8	59	90
Daily feed intake, kg	10.54	1.56	7.95	14.63
Average daily gain, kg/d	1.42	0.22	0.91	2.01
Residual feed intake, kg/d	-0.11	1.07	-2.03	2.48

ADG did not differ significantly (P = 0.89) between recording (1.50 kg/day) and non-recording (1.51 kg/day) weeks. This result indicates that heifers performed similarly with respect to weight gain during recording and non-recording periods, and that the method of periodic recording of feed intake had no impact on weight gain throughout the test used in this study.

There was no evidence of significant differences (P>0.05) between individual DFI early in recording weeks (days 1, 2 or 3) and individual DFI in the second half of the recording weeks (Table 2). This result indicates that heifers were not experiencing re-adaptation each time they rotated to the recording feeders, and that the method of periodic recording of feed intake had no impact on DFI throughout the test. Further confirmation of this finding is evident in Figure 1, where there was no evidence of greater number of animals having low feed intake on the first day of each recording week (Days 1, 8, 15, 22 and 29) compared to other days throughout the test.

Table 2. The t-tests of significance for difference in feed intake between days of week

Comparison	P value
Day 1 vs. Days 2-7	0.14
Days 1-2 vs. Days 3-7	0.22
Days 1-3 vs. Days 4-7	0.38

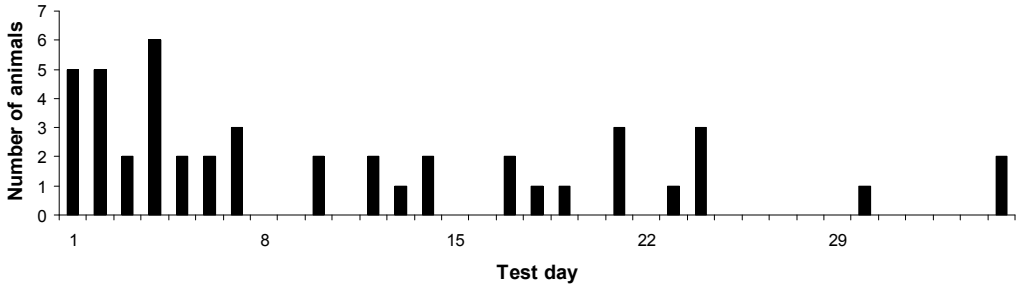


Figure 1. Number of animals with feed intake below 1.5% of liveweight

These results are in contrast to a study in pigs undergoing a similar week-on/week-off feed intake test (Schulze *et al.* 2001) where animals experienced "re-adaptation" to the feeders each recording week. This re-adaptation was manifested such that the periodically recorded pigs had a substantially higher percentage of animals with low feed consumption at the beginning of each recording week, in comparison with continuously recorded pigs. In our study, the only difference between the recording and non-recording feeders was the absence of a small electronic control box on the back (and out of view of the animal) of the non-recording feeders. Thus, heifers experienced a very similar feeding environment in recording and non-recording weeks. In the pig study, however, the feeding environments between recording (electronic feeders) and non-recording (conventional feed dispenser) weeks were quite different (Schulze *et al.* 2001), which may explain the different findings.

The amount of variation in DFI explained by ADG and metabolic mid-weight in this test (0.54; see Table 3) was similar to that for other recent RFI tests in which animals were continuously recorded for DFI (0.46-0.66; see Table 3). This would indicate that periodic recording of DFI did not compromise data collected for RFI computation, and that the amount of variation in RFI

explained by ADG and DFI was similar to that observed in studies where DFI was recorded continuously.

Table 3. Comparison of adjusted R^2 values for RFI tests

Animal source	Year of test	Number of animals	Adjusted R^2
Trangie bulls	2000	89	0.63
Trangie heifers	2000	92	0.46
Industry steers	2007	35	0.66
Industry heifers	2007	23	0.51
Industry bulls	2007	37	0.64
Trangie bulls	2008	102	0.59
This study	2008	48	0.54

CONCLUSIONS

Foregoing daily feed intake records when cattle were rotated on a week-on/week-off basis between pens with functional and non-functional feed intake recorders had no impact on the estimate of RFI over a 70-day test period. There was no evidence of re-adaptation to feeders during recording weeks or for significant differences in ADG between the recording and non-recording weeks. These results indicate a week-on/week-off strategy, when implemented as in this study, has potential to double the number of animals measured through a feed intake recording system without comprising data collected for RFI computation. Genetic improvement programs including feed efficiency in their breeding objective would benefit from the increased amount of information available on the trait.

ACKNOWLEDGMENTS

This work was funded by the Cooperative Research Centre for Beef Genetic Technologies. The capable assistance of D. Mula, K. Dibley and T. Snelgar is gratefully acknowledged.

REFERENCES

Archer, J.A., Arthur, P.F., Herd, R.M., Parnell, P.F. and Pitchford, W.S. (1997) *J. Anim. Sci.* **75**: 2024.

Arthur, P.F., Archer, J.A., Johnston, D.J., Herd, R.M., Richardson, E.C. and Parnell, P.F. (2001a) *J. Anim. Sci.* **79**: 2805.

Arthur, P. F., Archer, J.A., Herd, R.M. and Melville, G.J. (2001b) *Proc. Assoc. Advmt. Anim. Breed. Genet*. **14**: 135.

Bindon, B.M. (2001) *Aust. J. Exp. Ag.* **41**:843

Exton, S. (2001) Testing beef cattle for net feed efficiency – standards manual. Edition March 2001. Performance Beef Breeders Association, Armidale, NSW.

SAS Institute (2003). 'ONLINE Documentation for SAS/STAT Procedures.' (SAS Institute: Carey, NC.)

Schulze, V., Roehe, R., Looft, H. and Kalm, E. (2001) *J. Anim. Breed. Genet.* **118**:403.

FAT DISTRIBUTION IN ANGUS STEERS IS RELATED TO RESIDUAL FEED INTAKE ESTIMATED BREEDING VALUE

A.R. Egarr[1], W.S. Pitchford[1], M.J. Bottema[2], R.M. Herd[3], J.P. Siddell[3], J.M. Thompson[4] and C.D.K. Bottema[1]

Cooperative Research Centre for Beef Genetic Technologies
[1]Animal Science, University of Adelaide, Roseworthy, S.A. 5371
[2]School of Informatics and Engineering, Flinders University, Bedford Park, S.A. 5042
[3]NSW DPI Beef Industry Centre, Armidale, NSW, 2351
[4]Division of Animal Science, University of New England, Armidale, NSW, 2351

SUMMARY

The association between fat distribution and feed efficiency was investigated using Angus steers divergent in residual feed intake (RFI) estimated breeding values (EBV). The 208 steers were fed in a commercial feedlot in NSW for 250 days, entering at 13 – 16 months of age. Hot standard carcass weight, eye muscle area, marble score, intramuscular fat content, rib fat (subcutaneous) depth and seam (intermuscular) fat area were measured and regressed against the mid-parent RFI EBV of the steers. The results showed that rib fat depth was more strongly associated with RFI EBV than were the other fat depots.

INTRODUCTION

Feed costs comprise a significant part of the cost of beef production. In order to reduce these costs, research has been undertaken to identify cattle with genetically higher net feed efficiency measured as low residual feed intake (RFI), i.e. animals that consume less feed (energy) than would be expected for an animal at a particular live weight and growth rate. However, the impact on individual fat depots when breeding for increased net feed efficiency warrants further investigation as there have been reports of an association between feed efficiency and fat deposition (Richardson *et al.* 2001; Basarab *et al*. 2003; Bulle *et al*. 2007).

Pitchford (2004) reviewed selection for feed efficiency in other species and concluded that 'while improvements in feed efficiency can be made, they come at a cost. In poultry and mice, increased net feed efficiency has been associated with increased fatness but in pigs and beef cattle there is some evidence of the reverse. While a number of studies have predicted that selection for efficiency may result in lower proportions of crucial, metabolically active organs, there are few studies to support this.' Richardson *et al.* (2001) determined that cattle with genetically lower RFI also had a lower proportion of body fat, but there was no difference in subcutaneous fat or eye muscle area (EMA). Basarab *et al*. (2003) found a similar phenotypic relationship between RFI and general fatness, with an increase in RFI associated with increased marble score, removable fat and subcutaneous fat, but no relationship with muscle area (*longissimus thoracis)*. In an intensive trial using 24 steers, Bulle *et al.* (2007) demonstrated that phenotypically low RFI steers tended to gain less fat than their high RFI counterparts, yet there was no significant difference in carcass weight, marbling score, subcutaneous fat or internal fat between high and low RFI animals. This was also the finding of Baker *et al.* (2006).

The purpose of this work was to examine potential associations between genetic variation in RFI and phenotypic data for different fat depots (subcutaneous, intramuscular and intermuscular fat), in comparison to muscle traits, in a larger number of steers fed in a commercial feedlot trial.

MATERIALS AND METHODS

Cattle. 208 Angus steers from the NSW Department of Primary Industries Trangie RFI selection line with a large divergence in mid-parent residual feed intake estimated breeding values (RFI EBVs) were used in this trial to show that differences in residual feed intake can be achieved in a commercial setting. The steers were progeny of 26 sires with numbers of progeny ranging from 1 to 21. The feedlot trial comprised 3 pens, with steers allocated to the pens based on mid-parent RFI EBV (low RFI EBV = -0.85 to -0.52, n = 68; medium RFI EBV = -0.29 to 0.14, n = 72; high RFI EBV = 0.16 to 0.98, n = 68) and fed for 250 days in a large commercial feedlot in NSW. The steers had been managed together from birth until they entered the feedlot. Age at feedlot entry ranged from 13 to 16 months. Each pen was supplied with the same ration, *ad libitum,* adjusted weekly for over/under feeding. All steers were slaughtered on the same day.

Data. Aus-meat abattoir measurements and Meat Standards Australia chiller assessments were performed on the carcasses by accredited graders (Table 1). Traits included hot standard carcass weight (HSCW), rib fat thickness (RFT), eye muscle area (EMA) and marble score (MSAMB). EMA and MSAMB were assessed at the $5^{th}/6^{th}$ rib. In addition, a 15mm thick sample incorporating the *M. longissimus dorsi, Spinalis dorsi and Semi-spinalis dorsi* muscles was collected for calculation of seam fat area (SF) via image analysis and measurement of intramuscular fat content (IMF%) via chemical extraction according to the protocol described by Siebert *et al.* (2006).

Image analysis. All samples were photographed and images stored as separate jpeg files with a resolution of 180dpi. Subsequent image manipulation involved using Adobe® Photoshop® CS2 to trim the seam fat from the surrounding muscles and saving these as individual images, again at 180dpi. The seam fat area (SF) was measured using Matlab R2007a, an interactive software system.

Data analysis. All data were analysed using general linear regression (GenStat 10.1) with effects of age at slaughter (range 646 to 746 days), RFI EBV (range -0.85 to +0.98 kg/d), pen (high, medium, low), and the RFI EBV by pen interaction. Regression coefficients of RFI EBV are reported and percent change in each trait per kg/day change in RFI EBV was calculated by dividing the regression coefficient by the mean of that trait. Significance was defined as $P<0.05$.

RESULTS AND DISCUSSION

The range in mid-parent RFI EBV was 1.83kg/day. Substantial variation was observed for all traits (Table 1), but was most noticeable for rib fat depth, seam fat area and marble score.

Table 1. Summary of trait data.

	HSCW (kg)	*EMA (cm^2)*	*RFT (mm)*	*SF (cm^2)*	*IMF%*	*MSAMB*
Mean	415	77	17.9	24.2	14.5	504
Min	354	68	6	9.67	8.29	350
Max	494	85	34	44.2	22.7	830
St Dev	27.4	3.28	5.62	6.21	3.12	107

HSCW = hot standard carcass weight, EMA = eye muscle area, RFT = rib fat thickness, SF = seam fat area, IMF% = percent of chemically extracted intramuscular fat, MSAMB = Meat Standards Australia marble score.

Because of the experimental design, pen was completely confounded with RFI EBV group. However, as there was variation in the mid-parent RFI EBVs within the pens, each trait was regressed against the mid-parent EBV, pen, and the mid-parent EBV by pen interaction. Mid-parent EBV and pen were significant for most fat and muscle traits (Table 2). Mid-parent EBV by pen interaction was marginally significant for the eye muscle area but not for any of the other traits and was removed from the model (Table 2).

Of the 26 sires of steers in this trial, 10 had progeny in two pens. Therefore, an additional model was fitted with sire as random effect. This only reduced the significance levels slightly and did not remove the pen effect. Therefore, the mixed model results are not reported herein.

Table 2. Tests of significance for traits (F-probabilities).

	HSCW	***EMA***	***Rib fat***	***Seam fat***	***IMF %***	***MSAMB***
Model 1: final age + mid-parent RFI EBV + pen + mid-parent RFI EBV x pen interaction						
RFI EBV	0.026	0.550	<0.001	0.013	0.093	0.146
Pen	<0.01	<0.001	0.363	0.018	<0.001	<0.001
RFI EBV x pen	0.078	0.039	0.885	0.447	0.889	0.248
Model 2: final age + mid-parent RFI EBV and final age + pen						
RFI EBV	0.030	0.579	<0.001	0.015	0.103	0.191
Pen	0.004	<0.001	<0.001	0.001	<0.001	<0.001

Terms as defined in Table 1.

The relationship between mid-parent RFI EBV and rib fat was highly significant (<0.001); with HSCW and seam fat marginally significant (0.030 and 0.015, respectively), while other traits had no significant association with mid-parent RFI EBV (Table 2, Model 2). This result does not agree with the work by Bulle *et al.* (2007), who found no significant differences in subcutaneous fat or carcass weight. However, it is consistent with the subcutaneous fat results of Basarab *et al*. (2003) and Richardson *et al.* (2001). There are a number of differences in these experiments that might account for the conflicting results. Two of the previous experiments involved crossbred steers which may have influenced the result, as the genetics underlying the variation in RFI may differ between breeds. Indeed, Basarab *et al.* (2003) selected animals from five foundation breeds because of the large variation in genetic backgrounds. In addition to this, there was variation in the age at feedlot entry, length of time on feed and the number of animals in each trial. The trial reported here involved a larger number of steers that were fed a high energy diet for a longer period than the previous studies. In addition, the steers were older than all previous studies with the exception of the animals in the Bulle *et al.* (2007) trial. The variation in the results may reflect the added maturity of the steers and the longer time on a high energy diet, both of which are likely to increase the amount of fat deposited and therefore, increase the variation between the animals, as has been discussed in a review by Herd and Arthur (2009).

As pen was significant for the traits measured in this trial, it would suggest that there were important environmental effects involved. The most likely explanation in this instance would be differences in the feeding regimens between the pens. Therefore, careful analysis and interpretation of the data is required in order to draw conclusions regarding body composition and selection for RFI.

The regression coefficients (Table 3) indicate the measured change in the traits that was observed with a one kilogram/day increase in RFI EBV. The percentage change enables a comparison of the magnitude of the effect on each trait from the genetic change in RFI under the conditions of this experiment. The association between RFI EBV and the fat traits implies that

selection for lower RFI would decrease fatness. However, the change in rib fat thickness was much larger than the change in seam fat area, and the change in intramuscular fat content was statistically insignificant. This suggests that while improving (decreasing) RFI could reduce fatness, the magnitude of the effect differs between the adipose depots.

Table 3. Regression coefficients of RFI EBV with the standard errors and percent changes of traits.

	HSCW	***EMA***	***Rib fat***	***Seam fat***	***IMF %***	***MSAMB***
Regression coefficient	19.9 ± 12.0	1.04 ± 1.44	7.97 ± 2.4	2.17 ± 2.85	2.12 ±1.39	42.9 ± 44.5
Percentage change (% trait / unit RFI EBV)	5	1	45	9	15	9

Terms as defined in Table 1.

CONCLUSIONS

This experiment showed that rib fat thickness and seam fat area were associated with RFI EBV in these Angus steers. Rib fat thickness often has tight market specification and showed the most variation linked with genetic change in RFI. There was a significant relationship between genetic variation in RFI and area of seam fat, a previously unreported fatness trait, but the magnitude of the association was not nearly as large as that for rib fat thickness. The results confirm previously reported associations between improved RFI and decreased fatness. However, the results also show that substantial changes in fat deposition following selection for RFI will not necessarily affect all adipose depots equally. Reducing residual feed intake, and therefore, reducing a major cost of production, may change carcass fat composition but the consequences may not be as severe as previously thought as not all fat depots appear to be affected.

ACKNOWLEDGEMENTS

The authors thank the commercial feedlot for their collaboration and Stephen Lee, Xuemei Han, Yizhou Chen and Kim Quinn for assistance with sample collection.

REFERENCES

Baker, S.D., Szasz, J.I., Klein, T.A., Kuber, P.S., Hunt, C.W., Glaze Jr, J.B., Falk, D., Richard, R., Miller, J.C., Battaglia, R.A. and Hill, R.A. (2006) *J. Anim. Sci.* **84**:938.

Basarab, J.A., Price, M.A., Aalhus, J.L., Okine, E.K., Snelling, W.M. and Lyle, K.L. (2003) *Can. J. Anim. Sci.* **83**:189

Bulle, C.P., Paulino, V.P., Sanches, C.A. and Sainz, D.R. (2007) *J. Anim. Sci.* **85**:928.

Herd, R.M. and Arthur, P.F. (2009) *J. Anim. Sci.* **87** (E. Suppl.): E64.

Pitchford, W.S. (2004) *Aust. J. Exper. Agric.* **44**:371.

Richardson, E.C., Herd, R.M., Oddy, V.H., Thompson, J.M., Archer. J.A. and Arthur, P.F. (2001) *Aust. J. Exp. Agric.* **41**:1065.

Siebert, B. D., Kruk, Z. A., Davis, J., Pitchford, W. S., Harper, G. S., and Bottema, C. D. K. (2006) *Lipids,* **41**:365.

CHARACTERIZATION OF A HUMAN PERFORMANCE GENE IN THE HORSE

N.A. Ellis[1], P.C. Thomson[1], R.R. Coomer[2], A.J. Forhead[3], M.J. Head[4], I. Tammen[1] and H.W. Raadsma[1]

[1] Reprogen. Faculty of Veterinary Science, University of Sydney, Camden NSW Australia.
[2] Department of Clinical Veterinary Medicine, University of Cambridge, Cambridge CB3 0ES, UK.
[3] Department of Physiology, University of Cambridge, Cambridge CB2 3EG, UK.
[4] Rossdale and Partners, Beaufort Cottage Stables, Newmarket CB8 8JS, UK.

SUMMARY

Although estimates of heritability of performance in the Thoroughbred racehorse have been calculated to be between 10-40%, no genes influencing performance have been identified. Based on comparative studies with human, we identified the angiotensin-converting enzyme (ACE) gene as an equine candidate gene for performance. From characterization of nine SNPs, nine equine ACE haplotypes were identified in 34 horses. The haplotypes occurred at differing frequencies in breeds of different athletic capabilities, including endurance Arabians, Thoroughbred racehorses and Draught horses. One haplotype had a similar effect on equine ACE levels as the recognised *Alu* insertion/deletion polymorphism in the human gene. Further studies are suggested to test the association of this haplotype with a range of indices of performance in substantially larger cohorts of horses before these markers can be used for prediction of racing performance.

INTRODUCTION

The Thoroughbred racehorse has been selectively bred for its performance for hundreds of years. While it has been shown that aspects of racing performance including race times, earnings, handicaps and performance rates are up to 40% heritable (Langlois 1980; Tolley *et al*. 1983), the underlying characterization of genes contributing to this significant genetic variation is completely absent.

Angiotensin converting enzyme (ACE) is essential for the control of blood pressure. The human ACE gene contains an insertion polymorphism within intron 16, which accounts for nearly half of the variation in circulating enzyme levels (Rigat *et al.* 1990; Tiret *et al*. 1992). Associations have also been observed between the polymorphism and response to training, and elite endurance and/or sprint performance in human athletes (Gayagay *et al.* 1998; Montgomery *et al*. 1998; Myerson *et al.* 1999; Alvarez *et al*. 2000).

We present here a preliminary characterization of the equine ACE gene, whilst also demonstrating an association of equine ACE haplotypes with circulating enzyme levels.

MATERIALS AND METHODS

Sequencing the Equine ACE gene. Since the equine ACE gene sequence was not available, primers were selected based on the aligned cDNA sequences of the human (J04144), rabbit (X62551), rat (AF201332) and chicken (L40175) ACE genes. Primers were used to amplify regions of BAC clone 801F9 (supplied by Dr Francois Piumi, INRA) which was shown to contain the gene. Direct BAC sequencing was used to obtain the 5' and 3' UTRs, while exon/intron boundaries were identified by sequencing equine cDNA and alignment with human cDNA structure.

Screening for polymorphisms. Common polymorphisms were detected by sequencing three pools of DNA from 10 Thoroughbred racehorses (TB pool), 14 endurance Arabian horses (AR pool), and 10 horses of mixed breeds (MB pool). Polymorphisms were identified by comparing the chromatograms of pooled sequence with that of a single animal, and confirmed by genotyping of animals within the pool. Common SNPs (MAF > 5%) were also typed across a panel of 40 horses: 10 each of TB, AR, Standardbred (SB) and Draught (Heavy) horses (HH). RFLP or partial sequencing (when no relevant restriction enzymes were available) was used to genotype individuals.

Association Study. Racehorses (n = 203) from the UK for which circulating ACE levels were available (Coomer *et al.* 2003) were genotyped for ACE haplotypes. Initially, only those horses with ACE levels falling further than 1 SD from the mean were typed. The remaining horses were then genotyped for the SNPs comprising one haplotype that was significantly associated with enzyme level.

Statistical analysis. Haplotypes were generated from the SNP data derived from the multi-breed panel with the program PHASE version 2.0.2 (Stephens *et al*. 2001). χ^2 tests were used to assess the association between haplotypes and ACE level, while one-way ANOVA was used to determine the size of variation accounted for by haplotypes.

RESULTS and DISCUSSION

Equine ACE gene sequence, polymorphisms and haplotypes. The genomic sequence of the equine ACE gene was derived with the exception of the central region of three large introns (18, 20 and 23). The gene structure was as seen in other mammalian species and consisted of 26 exons. The sequence is 86% conserved between the horse and human. Over 10 kb of sequence, covering 73% of the cDNA, was screened for base changes. Sixteen sequence changes were identified with 11 SNPs in non-coding sequence, three silent SNPs, and one non-synonymous SNP (exon 26 G3872A). One SNP within intron 20 was triallelic. A poly-A stretch of variable length was identified in intron 14, associated with an equine repetitive element-2 (ERE-2), but proved difficult to reliably genotype and so was not further analyzed. Nine of the SNPs were found in more than one animal and were genotyped across the 40 horse panel. Inheritance of these SNPs conformed to nine predicted haplotypes (Figure 1), which differed in frequency between different breeds (Figure 2).

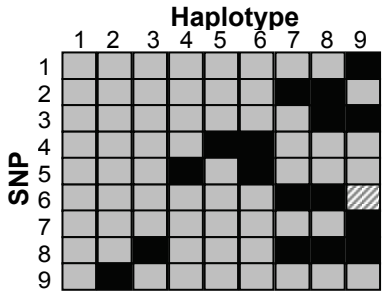


Figure 1: Nine common haplotypes found within the equine ACE gene. Grey blocks indicate the common allele, black blocks indicate the minor allele, and the striped bock in SNP 6 indicates the third allele.

Of the nine haplotypes, the haplotype represented by only common SNP alleles (haplotype 1) was the most frequent, found on 47 of the 80 possible chromosomes and in all breeds. Of the remaining haplotypes, six were observed more than once (haplotypes 2, 5, 6, 7, 8 and 9). More haplotypes with fewer representations were identified in SB and Draught horses, perhaps reflecting the higher level of genetic diversity seen in the founders and consequent generations of those populations. Haplotypes 7, 8 and 9 were seen only in the ARs and TBs, whilst haplotypes 2, 3, 4 and 5 were only found in the heavier breeds. While there appears to be significant differences between the distribution of ACE haplotypes between breeds it is difficult to say whether this is a true effect of selection, genetic drift, or sampling bias caused by the small numbers of each haplotype within each group, and the study needs to be expanded to determine this.

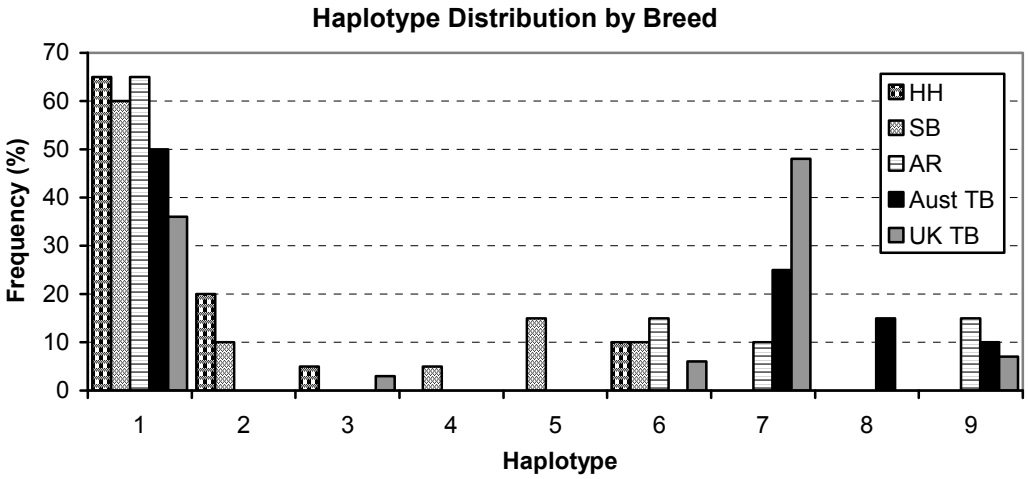


Figure 2: Distribution of ACE gene haplotypes across the UK TBs from the association study and 10 horses of different breeds: Thoroughbred, Arabian, Standardbred and Draught horses.

Association study. Sixty-two horses with circulating ACE levels representing the extreme ends of the distribution were genotyped for the nine haplotypes. Each horse had a previously determined ACE genotype: no new haplotype combinations were discovered (Figure 2). Haplotype 6 was associated with significantly lower ACE levels than the population average (P=0.007). After the remaining 141 horses in the population were genotyped for this haplotype, the association remained (P=0.000) and accounted for 10.26% of the variation in ACE concentration.

Five of the previously determined haplotypes were detected in the UK population of TBs. Haplotypes 1, 7 and 9 were common to both the Australian and UK TB populations, while 3 and 6 were found only in the UK group. This difference in haplotype distribution between the two populations may be due to local breeding trends or a simple sampling effect. It is unlikely that shuttle stallions (northern hemisphere stallions serving during the Australian breeding season) would have had the opportunity to make a large contribution to the Australian gene pool at this stage, since the majority of the horses in this study were conceived when shuttle stallions covered only 2.9 - 19% of the mare population. Local breed differences are likely to have occurred due to the popularity of certain sire lines in particular parts of the world. Again, due to the low numbers of each haplotype detected it was also impossible to determine whether the different haplotypes were associated with the different physical specialties of the breeds used, and their performances. More horses are needed to better investigate this.

The association of haplotype 6 with reduced circulating enzyme levels is similar to the documented association between the human ACE insertion allele and enzyme levels. In both cases, the polymorphisms involved are intronic and as such do not affect the coding sequence, and must influence gene expression through other methods. In both species the markers identified to date may not be responsible for the observed effects; and instead may be close to or in linkage disequilibrium with the causative polymorphisms. Alternately, the effect may be caused by changes to an unrecognized non-coding regulatory region or microRNA, either within the gene (cis acting) or elsewhere (trans acting). Further study is needed on both species to determine the mode of action of these polymorphisms.

CONCLUSION

We have characterized the equine angiotensin-converting enzyme gene, identifying nine commonly inherited haplotypes that appear to occur at different frequencies in different breeds. One haplotype was associated with circulating enzyme levels, and had a similar effect as the intronic polymorphism in the human ACE gene. Further study is needed to determine the mode of action of the equine markers, the extent of the effect on enzyme levels, and to elucidate whether there is also an association with racing performance in the horse.

ACKNOWLEDGEMENTS

This work was supported by the NSW Racing Research Fund. We are grateful to Dr. Francois Piumi, INRA, for supplying the BAC clone 801F9.

REFERENCES

Alvarez, R., Terrados, N., Ortolano, R., Iglesias-Cubero, G., Reguero, J.R., Batalla, A., Cortina, A., Fernandez-Garcia, B., Rodriguez, C., Braga, S., Alvarez, V. and Coto, E. (2000) *Eur. J. Appl. Physiol.* **82:** 117.

Coomer, R.P.C., Forhead, A.J., Bathe, A.P. and Head, M.J. (2003) *Equine Vet. J.* **35:**96.

Gayagay, G., Yu, B., Hambly, B., Boston, T., Hahn, A., Celermajer, D.S. and Trent, R.J. (1998) *Hum. Genet.* **103:**48.

Langlois, B. (1980) *Livestock Prod. Sci***. 7:**591

Montgomery, H.E., Marshall, R., Hemingway, H., Myerson, S., Clarkson, P., Dollery, C., Hayward, M., Holliman, D.E., Jubb, M., World, M., Thomas, E.L., Brynes, A.E., Saeed, N., Barnard, M., Bell, J.D., Prasad, K., Rayson, M., Talmud, P.J. and Humphries, S.E. (1998) *Nature* **393:**221.

Myerson, S., Hemingway, H., Budget, R., Martin, J., Humphries, S. and Montgomery, H. (1999) *J. Appl. Physiol.* **87:**1313.

Rigat, B., Hubert, C., Alhenc-Gelas, F., Cambien, F., Corvol, P. and Soubrier, F. (1990) *J. Clin. Invest.* **86:**1343.

Stephens, M., Smith, N. and Donnelly, P. (2001*) Am. J. Hum. Genet.* **68:**978.

Tiret, L., Rigat, B., Visvikis, S., Breda, C., Corvol, P., Cambien, F. and Soubrier, F. (1992). *Am. J. Hum. Genet.* **51:**197

Tolley, E.A., Notter, D.R. and Marlowe, T.J. (1983) *J. Anim. Sci.* **56:**1294

TARGETED MAPPING OF QTL ON CHROMOSOMES 1 AND 3 FOR PARASITE RESISTANCE IN SHEEP

N.A. Ellis[1], S.A. Kayis[1], K.J. Fullard[1], D.J. Townley[2], D. Khatkar[1], G. Attard[1], K. Beh[1], D. Piedrafita[3] and H.W. Raadsma[1]

[1] Reprogen, Faculty of Veterinary Science, University of Sydney, Camden, NSW, Australia.
[2] CSIRO Livestock Industries, 306 Carmody Rd, St Lucia. QLD, Australia.
[3] Animal Biotechnology Research Laboratories, Monash University, Clayton, VIC, Australia.

SUMMARY

The aim of this project is to identify ovine quantitative trait loci (QTL) affecting resistance and susceptibility to internal parasites. For this study, we focused on previously identified QTL regions on OAR1 and OAR3 by adding an extra 12 markers to the 15 formerly mapped on OAR1, while 14 were added to OAR3, making a total of 28. Significant QTL for FEC were identified on both chromosomes, at 188 and 184 cM respectively. Further reinforcing these findings, a meta-analysis of published QTL supported these locations. The meta-analysis was used to determine the density and location of SNPs used in the development of a 1536 custom SNP golden gate assay, for a combined LA/LD analysis.

INTRODUCTION

The cost of gastro-intestinal nematode (GIN) infection to the Australian sheep industry exceeds $200 million pa. Control of GIN has primarily focused on the use of anthelmintic drenches. However, the evolution of drench-resistant parasites, coupled with increasing consumer concern over chemical use in agricultural industries, has instigated research into alternative forms of parasite control. Certain breeds of sheep are naturally resistant to GIN (Vanimisetti *et al.* 2004), and significant genetic variation within breeds strongly indicates that genetic factors play a substantial role in host resistance to parasites. Characterization of genes responsible for variation in host resistance has been restricted to QTL and candidate gene analyses. However, few studies have identified causative mutations, primarily due to a lack of effort in fine mapping QTL. Here we present the results of progress towards fine mapping of two QTL previously identified from a whole genome scan.

MATERIALS AND METHODS

Background. Ten half sibling families (n = 694) were created by mating Indonesian Thin Tail (ITT) cross Merino (M) F1 sires to M, ITT and F1 ewes over 4 seasons. The sheep were raised and kept indoors on slatted floors and were subjected to two sequential 10 week challenges with thrice weekly administration of 2000 L3 *H. contortus* larvae for the first 3 weeks of each challenge period. Phenotypic characteristics measured were faecal egg counts (FEC), live weights (Wt) and packed cell volumes (PCV), which were recorded on the first day of the challenge and at fortnightly intervals thereafter for 10 weeks after initial challenge.

QTL analysis. FEC data was normalised using the accepted method of cube root transformation as log transformation over corrected some traits. All phenotypes were adjusted for fixed effects. QTL analysis was performed using QTLExpress (Seaton *et al.* 2002) with 5000 permutations and bootstraps for establishment of significance thresholds and QTL locations. As preliminary results from the whole genome screen (unpublished results) indicated the presence of QTL on OAR1 and OAR3, an additional 12 and 14 markers were added to those chromosomes respectively.

Meta-analysis and SNP selection. A meta-analysis was performed as described in Khatkar *et al.* (2004) to determine consensus locations for QTL on OAR1 and 3 for parasite traits. All papers that reported QTL on OAR 1 or 3 were included. The QTL location, confidence interval (CI) and a pre-assigned score according to the importance of the QTL were used to calculate a sum of scores at each putative QTL position along the chromosomes. These meta-scores were then used to select SNPs to be incorporated on a custom array designed to allow fine mapping of parasite traits in both ours and other populations, with the eventual aim of developing a diagnostic set of markers for use in breeding programs. The 1536 SNPs that could fit onto the array were distributed across the areas of interest according to the strength of the QTL score in that region, and were selected from the Virtual Sheep Genome resource (Dalrymple *et al.* 2007). Selection criteria were Illumina Score (to assess the likely success of the oligo design) >0.8, position, and MAF > 0.3. Genotyping is underway but was not completed at the time of publication.

RESULTS AND DISCUSSION

QTL Mapping. Following the addition of 12 markers to OAR 1 (27 markers in total) and 14 to OAR3 (n = 28), the information content increased markedly (Figure 1).

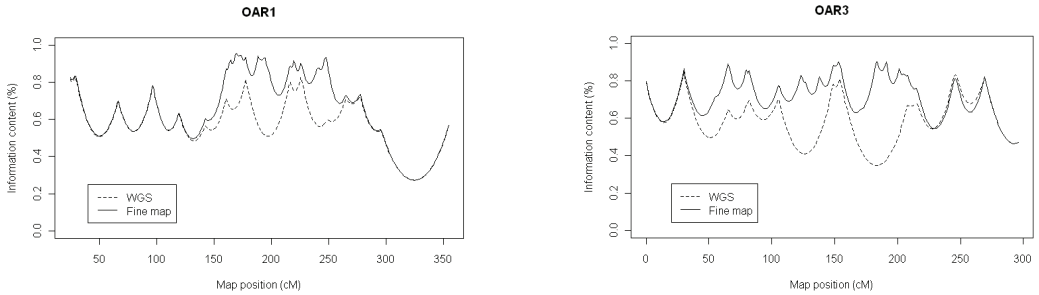


Figure 1: Information content along OAR1 and 3 before after fine mapping. The dotted line shows information content of the original scan, while the solid line shows the content after the addition of fine mapping markers.

Previously identified QTL for Wt change on OAR1 and FEC2 on OAR3 (data not shown) were not significant following the genotyping of extra markers. However, on each chromosome, one QTL for FEC during the first challenge remained significant (Table 1). The additional markers improved the F statistic from 2.95 (P<0.05 chromosome wide) to 4.37 (P<0.05 genome wide) for the QTL on OAR1, and from 1.64 (not significant) to 3.03 (P<0.05, genome wide) on OAR3. However, the confidence intervals remained large (Table 1), as illustrated by bootstrap analyses showing possible distribution of QTL locations in Figure 2.

Table 1: Significant QTL detected on OAR1 and 3. The framework map positions were taken from the sheep linkage map v4.7 (http://rubens.its.unimelb.edu.au/~jillm/jill.htm

Trait	QTL position (cM)	F	LR	Significance thresholds				95% CI (cM)
				Chromosome wide		Genome wide		
				0.05	0.01	0.05	0.01	
OAR 1 FEC1	188	4.37	42.55	2.945	3.616	3.770	4.660	68-352
OAR 3 FEC1	184	3.03	29.83	2.971	3.606	3.770	4.660	0-252

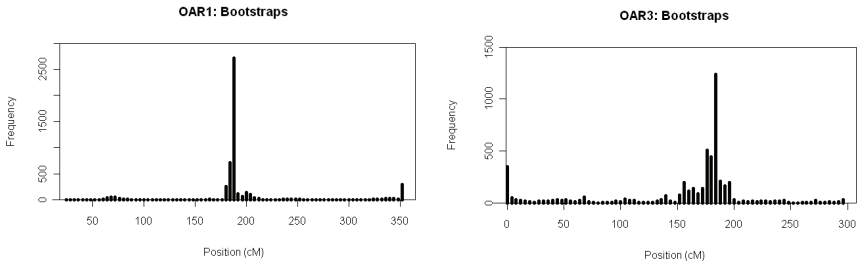


Figure 2: Distribution of possible QTL locations (5000 bootstraps)

Meta-analysis and SNP selection. The meta-analysis included 3 peer reviewed whole genome QTL studies (Crawford *et al.* 2006; Davies *et al.* 2006; Beraldi *et al.* 2007) and 8 unpublished reports, including milestone reports from collaborators and an update of the Beh (2002) project. The analysis indicated there may be a number of small QTL on each chromosome, instead of revealing strong evidence of a single QTL. The meta-scores were used to design the custom SNP array, with 1534 SNPs selected on OAR1 and 3 using the densities suggested in Table 1, while one SNP each from OARX and Y were included for quality control purposes.

Table 2: Suggested SNP positions and densities for custom parasite SNP array.

Area (cM)	Priority	SNP Density (/cM)	Flanking markers	Interval Length (cM)	SNPs
OAR1					
75–150	Med	3	OARCP93/UMJM36 - APPO10/BMS574	75	225
150–180	High	10	APPO10/BMS574 - MNS94/CSSM4	30	300
190–250	Med	3	MNS94/CSSM4 - MCM130/TEXAN6	60	180
300–325	Med	3	UWCA46/KD721 - MCM357/EPCDV13	25	75
Total				190	780
OAR3					
25–50	Med	3	DIK5391/ILSTS28 - BM746/BM1861	25	75
130–190	Med	3	BM2818/BMS1953 -UCD52/BL4	60	180
190–230	High	10	UCD52/BL4 - TEXAN15/PTHLH	40	400
230–260	Med	3	TEXAN15/PTHLH - ILSTS34/UCD14	30	90
Total				155	745
Overall				345	1525

Following the placement of additional markers on OAR1, the significant QTL mapped to a different position to previously published FEC QTL, found at around 270 - 340 cM and 79 cM respectively (Beh *et al.* 2002; Beraldi *et al.* 2007). However, the large CI spans the locations of all QTL identified in these studies so it is possible that the QTL detected by us and others in different positions may actually be caused by a common underlying genetic factor. A similar result was seen on OAR3, where previous studies have detected significant QTL for FEC and immune responses (Beh *et al.* 2002; Davies *et al.* 2006) as well as numerous suggestive QTL at positions between 0 - 35, 95 - 180 cM, and 200 - 240 cM (Crawford *et al.* 2006; Beraldi *et al.* 2007).

Due to the ambiguity of the QTL positions as indicated by large CI, which in both cases cover most of the chromosome, it was thought that performing the meta-analysis would be useful to determine the most suitable regions to add more SNP markers for fine mapping. However, on both chromosomes, the meta-analysis did not indicate the presence of a single QTL. It was thus decided to place SNPs over a larger area of the chromosome than just those regions surrounding the indicated QTL position, with increasing SNP density in the priority area as determined by the

combined weighted score from the meta-analysis. On OAR3, the priority area indicated by the meta-analysis was slightly downstream of the most likely position of our QTL (184 cM), and included the region surrounding the interferon-γ gene. Although a marker within interferon-γ (205.5 cM) has previously been implicated as contributing to variation in host resistance to parasites (Coltman *et al.* 2001, Paterson *et al.* 2001), other studies have failed to replicate these results, and we found no evidence to support this within our flock (Raadsma *et al.* 2008). The addition of SNP markers spread over a large area on both chromosomes will allow us to perform more rigorous LA/LD analyses. This method uses simultaneous multi-trait analyses and combines linkage analysis (LA) information with linkage disequilibrium (LD) information, thus preventing spurious associations, improving power and precision, and reducing confidence intervals, resulting in more specific QTL locations for the identification of causative polymorphisms (Meuwissen and Goddard, 2004).

CONCLUSION

We have mapped two QTL for FEC on OAR1 and 3, and performed a meta-analysis of other studies, published and otherwise, combining the information for targeted fine mapping using a custom designed 1536 SNP array. Fine mapping using LA/LD analyses methodologies to narrow down the region is critical before systematic positional candidate gene searches can begin. In the future we hope to identify and characterize genetic markers for parasite traits in sheep, allowing farmers the option to selectively breed their flocks for resistance and resilience to gastro-intestinal nematodes.

REFERENCES

Beh, K.J., Hulme, D.J., Callaghan, m.J., Leish, Z., Lenane, I., Windon, R.G. and Maddox, J.F. (2002) *An. Genet*. **33**:97

Beraldi, D., McRae, A.F., Gratten, J., Pilkington, J.G., Slate, J., Visscher, P.M. and Pemberton, J.M. (2007) *Int. J.Parasit.* **37**:121

Coltman, D.W., Wilson, K., Pilkington, J.G., Stear, M.J. and Pemberton, J.M. (2001) *Parasitology* **122**:571

Crawford, A.M., Paterson, K.A., Dodds, K.G., Diez-Tascon, C., Williamson, P.A., Roberts-Thomson, M., Bisset, S.A., Beattie, A.E., Greer, G.J., Green, R.S, Wheeler, R., Shaw, R.J., Knowler K. and McEwan, J.C. (2006) *BMC Genomics* **7**:178

Dalrymple, B.P., Kirkness, E.F., Nefedov, M., McWilliam, S., Ratnakumar, A., Barris, W., Zhao, S., Shetty, J., Maddox, J.F., O'Grady, M., Frank Nicholas, F., Crawford, A.M., Smith, T., de Jong, P.J., McEwan, J, Oddy, V.H. and Cockett, N.E. for the ISCG (2007) *Genome Biol.* **8:**R152

Davies, G., Stear, M.J., Benothman, M., Abuagob, O., Kerr, A., Mitchell S. and Bishop, S.C. (2006) Heredity **96**:252

Khatkar, M.S., Thomson, P.C., Tammen, I. and Raadsma, H.W. (2004) *Genet. Sel. Evol.* **36**:163

Meuwissen, T.H.E. and Goddard, M.E. (2004). *Genet. Sel. Evol.* 36, 261-279

Paterson, K.A, McEwan, J.C., Dodds, K.G. and Crawford, A.M. (2001). *Proc. Assoc. Adv. Anim. Breed. Genet.* **14**: 91

Raadsma, H.W., Fullard, K.J., Kingsford, N.M., Margawati, E.T., Estuningsih, E., Widjayanti, S., Subandriyo, Clairoux, N., Spithill, T.W. and Piedrafita, D. (2008) In "Genomics of Disease", p. 89, Editors J.P. Gustafson, J. Taylor and G. Stacey, Springer Science, New York.

Seaton, G., Haley, C.S., Knott, S.A., Kearsey, M. and Visscher, P.M. (2002) Bioinformatics **18**:339

Vanimisetti, H.B., Greiner, S.P., Zajac A.M. and. Notter D.R. (2004) *J. Anim. Sci.* **82**:595

ESTIMATING HERITABILITY OF SUBJECTIVELY ASSESSED OSTRICH LEATHER QUALITY TRAITS USING THRESHOLD MODELS

A. Engelbrecht[1,2], S.W.P. Cloete[1,3], K.L. Bunter[4] and J.B. van Wyk[5]

[1]Department of Animal Sciences, University of Stellenbosch, Matieland 7602, South Africa
[2]Institute for Animal Production, Oudtshoorn Research Farm, Oudtshoorn 6620, South Africa
[3]Institute for Animal Production, Elsenburg, Private Bag X1, Elsenburg 7607, South Africa
[4]Animal Genetics and Breeding Unit, University of New England, Armidale, NSW 2351, Australia
[5]Department of Animal, Wildlife and Grassland Sciences, University of the Free State, Bloemfontein 9300, South Africa

SUMMARY

Several ostrich leather traits cannot be assessed objectively within the present industry structure. A subjective scoring system was devised to allow the assessment of several qualitative leather traits for the estimation of genetic parameters. Nodule size and shape, along with the presence of hair follicles and pitting damage were scored on linear scales from 1 to 9 on 1021 processed ostrich skins. The data were analysed together with live weight and skin size using threshold-linear animal models, with compression of categories where required. Heritability estimates were 0.44 for live weight, 0.33 for skin size, 0.46 for nodule size, 0.33 for nodule shape, 0.49 for hair follicles and 0.15 for pitting. The 95% highest posterior density confidence intervals for all variance components excluded zero. The quantitative slaughter traits were favourably related to nodule shape and nodule size, while the presence of hair follicles or pitting was not significantly related to the other traits on a genetic level. Sufficient genetic variation was demonstrated to allow sustained genetic progress for objective and subjectively assessed traits.

INTRODUCTION

Ostrich leather is considered unique due to the feather follicles which gives it a distinctive nodulated appearance (Engelbrecht *et al*. 2009). The appearance of the nodules is consequently important when the value of ostrich leather is determined (Engelbrecht *et al.* 2005). Despite its importance, nodule traits are currently being assessed subjectively owing to a lack of practical objective evaluation methods (Cloete *et al.* 2006; Van Schalkwyk *et al*. 2005). The appearance and distribution of nodules, as well as the heritability of nodule size and density, varies between different locations on the skin (Cloete *et al.* 2006), making it difficult to accurately evaluate leather quality with objective methods.

Other aspects that influence the value of ostrich leather that are also being assessed subjectively are skin damage and skin defects. Skins are frequently downgraded due to pitting damage, defined as small superficial pits on the surface of the leather. This defect is clearly discernable on processed skins, and impairs the quality of the leather (Engelbrecht *et al.* 2009). An excessive occurrence of hair follicles among feather follicles is also considered a defect, resulting in skins being downgraded (Lunam and Weir 2006).

Slaughter age was previously shown to influence leather traits (Engelbrecht *et al.* 2005; Van Schalkwyk *et al.* 2005; Cloete *et al.* 2006). Van Schalkwyk (2008) later reported a significant interaction between age and weight for nodule traits, indicating that the effect of age and weight could not be partitioned accurately.

Cloete *et al.* (2008) recently demonstrated the genetic basis of most ostrich traits of economic importance, including slaughter and leather traits. Meyer *et al.* (2004) reported low heritabilities for nodule traits when measured with objective methods, while Engelbrecht *et al.* (2005) reported higher heritabilities for subjectively assessed nodule traits, analyzed with linear REML methods.

Improvement of these subjectively assessed leather traits through genetic selection was consequently further investigated. In order to better understand to what extent ostrich leather quality traits are heritable, linear-threshold models were used for data analysis in this study to properly account for the distribution and categorical nature of the data.

MATERIALS AND METHODS

Pedigree records and processed skins were obtained from South African Black ostriches in the research flock maintained at the Oudtshoorn Research Farm, near Oudtshoorn in the Klein Karoo region of South Africa. The pedigree file contained 2577 records, the progeny of 300 sires and 282 dams. Processed crust skins from birds slaughtered between 1997 and 2007, aged 240 to 480 days (mean ± s.d. = 367 ± 55 days), were scored for nodule traits (nodule size and shape), and for the prevalence of hair follicles and pitting damage. Live weight at slaughter, slaughter age and crust skin size were available for most of the birds included in the investigation.

Average nodule size (NSZ) and nodule shape (NS) on skins were evaluated on linear scales of 1 to 9, with one indicating poor quality and 9 indicating excellent quality. Scoring for nodule size was standardized by establishing visual standards for each score. A representative skin was chosen for each score from one to nine based on the average size of the nodules on the skin. These skins served as standards against which each skin could be compared during scoring. For nodule shape, each score was defined in terms of the general shape of the nodules on the skin. Scores of 1 to 3 were used for skins with mostly poorly developed or elongated nodules, 4 to 6 for skins with relatively rounded nodules and 7 to 9 for skins with well-rounded nodules.

The overall occurrence of hair follicles (HF) and pitting (PIT) on skins were also evaluated on a scale of 1 to 9, with one indicating none and 9 indicating an excessive occurrence of the relevant defect on a skin. The scoring system was devised by Engelbrecht *et al.* (2005), and was applied by the first author throughout to exclude variation between scorers.

Linear-threshold animal model analyses were applied for the estimation of (co)variance components and genetic parameters. Analyses were performed using THRGIBBS1F90 software, which is suitable for the estimation of (co)variance components and genetic parameters in threshold-linear animal mixed models, and allows for any combination of categorical and continuous traits (Misztal *et al.* 2002). A single chain of 200,000 samples were run, with 40,000 samples being used as the burn-in period. The remaining 160,000 samples were retained for post-Gibbs analysis with the POSTGIBBSF90 graphical program (Misztal *et al.* 2002). Every 10^{th} sample was saved and used to compute means (which were used as point estimates for parameters), posterior standard deviations and 95% highest posterior density (HPD) confidence intervals.

Fixed effects included in the analysis were contemporary group (n = 21) and gender (male and female). Slaughter age (240 to 480 days) was included as a linear covariate for all traits, while animal was added to the operational model as a single random effect. Preliminary analyses showed that thresholds could not be accurately discerned for all traits. The quality of threshold estimates depend on the number of records in categories, while too many categories (i.e. 9) tend to be similar to continuous traits and result in some thresholds having too little information (pers. comm. Ignacy Misztal). Due to the poor representation of certain categories (< 10 records), adjacent categories were merged to improve data distribution and the estimation of thresholds. Scores of 4 to 6 (average occurrence of pitting) and 7 to 9 (excessive occurrence of pitting) were pooled for pitting. After pooling adjacent categories based on the numbers of records per category, the number of categories entered in the final analyses amounted to five for pitting, six for nodule size and nodule shape (scores 1 and 2 pooled, and 7, 8 and 9 pooled), and seven for hair follicles (scores 7, 8 and 9 pooled).

Two sets of analyses were conducted: initially live weight (LW) and processed skin size (SSZ)

were analysed with each individual subjectively assessed skin trait in four three-trait analyses. A four-trait analysis involving all subjective traits was conducted subsequently.

RESULTS AND DISCUSSION

Means (± s.d.) for LW (n = 1194) and SSZ (n = 1451) were 92±17 kg and 137±14 dm² respectively. The distribution of records for scored traits before merging is shown in Table 1. The distribution of scores (n = 1021) for nodule traits and hair follicle prevalence did not deviate from normality, while scores for pitting were skewed and leptokurtic. Merging of pitting categories reduced skewness from 2.80 to 1.79 and kurtosis from 10.73 to 2.92. Pooling higher categories for pitting did not reduce the predominance of score 1, so pitting was absent for more than 50% of skins.

Table 1. Number of records per category for scored ostrich skin traits

Trait	Category								
	1	2	3	4	5	6	7	8	9
Nodule size	14	37	184	402	248	90	28	11	7
Nodule shape	14	75	233	333	232	101	30	3	0
Hair follicles	25	184	257	191	172	106	56	28	2
Pitting	648	271	49	31	9	8	3	1	1

Estimates of heritability (h^2), trait variances, along with environmental and genetic correlations are provided in Table 2. Covariance ratios did not differ significantly between the different three-trait analyses, or between the three-trait analyses and the four-trait analysis when the relevant standard errors were compared (not presented). Estimates reported for the scored traits are from the four-trait analyses, while the estimates for LW and SSZ provided in Table 2 were averaged across the four three-trait analyses. The genetic variance components and ratios for all traits were significant, as reflected by the 95% highest posterior density (HPD) confidence limits excluding zero. The 95% HPD intervals for the genetic variance components were 47.2 to 105 for LW, 26.3 to 58.3 for SSZ, 0.27 to 0.86 for NSZ, 0.16 to 0.90 for NS, 0.24 to 0.73 for HF and 0.04 to 0.36 for PIT.

Estimates of h^2 were moderate to high for all skin traits analysed, except for PIT, and were consistent across different sets of analyses. Estimates for NSZ, NS and HF were somewhat higher than previously reported by Engelbrecht *et al.* (2005) when data from a wider slaughter range were included in a linear model analysis (0.43 for NSZ, 0.20 for NS and 0.31 for HF). This could indicate that threshold analysis was more appropriate for estimation of parameters for subjectively scored skin traits. On the other hand, the use of more data from typical slaughter birds (narrower age range) could also have contributed to this result. In contrast, h^2 of LW were similar to the previous estimate of 0.46, while h^2 of SSZ was somewhat lower than the previous estimate of 0.47 (Engelbrecht *et al*. 2005).

The significant heritability estimate for pitting is somewhat surprising since pitting was previously thought to be related only to on-farm practices and environmental factors (Engelbrecht *et al.* 2009). It could, however, be an indication as to a parasitic origin of pitting damage, reflecting genetic differences in parasitic resistance.

The genetic correlation between live weight and skin size did not differ from unity, corresponding with previous studies conducted on the same resource population (Engelbrecht *et al.* 2005). Significant genetic correlations were also found between LW and NSZ, and between SSZ and NSZ or NS. The high correlation found between NSZ and NS confirmed previous results of Engelbrecht *et al.* (2005). It is noteworthy that genetic correlations of live weight and skin size with nodule traits were slightly lower in absolute terms when compared with the previous results

reported by Engelbrecht *et al.* (2005), who used a smaller data set with a wider range of slaughter ages and thus live weight and skin size. With more data at our disposal at present, it was possible to source sufficient records of animals with a more realistic slaughter age for this study. No other significant genetic correlations were found between the slaughter and skin traits investigated, although the standard errors were mostly so high that these estimates cannot be seen as conclusive.

Table 2. Additive genetic (σ^2_a) and environmental (σ^2_e) variance components, along with heritabilities, genetic and environmental correlations (± s.e.), for ostrich slaughter traits

	Trait					
	Live weight (LW)	Skin size (SSZ)	Nodule size score (NSZ)	Nodule shape score (NS)	Hair follicle score (HF)	Pitting score (PIT)
σ^2 (a)	75.92	42.29	0.56	0.53	0.48	0.20
σ^2 (e)	97.86	87.02	0.66	1.06	0.50	1.15
	Heritabilities (diagonal, bold), residual (above diagonal) and genetic (below diagonal) correlations					
LW	**0.44 ± 0.09**	0.64 ± 0.08	0.40 ± 0.10	0.32 ± 0.19	0.12 ± 0.10	-0.22 ± 0.08
SSZ	0.98 ± 0.19	**0.32 ± 0.07**	0.51 ± 0.10	0.49 ± 0.21	0.16 ± 0.08	-0.19 ± 0.07
NSZ	0.42 ± 0.16	0.49 ± 0.18	**0.46 ± 0.12**	0.38 ± 0.12	0.26 ± 0.11	-0.26 ± 0.10
NS	0.32 ± 0.19	0.49 ± 0.21	0.78 ± 0.27	**0.33 ± 0.12**	-0.07 ± 0.16	-0.12 ± 0.07
HF	-0.15 ± 0.16	-0.16 ± 0.08	0.04 ± 0.18	0.23 ± 0.20	**0.49 ± 0.13**	0.03 ± 0.08
PIT	0.24 ± 0.21	0.25 ± 0.21	0.55 ± 0.32	0.32 ± 0.27	-0.24 ± 0.25	**0.15 ± 0.06**

CONCLUSIONS

Improvement of ostrich leather quality is important for continued profitability in the commercial industry. Accurate evaluation of skin quality and directed genetic selection for improved leather quality are therefore a prerequisite.

This study demonstrated that genetic variation exists for nodule traits and hair follicle prevalence to allow sustained genetic progress for these traits, should it be desired as part of the overall selection objective. In practice, selection decisions will depend on progeny test results, as nodule traits cannot be scored on live birds at present. Alternatively progress in traits like nodule shape and nodules size are or could be achieved through favourable genetic correlations with live weight and/or skin size, evaluated on an age constant basis.

REFERENCES

Cloete, S.W.P., Engelbrecht, A., Olivier, J.J. and Bunter, K.L. (2008) *Aust. J. Exp. Agric.* **48**:1247.

Cloete, S.W.P., Van Schalkwyk, S.J., Engelbrecht, A. and Hoffman L.C. (2006) *S. Afr. J. Anim. Sci.* **36**:160.

Engelbrecht, A., Cloete, S.W.P. and Van Wyk, J.B. (2005) *Proc. 3rd Int. Ratite Sci. Symp. & XII World Ostrich Congr.,* 121.

Engelbrecht, A., Hoffman, L.C., Cloete, S.W.P. and Van Schalkwyk, S.J. (2009) *Anim. Prod. Sci*. **49**:549.

Lunam, C.A. and Weir, K.A. (2006) RIRDC Pub No. 06/054, Canberra, Australia.

Meyer, A., Cloete, S.W.P., van Wyk, J.B. and Van Schalkwyk, S.J. (2004) *S. Afr. J. Anim. Sci.* **34**:29.

Misztal, I., Tsuruta, S., Strabel, T., Auvray, B., Druet, T. and Lee, D.H. (2002) *Proc. 7th World Congr. Gen. Appl. Livest. Prod*. **33**:743.

Van Schalkwyk, S.J., Cloete, S.W.P., Hoffman, L.C. and Meyer, A. (2005) *S. Afr. J. Anim. Sci.* **35**:48.

Van Schalkwyk, S.J. (2008) PhD Thesis, University of Stellenbosch, South Africa.

IDENTIFICATION OF A POTENTIAL MARKER FOR ABSENCE OF DARK FIBRE IN VICUGNA PACOS (ALPACA)

Natasha L. Feeley and Kylie A. Munyard

Biomedical Science, Curtin University of Technology, Perth 6000 Western Australia

SUMMARY

The *Melanocortin-1 receptor* gene was sequenced in a group of 41 Australian alpacas and seven single nucleotide polymorphisms (SNPs) were identified within the coding region (D42D, N118N, L206L, E311E, T28A, G126S and R301C). Three of these SNP (T28A, G126S and R301C) showed an association with phenotypic colour variants when both skin and fibre colour were used to segregate animals into groups. We propose the identification of a haplotype (T28A/G126S), which appears to be a marker for the absence of dark pigment in alpaca fleeces. Animals with the G82/C126 combination did not have any dark pigment. Both A82G & C901T are potentially capable of altering *MC1R* function. It's therefore possible that we have identified wild type (dominant) and loss-of-function (recessive) alleles of the alpaca *MC1R* gene.

INTRODUCTION

Alpaca fibre is renowned for its strength and softness and is a highly valuable fibre in the textile industry. Colour is an important fibre characteristic because it influences the potential applications and value of the end product. Although using phenotype as a basis for breeding selection can be of assistance to breeders an understanding of the molecular characterization of coat colour would allow more effective selection.. The *Melanocortin-1 Receptor* is known to be a key regulator of pigment colour in mammals (Rouzaud and Hearing 2005; Hoekstra *et al.* 2006). *MC1R* encodes a receptor on the surface of the melanocyte (pigment producing cell) that mediates the proportions of dark (black-brown, eumelanin) or light (tan-red, pheomelanin) pigment in the granules transferred to the surrounding cells; so producing the visible coloration in fibre or skin epidermis (Hearing 2005; Hoekstra *et al* 2006; Tully 2007). The MC1R receptor can direct pigment production in response to external factors (e.g. melanocyte stimulating hormone and agouti signaling protein) or can be inherently defective so that it becomes constitutively active, dominant black, or deactivated, recessive yellow (Hearing 2005; Hoekstra *et al.* 2006; Tully 2007). Alpaca flocks often contain a variety of coat colour patterns and so present an opportunity to substantiate the molecular basis of coat colour. Such advancements would be highly advantageous to the alpaca industry as it could provide breeders with the knowledge to effectively select for preferred genotypes.

MATERIALS AND METHODS

Blood samples were collected from 41 alpacas. Initial sample analysis was performed on 9 entirely white and 14 entirely black animals. A second group of animals comprising a wider range of colour phenotypes were subsequently analysed (3 black/brown, 2 grey, 2 dark brown, 9 fawn, 1 rose/grey and 1 white animal).

Alpaca *MC1R* Primers MC1R-F (GGGAGAAGGTGAGTGTGAGG) and MC1R-R (GCTCTTCCTGGAGATTCGTG) were designed to hybridise to regions flanking the alpaca *MC1R* coding sequence. All polymerase chain reactions (PCR) were performed in an Eppendorf Mastercycler, in 10µl reactions. Amplified DNA was sequenced with Big Dye Terminator Technology (Applied Biosystems) and analysed on a 3730 DNA analyser (Applied Biosystems). Complete MC1R sequences for each animal were compiled into contigs using Vector NTI software (Invitrogen 2008), and compared with genes and proteins from other species by NCBI

BLASTn and BLASTx protocols (Invitrogen 2008; NCBI 2008). Initial sequence analysis aimed to determine the relationship between the SNP genotype and phenotype, which was defined as the presence of absence of black pigment in fibres and/or skin. This identified three SNP that appeared to have an association with phenotype. A Chi^2 test for association was performed on three of the SNPs, A82G, C126T and C901T. Additional animals of a range of intermediate phenotypic fibre/skin colour variants were then analysed and the genotype data segregated the animals in the same way.

RESULTS AND DISCUSSION

Seven SNPs were identified in the alpaca *MC1R* gene (Table 2). Four of the seven SNP caused no amino acid change (D42D, N118N, L206L and E311E) while the remaining three resulted in amino acid substitutions (T28A, G126S and R301C). No correlation was observed between fibre colour alone, and MC1R genotype in the 41 animals studied. However, when the animals were assigned to groups based on the presence or absence of eumelanin in fibre and skin, three SNP appeared to be associated with phenotype variation (Tables 2 & 3). A chi-squared test for association was performed to test the association between skin/fibre phenotype and SNP genotype. All three SNP were shown to have significant correlation at 2 degrees of freedom (Table 1).

Table 1. Results from chi-squared analysis of SNP genotype versus phenotype

SNP	Likelihood Ratio	df	Asymp. Sig (2-sided)
G82A	52.644	2	.000
C126T	52.644	2	.000
C901T	38.599	2	.000

Table 2. The phenotype and *MC1R* genotypes of the initial alpaca samples examined in this study. "*E*" denotes the proposed wild type allele and "*e*" denotes the proposed recessive alleles at *MC1R*. SNP in bold are those from which showed phenotypic correlations

SNP Genotype										
82 T28A	126 D42D	354 N118N	376 G126S	618 L206L	901 R301C	933 E311E	Fibre Colour	Eumelanin Present	Proposed MC1R alleles	n*
G/G	C/C	C/C	G/G	A/A	T/T	A/A	white	No	E^e/E^e	7
G/G	C/C	C/T	G/G	A/A	T/C	A/A	white	No	E^e/E^e	1
G/G	C/C	C/C	G/G	A/G	T/C	A/G	white	No	E^e/E^e	1
G/A	C/T	C/T	G/A	A/G	T/C	A/G	black	Yes	E^+/E^e	1
G/A	C/T	C/T	G/G	A/G	T/C	A/G	black	Yes	E^+/E^e	1
G/A	C/T	C/T	G/A	A/G	C/C	A/G	black	Yes	E^+/E^e	1
A/A	T/T	T/T	A/A	G/G	C/C	G/G	black	Yes	E^+/E^+,	4
A/A	T/T	T/T	G/A	G/G	C/C	G/G	black	Yes	E^+/E^+,	1
A/A	T/T	T/T	A/A	A/G	T/C	A/G	black	Yes	E^+/E^+,	2
A/A	T/T	T/T	A/A	A/G	C/C	G/G	black	Yes	E^+/E^+,	1
A/A	T/T	T/T	A/A	A/G	C/C	A/G	black	Yes	E^+/E^+,	3

* number of samples

Genotypes A82G and C126T were in complete concordance and hence are considered to be a haplotype. Analysis identified the G82/C126 combination as a possible marker for animals which had an absence of black pigment. These SNP were correlated with the presence or absence of eumelanin in skin and fibre. All animals with the G82/C126 combination were characterised by a lack of dark pigment, while animals that were either heterozygous or had the opposite combination, A82/T126, displayed dark pigment in skin and/or fibre. The animals expressing pheomelanin-only are proposed to have the genotype E^e/E^e representing the homozygous recessive genotype at MC1R, while the eumelanic animals are proposed to have the genotypes E^+/E^+, (homozygous wild-type) or E^+/E^e, which both allow normal eumelanin expression. It is not clear from the information we have gained so far whether these mutations are causative of a change in phenotype or merely linked to the absence of black pigment. It may be possible that these SNP are linked to a promoter mutation and don't necessarily cause the phenotypic change (Hornyak *et al.* 2001; Rouzaud and Hearing 2005; Smith *et al.* 2001). Nevertheless these polymorphisms may serve as a good predictor of pheomelanic animals for breeding purposes.

Table 3. The colour phenotype and *MC1R* genotypes of the additional alpaca samples examined in this study at the three significant polymorphisms. ? denotes that eumelanin status could not be determined

SNP Genotype						
82 T28A	126 D42D	901 R301C	Fibre colour	Eumelanin present	Proposed MC1R alleles	Number of samples
A/A	T/T	C/C	fawn	Yes	E^+/E^+,	3
A/A	T/T	C/C	silver/grey	Yes	E^+/E^+,	1
A/A	T/T	C/C	dark brown	Yes	E^+/E^+,	2
A/A	T/T	C/C	medium grey	Yes	E^+/E^+,	1
A/G	T/C	C/T	black/tan	Yes	E^+/E^e	3
A/G	T/C	C/T	fawn	Yes	E^+/E^e	2
A/G	T/C	C/T	white	Yes	E^+/E^e	1
G/G	C/C	T/T	rose/grey	No	E^e/E^e	1
G/G	C/C	T/T	fawn	No	E^e/E^e	3
G/G	C/C	C/T	fawn	?	E^e/E^e	1

C901T also appeared to be a significant candidate polymorphism for phenotype effect. Animals with the A82/T126/C901 combination were capable of producing eumelanin while animals with the G82/C126/T901 combination lack any eumelanic pigment in skin or hair fibres. The C901T polymorphism occurs in an extremely significant domain pertaining to structural integrity and function of the receptor (Strader *et al.* 1994; Tao 2006). Polymorphisms in this domain are reported to impair receptor function severely (Everts *et al.* 2000; Garcia-Borron *et al.* 2005; Sanchez-Mas *et al.* 2005). If these interactions are not properly carried out, downstream processes essential for the production of eumelanin are not initiated, resulting in the default colour, pheomelanin, being produced (Hoekstra *et al* 2006; Logan *et al.* 2003; Newton *et al.* 2000).

This study has provided new information on the possible effects of MC1R alleles in alpaca fibre pigmentation. The results have highlighted a significant haplotype that appears to be a marker for the absence of black pigment. This haplotype holds significant potential for use as a marker in breeding stock selection. While this study has provided significant new information

about MC1R, the nature of pigment gene interactions means that genetic analysis of a number of other pigment genes will be necessary before the nature of colour inheritance in this species is completely understood. Investigation and characterisation of the MC1R promoter may also yield useful information about the differences in MC1R expression in animals with identical genotypes that display varying degrees of pigmentation.

ACKNOWLEDGMENTS

This research was jointly funded by the Rural Industries Research & Development Corporation and Curtin University. Thanks to the Australian Alpaca Association for their support of this research. We thank the following alpaca breeders for supplying samples from their alpacas, Sue Leitch (Brookwood Alpacas), Carolyn Emery (Windella Alpacas), Pinjarra Alpacas, Aviana Farms, Bolero Alpacas, Mangowood Farm, and Tularosa Alpacas. Thanks also to Dr Katherine Morton (The University of Sydney) and Dr Belinda Appleton (The University of Melbourne/ Alpaca Genomics Australia Pty Ltd), who supplied DNA samples for nine of the animals.

REFERENCES

Everts, R.E, Rothuizen J. and AvanOost, B. (2000) Identification of a premature stop codon in the melanocyte-stimulating hormone receptor gene (MC1R) in Labrador and Golden retrievers with yellow coat colour. *Animal Genetics* **31**:194.

Garcia-Borron, J.C, Sanchez, B.L, Jimenez-Cervantes, C. (2005) Melanocortin-1 receptor structure and functional regulation. *Pigment Cell Research* **18**:393.

Hearing, V.J. (2005) Biogenesis of pigment granules: a sensitive way to regulate melanocyte function. *Journal of Dermatological Science* **37:** 3.

Hoekstra, H, Hirschmann, R.J, Bundey, R.A, Insel, P.A and Crossland, J.P. (2006) A single amino acid mutation contributes to adaptive beach mouse colour pattern. *Science* **313**:101.

Hornyak, T.J, Hayes, D.J, Chiu, L.Y and Ziff, E.B. (2001) Transcription factors in melanocyte development: distinct roles for Pax-3 and Mitf. *Mechanisms of Development* **101**:47.

Invitrogen (2008) Vector NTI Advanced.

Logan, D.W, Bryson-Richardson, R.J, Pagan, K.E, Taylor, M.S, Currie, P.D and Jackson, I.J. (2003) The structure and evolution of the melanocortin and MCH receptors in fish and mammals. *Genomics* **81**:184.

Newton, J.M, Wilkie, A.L, He, L, Jordan, S.A, Metallinos, D.L, Holmes, N.G, Jackson, I and Barsh, G.S. (2000) Melanocortin 1 receptor variation in the domestic dog. *Mammalian Genome* **11**: 24.

NCBI (2008) GenBank. In (National Center for Biotechnology Information)

Rouzaud, F and Hearing, V.J. (2005) Regulatory elements of the melanocortin 1 receptor. *Peptides* **26**:1858.

Sanchez-Mas, J, Sanchez-Laorden, B.L, Guillo, L.A, Jimenez-Cervantes, C and Garcia-Borron, J.C. (2005) The melanocortin-1 receptor carboxyl terminal pentapeptide is essential for MC1R function and expression on the cell surface. *Peptides* **26**:1848.

Smith, A.G, Box, N.F, Marks, L.H, Chen, W, Smit, D.J, Wyeth, J.R, Huttley, G.A, Easteal, S and Sturm, R.A. (2001) The human melanocortin-1 receptor locus: analysis of transcription unit, locus polymorphism and haplotype evolution. *Gene* **281**:81.

Strader, C.D, Fong, T.M, Tota, M.R, Underwood, D. (1994) Structure and Function of G Protein Coupled Receptors. *Annual Review of Biochemistry* **63**:101.

Tao, Y-X. (2006) Inactivating mutations of G protein-coupled receptors and diseases: Structure-function insights and therapeutic implications. *Pharmacology & Therapeutics* **111**:949.

Tully, G. (2007) Genotype versus phenotype: Human pigmentation. *Forensic Science International: Genetics* **1**:105.

ASSOCIATION OF FIBRE DIAMETER WITH WOOL COLOUR IN A SOUTH AUSTRALIAN SELECTION FLOCK

M. R. Fleet[1], K.R. Millington[2], D.H. Smith[1]and R.J. Grimson[1]

[1] South Australian Research and Development Institute,
Turretfield Research Centre, Rosedale SA 5350
[2] CSIRO Materials Science and Engineering, PO Box 21, Belmont Vic 3216

SUMMARY

Bulk colour and its photostability are recognised limitations of wool relative to competing fibres and form part of the research program of the CRC for Sheep Industry Innovation. This paper shows effects of the association between fibre diameter and wool colour traits (greasy yellowness score, clean brightness and clean yellowness) and how adjustment for this association can change dam line and sire differences. The 323 wool samples used in this study were from yearling progeny of ewes from the South Australian Selection Demonstration Flocks (SDF) and 6 industry representative Merino sires. Progeny of dams from the Fine Wool (FW) component flock is separated in the analysis from that of dams from the other flocks involved (SD). Clean wool yellowness (Y-Z) ranged between 8.1 and 11.3. The mean fibre diameter of the progeny of the FW ewes was 16.0μm while for the SD ewes it was 17.6μm. Sire and dam line (FW and SD), when fitted in the analysis together with rearing type and day of birth, were significant for greasy visual yellowness and the clean colour traits (P<0.01). When mean fibre diameter, standard deviation of diameter and coarse fibre deviation were added as covariates this removed the dam line effect (P>0.29) for the clean colour traits; while for greasy visual yellowness it remained significant (P<0.0001). Furthermore, the differences in clean colour between the progeny groups of the rams were also reduced. Therefore, it is important to take into account the association of fibre diameter when assessing differences in clean colour.

INTRODUCTION

Whiteness and photostability of colour after processing are limitations of wool relative to other apparel fibres (Millington *et al.* 2008). Despite its importance to the textile industry, routine measurement of clean colour for greasy wool sale lot tests in Australia has had poor adoption. Similarly, when colour is considered in sheep selection it is usually limited to assessments of the visual appearance of the greasy wool (Brown 2006). Clean wool yellowness is highly heritable in Merino sheep (James *et al.* 1990). However, strong positive genetic correlations exist between Y-Z and mean fibre diameter (Hebart and Brien 2009; Smith and Purvis 2009). Millington *et al.* (2008) suggest that surface area and optical differences between fine and coarse fibres influence clean colour. Also, fibre medullation can result in a whiter fibre appearance due to the diffraction caused by this hollow component (IWTO 1998).

This paper examines the association between fibre diameter and wool colour traits (greasy yellowness, clean brightness and Y-Z) in terms of dam line and sire differences involving progeny of ewes from the South Australian Selection Demonstration Flocks.

MATERIALS AND METHODS

Location and wool. The 323 yearling ewes born in 2007 and not previously shorn as lambs were located at Turretfield Research Centre. The yearlings were the progeny of 6 industry representative Merino sires used via artificial insemination across ewes of component flocks from the SARDI's South Australian Selection Demonstration Flocks (AWI 2007). The animals were run as a single

mob and progeny of ewes from the Fine Wool Flock component (FW) represented 41% of the total and each sire was represented by 10 – 23 individuals. Progeny of ewes from the other 5 component flocks from the Selection Demonstration Flocks together (SD) involved 28 – 36 animals per sire. The wool came from the rump region on the sheep and was collected from shorn fleeces.

Assessments and measurements. Each wool sample was subdivided to provide 40g for the AWTA Ltd tests of washing yield and clean colour and a 40g replicate sample for the OFDA100 test (IWTO 1998; 2000) and the remainder was used to make visual greasy wool assessments (AWI/MLA 2007). The AWTA Ltd tests provided a washing yield and tristimulus values X (red), Y (green) and Z (blue) for each sample (IWTO 2003). Clean yellowness is derived by Y-Z (higher value is more yellow) while brightness Y (higher indicates brighter or more intensity of reflection). A single slide with an average of 25,878 fibre snippets was tested by OFDA100 to measure the diameter and medullation variables (IWTO-57-98). The Visual Sheep Scores guide (AWI/MLA 2007) was used to score greasy yellowness (VC), fleece rot, dust penetration, crimp definition (CD), staple weathering, and staple structure. Crimp frequency was taken at the staple base and the average staple length based on 5 staples.

Statistical analysis. The data were analysed using procedures in SAS version 9.1. Regression analysis using the stepwise option with P=0.10 selected only mean fibre diameter (MFD) and VC as traits associated with Y-Z. As there was a limited range in Y-Z and some of the visual and medullation traits also had limited levels, a discriminant analysis was undertaken with 3 classes for Y-Z (<9.0; 9.0 – 9.9 and ≥10.0). GLM Model 1 involved Dam line (either FW or SD); Sire (1 to 6); Rearing type (1, single born or multiple born and single reared; 2, multiple born and reared) as the main effects and the first order interactions; and date of birth (day 13 – 26) was fitted as the covariate. GLM Model 2 involved simultaneously adding MFD, SDFD and CE (based on the discriminant analysis for Y-Z) as additional covariates to Model 1 (Table 1).

RESULTS AND DISCUSSION

The discriminant analysis selected the traits in Table 1 as associated with Y-Z class. MFD had the major effect while CVD, CF, CE and O had smaller significant effects. Figure 1 shows the relationship between MFD and Y-Z for all data. The correlation between the OFDA100 measures (Step 1-5) and Y-Z class was r = 0.71 (Table 1) while the regression shown in Figure 1 has a correlation of r = 0.67; which are higher than other reported phenotypic correlations (Smith *et al.* 2007) due to the added diameter components and perhaps also the large sample size (mean of 25,878 fibre snippets).

Table 1. Traits affecting fibre clean yellowness (Y-Z)

Step	Trait	Partial r^2	Probability >F
1	Mean fibre diameter (MFD)	0.385	<0.0001
2	Comfort factor (CF)	0.032	0.0052
3	Coarsest 5% of fibres deviation (CE)	0.033	0.0048
4	Coefficient of variation of diameter (CVD)	0.036	0.0031
5	Objectionable medullated fibres/10K (O)	0.023	0.0258
6	Greasy visual colour (VC)	0.019	0.0532
7	Crimp definition (CD)	0.017	0.0716

In GLM Model 1, Sire and Dam line were significant (P<0.01) for VC and clean colour X, Y, Z and Y-Z, and the diameter variables MFD, SDFD and CE. In Model 2 the covariates for

diameter were significant for all of the clean colour variables (MFD, $P<0.0001$; SD, $P<0.05$; CE, $P<0.01$), improving model r^2 and completely removing the effect of Dam line (Table 2), but these covariates were not important for VC. The effect of Sire in GLM Model 2 remained significant ($P<0.001$) for VC, X, Y and Z but became non-significant for Y-Z ($P=0.076$). Sire*Dam line interactions were significant ($P<0.05$) for Z and Y-Z in Model 1 and 2 (Table 3).

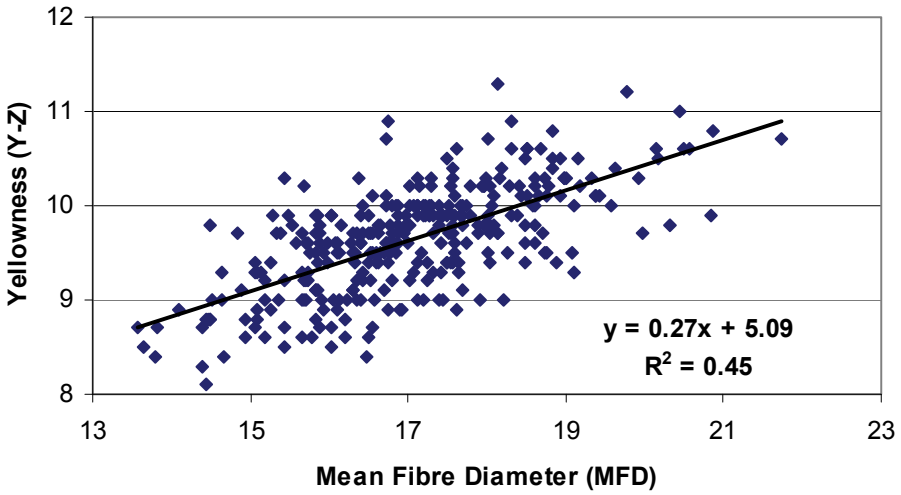


Figure 1: Relationship between mean fibre diameter (MFD) and yellowness (Y-Z)

A significant effect of Rearing Type (RT) occurred for X and Y ($P<0.05$) using Model 1 (lambs reared as multiples having lower values than singles) was removed in Model 2. RT also affected SDFD ($P=0.022$) and CE ($P=0.007$), the lambs reared as multiples had higher mean values than singles (3.94 v. 3.82µm and 7.42 v. 7.15µm; respectively), but not MFD. This result may reflect nutritional restriction on secondary follicle maturation for multiple born/reared lambs allowing more of the diameter variability associated with primary follicles to persist in the subsequent lamb/yearling fleece (Schinckel 1958). Sire*RT or Dam*RT interactions were not significant except Sire*RT for MFD in Model 1 ($P=0.048$); likely due to sampling. Day of birth as covariate, being restricted by artificial insemination to a range of 11 days, was non-significant.

Table 2 shows the GLM means for VC, Y and Y-Z and the associated difference in MFD, SDFD and CE of the progeny of SD and FW dam lines using Model 1 and in parenthesis the means from Model 2. Adjustment of the data for fibre diameter (MFD, SDFD and CE) reduced the difference between FW and SD dam lines for clean colour (Y and Y-Z) but not greasy colour (VC). Table 3 shows the sire mean values for VC and Y and the Sire*Dam line means for Y-Z from Model 1 and Model 1. Differences between the sire progeny groups for clean colour (Y and Y-Z) were reduced after adjustment for fibre diameter in Model 2.

Table 2. GLM means from Model 1 and Model 2 (in parenthesis)

Dam line	MFD	SDFD	CE	VC	Y	Y-Z
SD	17.6 a	4.10 a	7.75 a	2.07 a	72.2 a	9.81 a
(Model 2)				(2.09 a)	(72.5 a)	(9.65 a)
FW	16.0 b	3.65 b	6.82 b	1.64 b	72.8 b	9.37 b
(Model 2)				(1.61 b)	(72.2 a)	(9.64 a)

Means in columns with a different letter (a,b) are different at the 95% confidence level ($P<0.05$)

Table 3. Means for sire or sire within dam line for Model 1 and Model 2 (in parenthesis)

Sire		VC		Y		Dam	Y-Z	
	Model	1	2	1	2	line	1	2
1		2.08 bc		73.3 d		SD	9.73	(9.55)
			(2.06 bc)		(73.1 d)	FW	9.05	(9.45)
2		1.82 b		71.7 ab		SD	9.89	(9.62)
			(1.84 b)		(71.8 a)	FW	9.50	(9.64)
3		1.81 b		72.3 bc		SD	10.03	(9.82)
			(1.82 b)		(72.3 abc)	FW	9.35	(9.59)
4		2.10 c		71.4 a		SD	9.95	(9.65)
			(2.11 c)		(71.6 a)	FW	9.83	(9.96)
5		1.82 b		73.3 d		SD	9.62	(9.64)
			(1.80 b)		(72.8 bcd)	FW	9.10	(9.56)
6		1.49 a		72.8 cd		SD	9.63	(9.60)
			(1.48 a)		(72.6 cd)	FW	9.37	(9.65)
Range		0.61	0.62	1.9	1.5		0.99	0.51

Means in the VC and Y columns with a different letter (a,b,c,d) are different ($P<0.05$)

CONCLUSION

The basis of the relationship between fibre diameter and clean colour could involve several factors (e.g. surface area and optical properties; fibre structure and composition, fibre and staple density, follicle type ratio, wax and suint ratio and composition) that either affect the colour measurement or susceptibility of the fibre to yellowing. Whatever the cause it is clear that fibre diameter should be taken into account when considering clean colour differences.

ACKNOWLEDGEMENTS

Thanks to other staff at Turretfield and everyone else that had input or supported SARDI's Selection Demonstration Flocks project; AWI Ltd. and SARDI being the main funding providers. Helen Rimington assisted the wool measurements at Roseworthy Campus and AWTA Ltd. through Dr Trevor Mahar provided the clean colour data. In addition to the referee inputs, Dr Henry Wang (AWTA Ltd) and Dr Forbes Brien (SARDI) provided comments. This project was supported by the CRC for Sheep Industry Innovation and by Australian wool producers and the Australian Government through AWI Ltd.

REFERENCES

AWI (2007) Selection Demonstration Project. Merino Breeding for Profit and
AWI/MLA (2007) Visual Sheep Scores (http://www.wool.com.au/Publications/page__2272.aspx)
Brown D.J. (2006) *Int. J. Sheep Wool Sci.* **54**: 1.
Hebart M.L. and Brien F.D. (2009) *Assoc. Advmt. Anim. Breed. Genet*.**18**:500.
IWTO (1998) Specification IWTO-57-98. International Wool Textile Organisation.
IWTO (2000) Specification IWTO-47-00 International Wool Textile Organisation.
IWTO (2003) Specification IWTO-56-03 International Wool Textile Organisation.
James P.J., Ponzoni R.W., Walkley J.R.W. and Whitely K.J. (1990) *Aust. J. Agric. Res.* **41**: 583.
Millington K.R., Dyer J.M., Fleet M.R., Mahar T.J., Smith J. and Swan P. (2008) The 86th Textile Institute World Conference: p. 18 November, Hong Kong.
SAS Statistics package version 9.1 for Windows (SAS Institute Inc.; Cary, NC, USA).
Schinckel P.G. (1958) *Aust. J. Agric. Res*., **9**: 567.
Smith J. and Purvis I. (2009) *Assoc. Advmt. Anim. Breed. Genet*. **18**:390.

A NEW SYSTEM FOR COLLECTING AND PROCESSING PHENOTYPIC AND GENETIC INFORMATION FROM SHEEP FOR IMPROVED SELECTION TOOLS

K.G. Geenty[1], J.H.J. van der Werf[1], K.P. Gore[2], A.J. Ball[3] and S. Gill[3]

[1] CRC for Sheep Industry Innovation and the University of New England, Armidale, NSW 2351
[2]Animal Breeding and Genetics Unit, University of New England, Armidale, NSW 2351
[3]Meat and Livestock Australia, University of New England, Armidale, NSW 2351

SUMMARY

Current genetic improvement schemes are well developed in the Australian sheep industry for standard quantitative production traits. With the advent of new molecular genetic technologies, however, a new sheep industry linked system is being developed by the Sheep CRC based on the Information Nucleus (IN) concept. This sytem will deliver new and improved Australian Sheep breeding Values (ASBVs) for difficult and expensive to measure traits for product quality and production efficiency. The IN system is a world first and uses a network of research flocks representing the wider industry breed structure and combines extensive phenotyping and genotyping of up to 5,000 progeny annually. The IN system has developed new phenotyping protocols with modern electronic data capture, a sophisticated database and is linked with contributing stud breeders via Sheep Genetics (SG). Effective delivery of next generation genetic tools to the industry is facilitated by the stud breeder participation including provision of semen from selected sires and involvement in field activities.

INTRODUCTION

The Australian sheep industry has well developed and effective genetic improvement schemes in MERINOSELECT and LAMBPLAN based on delivery of Australian Sheep Breeding Values (ASBVs) by Sheep Genetics (SG) (Brown *et al.* 2007). The current SG system caters for approximately 52 traits routinely measured by commercial stud breeders. However other traits relevant to the breeding objective such as wool and meat quality that are hard or expensive to measure are not included. Such traits can still be improved if they are correlated with measured traits, but progress is likely to be slow unless the traits are highly correlated.

The Information Nucleus (IN) concept was first proposed by Banks *et al.* (2006) and the world's first IN materialized with a new CRC for Sheep Industry Innovation (Sheep CRC) in 2007. The concept involves intensive measurement of many new traits on a large group of animals linked to the sheep industry through selected sires related to other sire families. Sufficient offspring for each sire in different environments allows estimation of heritabilities and genetic correlations for new traits, and assessment of genotype by environment interactions. The system therefore provides genetic parameters and breeding values for important industry sires including important new traits not currently measured in commercial studs.

The recent development of new molecular genetic technologies such as single nucleotide polymorphism (SNP) chips for genome wide genotyping of dense markers gives the IN potential to add value to the current ASBV system. It allows extension of information on ASBVs for new traits to unrelated animals. Such *genomic selection* is already being implemented in young dairy bulls for milk production traits. The combination of intensive phenotyping of animals, potentially accelerated with genomic selection, will provide commercial stud breeders with next generation genetic tools to enhance genetic improvement, particularly for new traits related to efficiency and product quality.

The IN forms a major platform for research programs in the Sheep CRC, outlined previously by Banks *et al.* (2006) and Fogarty *et al.* (2007). This paper describes in more detail how the

Sheep CRC IN has been developed during the initial two years. The system for collection and processing of phenotypic and genetic information for development of next generation genetic tools is outlined.

SYSTEM DEVELOPMENT

The Sheep CRC system is based on data from five annual matings of approximately 100 new sires each year to 5,000 ewes across eight Australian research stations. Sheep breeds and environments represent a cross section of the Australian sheep industry (Fogarty *et al.* 2007). Seventy percent of progeny are slaughtered in the first year and the remaining 30% are females retained for reproductive evaluation.

Components of the system include-

- Development of data capture and management protocols including quality assurance
- Phenotypic data capture and collection of biological samples from some 5,000 progeny each year
- Genotyping and construction of SNP chips on approximately 2,500 progeny each year
- Construction of a database to accommodate the phenotypic and genetic data
- Association analysis with Sheep Genomics Program (SGP) for new new molecular ASBVs
- Establishment of linkages with other industry databases including SG and the SGP
- Delivery of new genetic tools to the sheep industry

IN flocks and data flow. The eight Information Nucleus flocks, representing the University of New England (Armidale), NSW DPI (Trangie and Cowra), Vic DPI (Rutherglen and Hamilton), SARDI (Struan and Turretfield) and DAFWA (Katanning), have 1,000 base ewes per organization (800 Merino and 200 Maternal except for DAFWA with 1,000 Merinos) joined annually to approximately 40 Merino, 40 terminal, and 20 maternal sires (Fogarty 2007). Quality assurance for phenotypic data is through protocols at IN sites, customised validation routines on entry to the IN database, during preliminary statistical analysis by IN data managers for each trait group and by scrutiny of preliminary IN ASBVs from SG. The IN ASBVs are then combined with industry information for standard traits by SG before delivery to the sheep industry (Figure 1). In addition data for new traits will be developed via prototype ASBVs before validation and delivery to industry by SG.

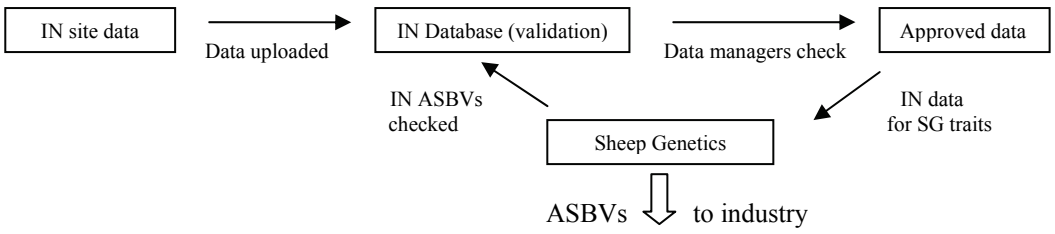


Figure 1. The flow of information from IN sites to SG and then to the sheep industry

Protocols

Detailed data collection protocols, including animal management guidelines and quality assurance, are in the IN Operational Manual (www.sheepcrc.org.au/insite). Also included in the manual is an outline of the IN program, a summary checklist of activities, trait dictionaries,

procedures for collection of biological samples and protocols for nutritional and animal health monitoring and management. Procedures for data transfer with the IN database are included. The Operational Manual is web based with regular updates.

Phenotyping. Phenotypic data for the 4,500-5,000 progeny each year for 163 traits are summarized in Table 1. In the database there are 266 traits including components of those included (eg. multiple values for wool micron profile data, meat colour and fatty acid data). New trait groups are in italics below.

Table 1. Trait groups representing phenotypic data captured from progeny

Program (no. traits)	Main trait groups
Sheep (81)	Lambing including *autopsies*, *visual scores*, live weight/condition score, worms
Meat (44)	Carcase measures, muscling/*yield*/fat, *meat colour/minerals/tenderness/flavo*
Wool (38)	Fleece weight, wool length/strength/micron profile, *colour, UV stability, comfort*

Newly developed e-sheep techniques are used for data capture including RFID tags and readers, auto weighing and drafting and blue tooth technology for wool information and sheep scores.

Genotyping. The first ovine SNP chip was made available in August 2008 by Illumina Inc. This SNP50 BeadChip was developed by an international consortium, with major inputs from Australia and New Zealand allowing genotyping with 55,000 SNPs. The CRC has genotyped 3650 animals using this SNP and along with 4200 genotypes from Sheep Genomics (Oddy *et al.* 2005) an association study is currently being carried out to derive a SNP set and prediction equation for genetic merit of individual traits. A set of about 700 industry sires with accurate ASBVs for commercial traits will be used for validation. Animals genotyped were mainly from the major breeds including Merino, White Suffolk, Poll Dorset and Border Leicester, as well as Texel and Suffolk, to assess genomic diversity and across breed application of genomic selection. The whole genome analysis will provide an insight into the ability to predict ASBVs from genomic data. It may also give further information about the segregation of quantitative trait loci (QTL) in certain regions, potentially providing further information about the genetic regulation of traits.

Database. The INF database (infdata.une.edu.au) is hosted by the Animal Breeding and Genetics Unit (AGBU) at the University of New England and is a relational SQL compliant platform with an ASP.NET front end allowing simultaneous multi-users. Data uploads, extracts and reports are available to acredited IN sites and researchers while technical aspects are controlled by the database manager. Data uploads are verified at submission according to data ranges, allowed values, specified trait names and correct pedigree information. An automated response is sent by email for error reports and/or final confirmation of upload. Real time web based reporting includes data collection statistics, summary statistics and phenotypic animal performance. Data extracts are generic for users or specialized for SG. The many and varied phenotypic traits and large volume of genotyping information are accommodated and configured at the front end.

Industry links and delivery. The major link with the sheep industry is through SG both in sourcing sires for the IN flocks from their database and developing and delivering ASBVs. Involvement by breeders through semen provision and attendance at field events has created significant 'buy in' to the IN program and this helps facilitate effective delivery and uptake of new and improved ASBVs, especially for new traits that are difficult or expensive to measure by stud breeders. In addition the combining of IN data with that from breeders for standard traits provides more data and sire linkage that strengthens the accuracy of ASBVs. The industry 'footprint' of the IN provided by common use of sires and their close relatives is significant. Current estimates are that over 35% of Merino, 30% of Border Leicester and 70 % of terminal breed sheep will have direct links to sires used in the IN. The link with SGP will facilitate verification of individual gene markers and development of enhanced ASBVs with use of molecular technologies (Hynd 2006).

CONCLUSIONS

The IN will impact on genetic gain in the Australian sheep industry by –

- developing an efficient system for collection and processing of phenotypic data and combining this with molecular genetic information for new and enhanced ASBVs
- modeling the new system on the industry breed structure with a significant genetic footprint and effective linkage with Sheep Genetics
- ensuring effective delivery of new genetic tools through stud breeder and wider sheep industry engagement

ACKNOWLEDGMENTS

The IN is supported by the Australian Government's Cooperative Research Centres Program, MLA and AWI. Considerable resources are provided at each site by NSW DPI, UNE, Vic DPI, SARDI and DAFWA. Valuable participation by SG and SGP is integral to the system. Development and maintenance of the IN database is supported by AGBU.

REFERENCES

Banks, R.G., van der Werf, J.H.J., and Gibson, J.P. (2006) *Proc. 8th Wld Congr. Genet. Appld. Livest . Prod.* Belo Horizonte, Brazil, August, CD-ROM Communication 30-12.

Brown, D.J., Huisman, A.E., Swan, A.A., Graser, H-U., Woolaston, R.R., Ball, A.J., Atkins, K.D., Banks, R.G. (2007) *Proc. Assoc.Advmt. Anim. Breed. Genet*. **17:**187.

Fogarty, N.M., Banks, R.G., van der Werf, J.H.J., Ball, A.J. and Gibson, J.P. (2007) *Proc. Assoc. Advmt. Anim. Breed. Genet.* **17:**29.

Hynd, P.I. (2006) *Proc. 8th Wld Congr. Genet. Appld. Livest . Prod.* Belo Horizonte, Brazil, August. CD-ROM Communication 05-06.

Oddy, V.H., Longhurst, T.J., Nicholas, F.W., Maddox, J.F. and McDonagh, M.B. (2005) *Proc. Assoc.Advmt. Anim. Breed. Genet*. **16:**209.

EFFECT OF THE ACCURACY OF AN ESTIMATED QTL EFFECT ON RESPONSE TO MARKER-ASSISTED SELECTION

N. Moghaddar[1,2] and J.H.J van der Werf[2]

[1]Ferdowsi University of Mashhad, Iran. [2]University of New England, Armidale, NSW, Australia

SUMMARY

The effect of the accuracy of estimated QTL effects on relative efficiency of marker-assisted selection (MAS) to non-MAS was investigated in a simulation study in a single trait selection scheme under variable heritability, different QTL effect and different initial QTL allele frequencies. The results showed that the probability of realized MAS response being less than expected response increased with higher standard error associated with an estimated QTL effect. This probability was significantly higher when a high standard error was associated with large QTL effects in lower QTL allele frequencies and/or lower heritability. The MAS responses could be equal or even less than non-MAS when the QTL variance from an inaccurate QTL effect constitutes a considerable proportion of overall genetic variance. This study showed that incorporating prior information about the QTL in a Bayesian approach can effectively remedy the problem of over prediction of MAS response.

INTRODUCTION

Information about quantitative trait loci (QTL) in economically important traits in livestock species, can be utilized to improve the rate of genetic gain via marker assisted selection. To achieve a high accuracy in estimation of QTL effects, the QTL mapping experiment needs to have a large number of progeny. However, because of the cost of experiments, QTL detection studies generally have a limited size, which restricts the accuracy of parameter estimation (Alison *et al.* 2002). Moreover, QTL detection experiments have considered a continuous distribution, normal distribution or no distribution for QTL effects (Hayes and Goddard 2001). The limited sample size along with ignoring the distribution of QTL effects result in upward biased estimation of QTL effects in QTL mapping experiments. The objectives of this study were to determine how the accuracy of estimated QTL parameters influence the realized response in MAS and how we can incorporate prior information about the QTL in prediction of realized MAS response.

MATERIALS AND METHODS

A stochastic simulation method was used to produce a base population of 500 unrelated, non inbred and unselected individuals with phenotypic variance equal to unity. The phenotypic values (y_p) consisted of polygenic effect (*poly*), a QTL genotype effect (g) and residual effect (e): $y_p=poly+(g-\mu)+e$. One QTL was assumed with two alleles with the difference between the two homozygotes being 2a, and assuming no dominance effect at the QTL. The QTL variance was therefore $V_g=2pq(a)^2$, where, p and q represent initial QTL allele frequencies and α is the QTL allele substitution effect (Falconer and Mackay, 1996). The heritability of the QTL and the proportion (in percent) of the QTL in overall additive genetic variance were calculated respectively as: $h^2_{(g)}= V_{(g)} /(V_{(poly)}+ V_{(g)} +V_{(E)})$, $\%V_{(g)}= V_{(g)} /(V_{(poly)}+ V_{(g)})*100$. We assumed to genotype directly for the QTL but we will use the term marker assisted selection (MAS) throughout this paper.

We considered variable heritability (0.1, 0.3 and 0.5), different standard errors, (SE: 0.0, 0.05, 0.1 and 0.15) associated with the estimated QTL effects and different initial allele frequency. Given an estimate of QTL effect and its standard error, we sampled a true QTL effect according to an appropriate posterior distribution and assigned a genotype and genotypic effect to each animal

according to the allele frequency. A polygenic and a residual effect was allocated randomly to each animal from a normally distributed polygenic effect with mean and variance equal to 0.0 and σ^2_A and from a normal distribution of residual effects, with mean and variance equal to 0.0 and residual variance σ^2_E respectively.

For each true QTL effect a MAS and non-MAS selection index were calculated using the selection index theory (Lande and Thompson 1990) and the top twenty percent of animals were selected. To find realized MAS response a gamma distribution was considered as a prior for the true underlying QTL effects (Hayes and Goddard 2001) and based on the Bayes theorem a posterior QTL distribution of the true QTL effect was derived given the prior distribution and the likelihood associated with the estimated QTL effect with its SE. The posterior probability of true QTL effect according to the Bayes theorem can be shown by this equation: $P(x_i|y_i) = P(y_i|x_i).P(x_i)/P(y_i)$ where x_i and y_i are the i^{th} true and estimated QTL effect respectively, $P(x_i|y_i)$ is the posterior probability of true QTL effect given the estimated QTL effect, $P(y_i \mid x_i)$ is the probability density function of the estimated QTL given the true value, which is the same as the likelihood function of the QTL effect, $P(x_i)$ and $P(y_i)$ are the prior probability of the true and estimated QTL effect respectively. The relative response of MAS over non-MAS for each estimated QTL value and associated SE was calculated based on deterministically sampling true QTL effects with 0.02 increments from the posterior distribution of QTL effects. For each value of possible QTL effect we assigned a polygenic value, residual value and genotypic value to animals stochastically according to polygenic variance, residual variance and genotypic value and response was calculated based on average of 500 iterations.

RESULTS

Tables 1 and 2 show, and Figure 1 plots the ratio of expected and realized MAS to non-MAS response for different combinations of QTL size, QTL allele frequency, standard error of the estimated effect for heritability values 0.1 and 0.3. The results show that in all cases the realized response of an estimated QTL associated with some degree of error was less than expected response. The difference was highly dependent on the accuracy of estimated effect. At each SE the difference between expected and realized response was bigger when the QTL variance was higher (bigger QTL effect, higher favourable allele frequency and lower heritable trait). As expected the probability of MAS response to be more than non-MAS response (P1) and probability of true QTL being smaller than estimated effect (P2) increased with higher QTL variance.

Table 1. Ratio of realized MAS over non-MAS response, probability of MAS response to be more than non-MAS response (P1) and probability of true QTL to be less than estimated effect (P2) for estimated QTL effect at different accuracy, (*heritability = 0.1*)

α^1	$p^{(2)}$	$V_{QTL}{}^{(3)}$	SE =0.0	SE=0.05			SE=0.1			SE=0.15		
			Resp$^{(4)}$	Resp	P1	P2	Resp	P1	P2	Resp	P1	P2
0.1	0.1	1.8	1.091	1.065	0.804	0.699	1.059	0.605	0.656	1.051	0.511	0.625
0.3	0.1	16.2	1.762	1.663	0.994	0.748	1.358	0.751	0.841	1.125	0.523	0.891
0.5	0.1	45.0	2.683	2.572	0.993	0.753	2.241	0.991	0.854	1.725	0.826	0.930
0.7	0.1	88.2	3.350	3.232	0.993	0.657	2.967	0.998	0.869	2.480	0.984	0.951
0.1	0.3	4.20	1.174	1.131	0.804	0.699	1.128	0.725	0.655	1.117	0.592	0.625
0.3	0.3	37.8	2.094	1.970	0.992	0.748	1.626	0.870	0.843	1.370	0.670	0.896
0.1	0.5	5.0	1.188	1.139	0.804	0.699	1.119	0.725	0.655	1.082	0.72	0.625
0.3	0.5	45.0	2.270	2.122	0.992	0.748	1.682	0.869	0.847	1.321	0.667	0.902

(1) = Estimated QTL Effect, (2) = QTL favourable allele frequency, (3) = Variance of QTL (as percent of heritability), (4) = Ratio of realized MAS to non-MAS response, SE = Standard Error of estimated QTL effect.

Table 2. Ratio of realized MAS over non-MAS response, probability of MAS response to be more than non-MAS response (P1) and probability of true QTL to be less than estimated effect (P2) for estimated QTL effect at different accuracy, (*heritability = 0.3*)

α^{1}	$p^{(2)}$	$V_{QTL}{}^{(3)}$	SE=0.0	SE=0.05			SE=0.1			SE=0.15		
			Resp$^{(4)}$	Resp	P1	P2	Resp	P1	P2	Resp	P1	P2
0.1	0.1	0.60	1.007	1.003	0.643	0.699	1.007	0.585	0.655	1.010	0.591	0.624
0.3	0.1	5.40	1.058	1.044	0.912	0.748	1.006	0.597	0.843	0.978	0.385	0.891
0.5	0.1	15.0	1.170	1.151	0.998	0.753	1.078	0.776	0.854	0.969	0.411	0.928
0.7	0.1	29.4	1.356	1.318	0.999	0.809	1.229	0.959	0.859	1.062	0.655	0.934
0.1	0.3	1.40	1.013	1.010	0.643	0.699	1.007	0.585	0.655	1.009	0.591	0.625
0.3	0.3	12.6	1.119	1.092	0.960	0.748	1.018	0.598	0.841	0.961	0.385	0.891
0.5	0.3	35.0	1.319	1.282	0.999	0.753	1.171	0.877	0.854	0.998	0.522	0.928
0.1	0.5	1.66	1.016	1.008	0.641	0.699	1.008	0.586	0.655	1.010	0.591	0.624
0.3	0.5	15.0	1.125	1.094	0.960	0.604	1.019	0.597	0.841	0.960	0.385	0.891
0.5	0.5	41.6	1.357	1.309	0.998	0.753	1.168	0.832	0.854	0.958	0.410	0.958

(1) = Estimated QTL Effect, (2) = QTL favourable allele frequency, (3) = Variance of QTL (as percent of heritability), (4) = Ratio of realized MAS to non-MAS response, SE = Standard Error of estimated QTL effect.

Figure 2 shows the effect of heritability on realized response of inaccurate QTL effect. In some cases when a considerable QTL variance was due to a QTL with high SE the realized response was even less than non-MAS response.

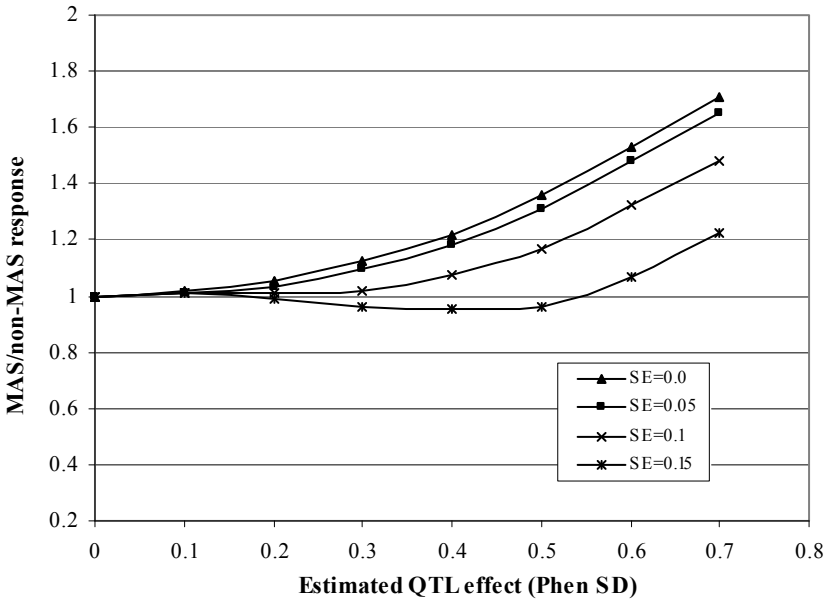


Figure 1. Ratio of MAS to non-MAS response for different SE associated with QTL effect, (h2=0.3, QTL allele frequency=0.5)

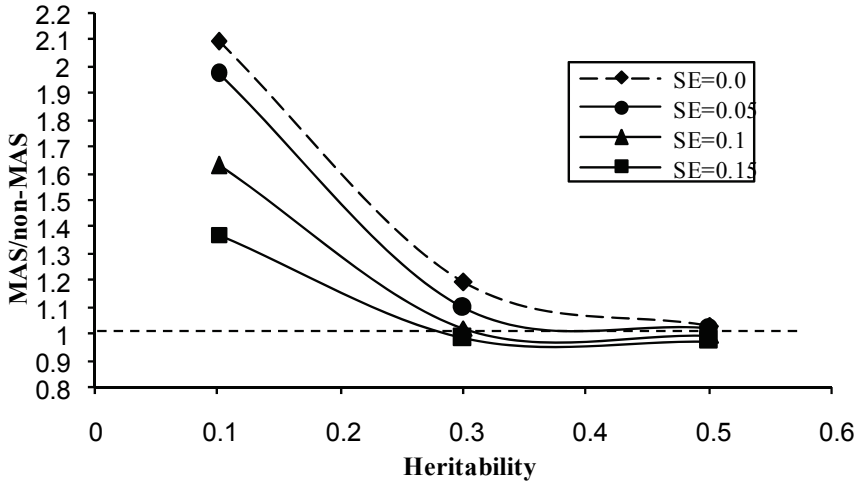


Figure 2. Effect of heritability on the relative efficiency of realized MAS GAS to non-MAS response (QTL effect=0.3 σP, Frequency=0.3)

DISCUSSION

This study showed that the realized MAS response from incorporation of information from an inaccurately estimated QTL is lower than the expected response and the deviation is bigger when the QTL constitutes a bigger part of the additive genetic variance. The simulation revealed that the difference between realized and expected MAS response is due to both incorrect prediction of QTL response and suboptimal prediction of response from polygenic effect.

This study also showed that using prior information about QTL effect distribution can give more accurate predictions of realized MAS response. In this study an average of derived parameters of QTL distribution in literature was used. However, using other parameters within a confidence interval had a small effect on our results except when QTL effects were large and had a high SE.

REFERENCES

Alison, D.B., Fernandez, J.R. Moonseong, H. Zhu, A, Etzel, C. Beasley T. M. and Amos, C.I. (2002) *Am. J. Hum. Genet* **70**:575.

Falconer, D. S. and Mackay, T.F.C (1996) "Introduction to Quantitative Genetics. 4th ed." Prentice Hall, Harlow, U.K.

Hayes, B. J. and Goddard, M.E (2001) "The distribution of effects of gene affecting quantitative traits in livestock." *Genet. Sel. Evol.* **33**:209.

Lande, R. and Thompson, T (1990) "Efficiency of Marker-Assisted Selection in the improvement of Quantitative Traits." *Genetics* **124**:743.

A COMPARISON OF METHODS FOR GENOMIC SELECTION IN AUSTRIAN DUAL PURPOSE SIMMENTAL CATTLE

B. Gredler[1], K. G. Nirea[1,2], T. R. Solberg[3], C. Egger-Danner[4], T. Meuwissen[2] and J. Sölkner[1]

[1] University of Natural Resources and Applied Life Sciences Vienna, Department of Sustainable Agricultural Systems, Division of Livestock Sciences, Gregor Mendel Str. 33, A-1180 Vienna
[2] Norwegian University of Life Science, Department of Animal and Aquacultural Sciences, Box 5003, N-1432 Ås
[3] Geno Breeding and AI Association, Box 5003, N-1432 Ås
[4] ZuchtData EDV-Dienstleistungen GmbH, Dresdner Str. 89/19, A-1200 Vienna

SUMMARY

The objective of this study was to compare partial least squares regression (PLSR), multivariate regression analysis using least absolute shrinkage and selection operator (LASSO), a Bayesian approach (BayesC) and an ordinary BLUP method (GS-BLUP) for the estimation of genome-wide breeding values for dual purpose Simmental Fleckvieh in Austria. A five-fold cross validation and a forward prediction were carried out for the traits protein yield, fat percentage, somatic cell count, and non return rate after 56 days in cows. Using cross validation, accuracies of genome-wide breeding values were in the range of 0.30 to 0.74. In the forward prediction, obtained accuracies were between 0.21 and 0.60. BayesC gave slightly better accuracies in forward prediction than the other methods.

INTRODUCTION

The use of molecular markers for improvement of genetic evaluation has been a major issue in animal breeding for many years. High throughput genotyping technologies enable the genotyping of more than 50,000 single nucleotide polymorphisms (SNP). Genomic selection, first introduced by Meuwissen *et al.* (2001), refers to the use of dense markers covering the whole genome to estimate genome-wide breeding values. In this simulation study the authors reported that it was possible to reach accuracies of genome-wide breeding values of 0.85 using markers only. So far, very few studies have reported results of genomic selection using real data. The objective of this study was to carry out a first comparison of methods for the estimation of genome-wide breeding values in dual purpose Simmental cattle in Austria.

MATERIAL AND METHODS

Data. 1,363 dual purpose Simmental (Fleckvieh) bulls, genotyped with the Illumina Bovine SNP50^{TM} Beadchip with a call rate >=95%, were included in the analysis. Bulls were born from 1990 to 2003. The distribution of bulls across birth years is shown in Figure 1. For method validation a five-fold cross validation was carried out. Bulls for the training and test set were randomly sampled across all birth years that 1,091 and 272 bulls were in the training and test set, respectively. Five replicates were carried out resulting in five pairs of training and test sets. In addition, the data set was split into a reference population (training set) including bulls born before 2001 (1,037 bulls) and a test set of bulls born between 2001 and 2003 (326 bulls) for forward prediction.

To be included in the analyses, a minimum minor allele frequency of 1 % was required for each SNP. To test for Hardy-Weinberg equilibrium, the deviation of observed genotype frequencies from expected genotype frequencies based on allele frequencies was calculated. SNP were included if Hardy Weinberg χ^2 values were below 600. In total, 45,519 SNP met all the SNP

selection criteria. As none of the methods described below allows for missing values, missing genotypes were filled in according to allele frequencies by sampling random numbers from a uniform distribution. The phenotypes used were estimated breeding values based on progeny testing obtained from the joint routine genetic evaluation in Austria and Germany for protein yield (Prot-kg), fat percentage (Fat%), somatic cell count (SCC) and non return rate after 56 days (NR56) for cows. Progeny testing involves samples of 50-100 daughters in Austria and Germany.

Model of analysis. In this study, we compared partial least square regression (PLSR), regression analysis using least absolute shrinkage and selection operator (LASSO), a Bayesian approach (BayesC; Meuwissen, 2009) and a BLUP approach as described by Meuwissen *et al.* (2001). For running PLSR and LASSO, the SAS procedures PROC PLS and PROC GLMSELECT were used (SAS, 2007). To assess the accuracy of genomic selection, the correlation between estimated genome-wide breeding values (GEBV) and current estimated breeding values (EBV) based on progeny testing was calculated. The regression coefficient of the current breeding value on the genome-wide breeding value was computed to assess the bias of genome-wide breeding values.

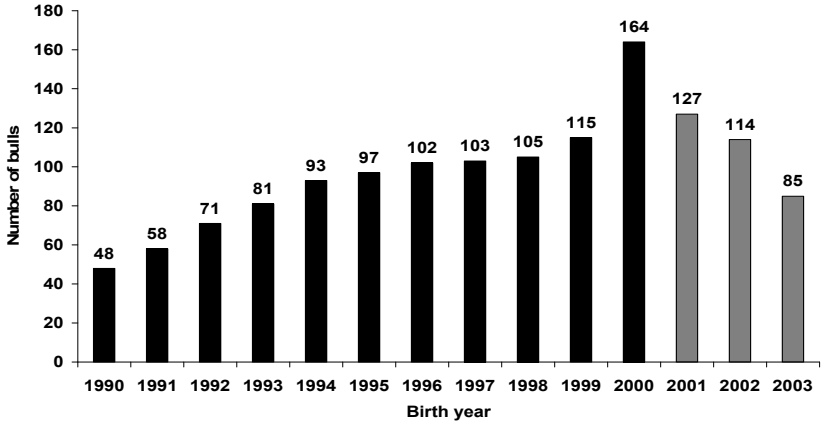


Figure 1. Distribution of bulls across to birth year (black and grey bars represent the number of bulls in the training and test set in forward prediction, respectively).

RESULTS AND DISCUSSION

Accuracies of genome-wide breeding values and regression coefficients for all traits using the 4 different methods applying cross validation are presented in Table 1. Accuracies were in the range of 0.30 to 0.74. BayesC was best to predict genome-wide breeding values for Fat%, whereas GS-BLUP, PLSR and LASSO gave similar, but lower accuracies. The highest accuracies were obtained for Prot-kg applying BayesC, GS-BLUP and PLSR. For the lowly heritable traits NR56 and SCC, all methods except LASSO performed equally in terms of accuracy. Using LASSO, where only a few numbers of SNP were selected to fit the data, accuracies for NR56 and SCC were 0.30 and 0.31, respectively.

Table 1. Accuracy (r) of genome-wide breeding values and regression coefficients (b) of the current estimated breeding value on the genome-wide breeding value with the genomic selection methods BayesC, GS-BLUP, PLSR and LASSO applying cross validation[1]

Trait[2]	BayesC		GS-BLUP		PLSR		LASSO	
	r	b	r	b	r	b	r	b
Fat%	0.64	0.92	0.52	0.93	0.50	0.83	0.56	1.55
Prot-kg	0.72	1.04	0.73	1.18	0.74	1.05	0.54	3.01
NR56	0.49	0.74	0.50	0.81	0.50	0.77	0.30	2.81
SCC	0.52	0.79	0.55	0.90	0.54	0.86	0.31	3.95

[1] Results are arithmetic means of the five-fold cross validation.
[2] Fat% = fat percentage; Prot-kg = protein kilogram; NR56 = non return rate after 56 days; SCC = somatic cell count.

From a practical point of view, animal breeders are more interested in forward prediction, i.e. prediction of genome-wide breeding values for young bulls which were not included in the derivation of the prediction equations. In Table 2, accuracies and regression coefficients for all traits and methods applied are shown for the forward prediction. In general, BayesC slightly outperformed the other methods. For Fat%, accuracy of genome-wide breeding values was 0.60, where all the other methods resulted in accuracies between 0.37 and 0.47. GS-BLUP gave similar accuracies for all traits, where surprisingly the highest accuracy was obtained for the very low heritable trait NR56. The same pattern was observed applying PLSR (Table 2). Lowest accuracies were calculated for Prot-kg, NR56 and SCC with LASSO. Using LASSO only a subset of SNP is included in the model (Tibshirani, 1996). For Fat%, Prot-kg, NR56, and SCC 25, 21, 20, and 21 SNP were selected, respectively. LASSO gave the highest accuracy of 0.47 for Fat% which might be in relation with the polymorphism in the DGAT1 gene which has a large effect on fat percentage (Grisart *et al*. 2004). Similar results were reported by Hayes (2009) where LASSO along with BayesC gave the highest accuracy for fat%. Regression coefficients for all traits and methods were below 1 indicating that genome-wide breeding values were overestimated with all methods used.

So far, only a very few results of genomic selection studies dealing with real data are available. Accuracies in this study were considerably lower compared to other studies. Sölkner *et al*. (2007) reported accuracies for Australian Holstein Friesian bulls in the range of 0.65 to 0.8 for different traits, including fertility, a trait with very low heritability, using different regression methods. Harris *et al*. (2008) have shown reliabilities (r^2) of genome-wide breeding values for young bulls without any daughter information in the range of 0.50 to 0.67 for milk production traits, live body weight, fertility, SCC, and longevity. In that study, Bayesian methods gave also slightly higher reliabilities compared to BLUP and regression methods. Hayes *et al*. (2009) observed reliabilities for Australian Holstein Friesian bulls for different traits between 0.14 and 0.55 using GS-BLUP and a Bayesian method (BayesA). A common finding of these studies was that GS-BLUP gave only slightly worse accuracies compared to Bayesian methods (Hayes *et al.* 2009; VanRaden *et al*. 2009). This is in agreement with the findings in this study, where, compared to BayesC, GS-BLUP resulted in similar accuracies for all traits except for Fat%. The GS-BLUP approach assumes a normal distribution of marker effects with the same variance for each marker (Meuwissen *et al*. 2001) whereas BayesC uses prior information about the distribution of marker effects allowing some markers having a big effect, whilst others having small effects (Meuwissen 2009). From this result, Hayes *et al*. (2009) conclude that the GS-BLUP assumptions, that most traits are influenced by many markers having a small effect and few with moderate to large effects, may be close to the truth.

Table 2. Accuracy (r) of genome-wide breeding values and regression coefficients (b) of the current estimated breeding value on the genome-wide breeding value with the genomic selection methods BayesC, GS-BLUP, PLSR and LASSO applying forward prediction

Trait[1]	BayesC		GS-BLUP		PLSR		LASSO	
	r	b	r	b	r	b	r	b
Fat%	0.60	0.81	0.42	0.78	0.37	0.64	0.47	0.70
Prot-kg	0.46	0.54	0.42	0.59	0.34	0.52	0.21	0.34
NR56	0.52	0.76	0.47	0.71	0.46	0.77	0.25	0.50
SCC	0.43	0.66	0.43	0.73	0.40	0.74	0.23	0.41

[1] Fat% = fat percentage, Prot-kg = protein kilogram, NR56 = non return rate after 56 days, SCC = somatic cell count

CONCLUSIONS

Considering the results for forward prediction, which are most relevant, BayesC turned out to predict the genome wide breeding values slightly more accurately than the other methods in this study. The LASSO method did not predict the genome wide breeding values very well except for Fat%. Results should be interpreted with caution as the analyses were based on a limited number of bulls. Further study is under way with the same methods and number of SNP increasing the number of bulls in the training set.

ACKNOWLEDGMENTS

The authors are grateful to ZuchtData EDV-Dienstleistungen GmbH and the Federation of Austrian Simmental Fleckvieh Cattle Breeders for providing the estimated breeding values and the genotypes.

REFERENCES

Grisart, B., Farnir, F., Karim, L., Cambisano, N., Kim, J.-J., Kvasz, A., Mni, M., Simon, P., Frère, J.-M., Coppieters, W., and Georges, M. (2004) *Proc. Natl. Acad. Sci. USA* **24**: 2398.

Hayes, B. J. (2009) *Symposium Statistical Genetics of Livestock for the Post-Genomic Era, Wisconsin-Madison, USA*

Hayes, B.J., Bowman, P.J., Chamberlain, A.J. and Goddard, M.E. (2009) *J. Dairy Sci*. **92**:433.

Harris, B.L., Johnson, D.L., and Spelman, R.J. (2008) Proc. *Interbull Meeting, Niagara Falls, Canada*

Meuwissen, T.H.E., Hayes, B.J. and Goddard, M.E. (2001) *Genetics* **157**:1819.

Meuwissen, T.H.E. (2009) *Genet. Sel. Evol*. **41**:35.

SAS Institute Inc. (2007) *SAS/STAT® User's Guide, Version 9.2. Cary, NC.*

Sölkner, J., Tier, B., Crump, R., Moser, G., Thomson, P.A. and Raadsma, H. (2007) *Book of Abstracts of the 58th Annual Meeting of the European Association for Animal Production*, p. 161.

Tibshirani, R. (1996) *J. R. Statist. Soc. B* **58**:267.

VanRaden P.M., Van Tassel, C.P., Wiggans, G.R., Sonstegard, T.S., Schnabel, R.D., Taylor, J.F. and Schenkel, F. (2009) *J. Dairy Sci*. **92**:16.

IDENTIFICATION OF SEX SPECIFIC DNA REGIONS IN THE SNAKE GENOME USING A SUBTRACTIVE HYBRIDIZATION TECHNIQUE

R.P. Harris, D.M. Groth, J. Ledger and C.Y. Lee

School of Biomedical Science, Curtin University of Technology, Bentley, WA 6021

SUMMARY

Karyotypic studies have shown that a ZZ/ZW sex chromosome system is used by snakes, which chromosomally resembles the ZZ/ZW system used by birds. However genetic studies have shown that *SOX 3*; the sex determining gene in mammals and *DMRT1*; which is the hypothesised sex determining gene in birds, are both located on autosomal chromosomes in the snake. Therefore it has been suggested that the snake sex chromosome system is unique and has evolved independently of the bird's system. This paper describes a subtractive hybridization method, using physical separation of biotinylated 'driver' DNA. The novel application of this technique was its use in identifying sex specific DNA regions within the genome of the Australian python; *Morelia spilota imbricate*. Female DNA enrichment was achieved using this technique and resulted in the identification of two non-sex specific repeating elements. The conclusion from this work is the identification of female specific DNA in snakes requires further subtractive hybridization enrichment and a more efficient screening procedure.

INTRODUCTION

Through karyotyping it has been shown that all species of snake use a common sex chromosome system, which consists of a single pair of ZZ (male) or ZW (female) sex chromosomes (Ezaz *et al.* 2006). Examination of the snake ZZ/ZW system has revealed that it is karyotypically similar to the ZZ/ZW sex chromosome system found in birds (Ezaz *et al* 2006). Therefore a theory that both bird and snake sex chromosomes have evolved from a common ancestor was conceived (Ezaz *et al.* 2006). However recent genetic research has revealed that *SOX 3*; the primary sex determining gene in mammals and *DMRT1*; which is hypothesised to be the sex determining gene in birds, are both located on autosomal chromosomes in the snake (Matsubara *et al.* 2006). It has now therefore been suggested that the snake sex chromosome system is unique and has evolved independently of the avian system despite its similarities (Matsubara *et al.* 2006). Our current research involves using a subtractive hybridization method, for the novel application of identifying sex specific DNA regions within the genome of the Australian carpet-python; *Morelia spilota imbricate*. By identifying sex specific regions we hope to assist in the study of snake sex chromosomes and the evolution of the snake's sex determination mechanism, which to date is still largely a mystery.

MATERIALS AND METHODS

Subtractive hybridization preparation. Male and female genomic DNA was isolated from ethanol preserved blood samples using white cell lysis buffer (10mM Tris.Hci, pH 7.6; 10mM EDTA, pH 8.0; 50mM NaCl; 200mg/ml Proteinase K (Promega); 0.1% SDS) followed by sequential phenol/chloroform extractions and isopropranol precipitation (Mathew 1985; Taggart *et al.* 1991). Male and female DNA was then double digested with 2.5 units *AluI* and *XmnI* and ligated with 50 units of T4 DNA Ligase to specifically designed double-stranded linkers RHZ (male linker) and RHW (female linker) respectively. This was performed as a single cycled reaction, to increase ligation efficiency, under the following thermal profile, 26× (16˚C for 15min, 37˚C for 30min); and heat inactivated at 65˚C for 20min.

Subtractive hybridization. A unique feature of the male linker RHZ, was biotin bound to the 5 prime end of the forward strand. This allowed the RHZ linker and ligated male DNA fragments to be almost irreversibly bound to streptavidin coated magnetic beads (Dynabeads®, Invitrogen), following manufactures instructions. Then female linker ligated DNA (2.5µg) was heat denatured and added to the magnetic beads coated with male DNA at a molar ratio of 2:1 male to female. The beads in the hybridization mixture were removed from the solution, after incubation with constant rotation at 55˚C for 1 hour, using a preheated magnetic particle concentrator (55˚C). The removed male DNA bound beads at this point contained female DNA fragments which have hybridized to the complementary male DNA and were discarded. The supernatant, enriched for unique female-specific DNA was used for repeated subtractive hybridization enrichments. This was achieved by replacing the female linker ligated DNA in the process described above with the enriched product from the previous subtractive hybridization. For subsequent enrichments fresh male DNA attached beads were used. The subtractive hybridization process was performed a total of 3 times at a stringency of 0.8×SSC (1.5M NaCl; 0.15M sodium citrate).

Screening. The sample enriched for female specific DNA was amplified with a RHW-F primer. The PCR was performed under the following conditions; volume equivalent to 15ng of template DNA before the subtractive hybridization cycles, 1.5mM $MgCl_2$, 200µM dNTPs, 0.6mM RHW-F primer, 1 unit Bioline DNA Polymerase, 1×PCR buffer (Promega) as recommended by the manufacturer and water to a final volume of 50µl. The PCR cycling conditions were: 1× (95˚C for 10min); 30× (94˚C for 45s, 55˚C for 45s, and 72˚C for 1.5 min); 1× (72˚C for 1hour extra extension for TA cloning); and held at 14˚C. Purified PCR product (3µl) was cloned using the TOPO TA Cloning Kit for Sequencing (Invitrogen). The TOPO cloning reaction (2µl) was used to electroporate TOP10 *E.coli* (20µl) (Invitrogen). The subsequent plating out of the transformed E.coli resulted in the isolation of 102 plasmids with potential sex specific DNA inserts. A dot blot analysis was performed to identify any sex specific DNA inserts. Only 51 of the 102 isolated identified were screened due to time constraints. Radioactive DNA probes of each plasmid isolate were prepared by incorporating ^{32}P label dCTP into a PCR reaction using M13 primers. Then each radioactive probe was hybridized to 2 male and 2 female genomic DNA samples. Subsequent DNA sequencing of selected plasmids followed.

RESULTS AND DISCUSSION

Subtractive hybridization preparation. An advantageous modification to the subtractive hybridization method was the combining of the genomic DNA digestion and linker ligation reactions together in a single cycled reaction. This cycling method was more efficient and less time and labor intensive then separating the digestion and ligation reactions. Firstly, de-phosphorylation of the genomic DNA to prevent concatamerization is not required in this method and secondly the introduction of XmnI prevents linker dimer formation, thereby increasing the efficiency of the linker ligation.

Subtractive hybridization. Using a molar ratio of 2:1, male to female DNA, three consecutive subtractive hybridization cycles were performed. At the end of each cycle a small aliquot of elution was kept for analysis, which can be seen in figure 1. When comparing the initial female linker ligated starting material (figure 1, lane 2), to the three sequential subtractive hybridization cycles (figure 1, lanes 3-5) a reduction in intensity of the DNA products is observed. This indicates a reduction in the concentration of female DNA, which decreases with each subsequent enrichment cycle. There is also a constant reduction in the observed fragment size range across the

three subtractive hybridization cycles.

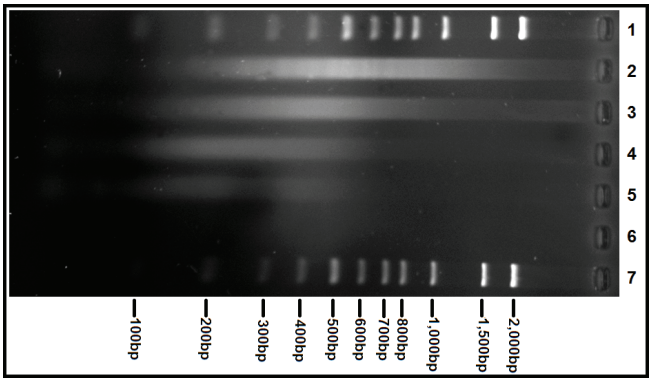


Figure 1. PCR of three consecutive subtractive hybridization cycles

Figure 1 shows the amplified product of a PCR using a RHW-F primer and the DNA from a RHW linker ligation (lane 2), elute from subtractive hybridization cycle 1 (lane 3), elute from subtractive hybridization cycle 2 (lane 4), elute from subtractive hybridization cycle 3 (lane 5) and a RHW negative control (lane 6). Lanes 1 and 7 are a molecular size marker (100kb ladder from Axygen).

Screening. Six of the 51 isolates screened for sex specific DNA gave a positive hybridization result, when analyzed using snake genomic DNA dot blots. This indicates an efficiency of 11.76% of the screening procedure, which was considerably lower than expected. Each blot had a probe positive control, which was found in virtually all instances to be very strong, indicating that the probe incorporated the radioactive tag. Subsequent investigations into this lowered efficiency revealed that the genomic DNA used for the membrane blots was of low quality and highly fragmented. Initial the DNA extractions were tested and concluded to be of highest quality; however, the DNA had degraded over time, due to its improper storage in purified water not a suitable buffer as required. Therefore, it is highly likely that the reason for the apparent lack of hybridization, in the negative blots was due to the use of degraded genomic DNA. It would be appropriate to repeat these experiments using pristine DNA, as it may reveal potential sex specific target sequences, however due to time limitations this was not possible.

Three clones, which did react with the genomic DNA on the dot blots, albeit at a lower signal strength, were DNA sequenced. DNA sequence analysis of the 731bp sequence from isolate 3E showed that it formed part of a probable repeating element. The repeating element is most likely a SINE or a LINE and was identified by BLAST analysis to match with a number of non-coding regions in a variety of unrelated organisms. This element is not unique to snakes and from the dot blot analysis was not identified to be sex specific. Sequence analysis of the isolate 5H resulted in the identification of another repeating element 140bp in length. This region was identified by BLAST analysis to occur in the intronic region of genes from two snake species and in a microsatellite region from a third snake species. In *Trimeresurus flavoviridis,* the sequence was found to have a high degree of similarity to sequences from two different intronic regions within the phospholipase A2 inhibitor gene. Furthermore, its coupling to a microsatellite sequence from a species of snake would indicate that it is most likely a SINE. Coupled consecutive SINE and microsatellite sequences have been reported to occur in other species. For instance 24% of dG-dT and dC-dA microsatellites found in the porcine genome are associated with the PRE-1 SINE and 5.6% are associated with Artiodactyl Repetitive Element 1 and 2 Porcine SINE (Alexander *et al*

1995). Furthermore, in cattle, horses and whales similar SINE elements have been found to be associated with microsatellites (Band and Ron 1996; Gallagher *et al.* 1999; Buchanan *et al. 2*001). Therefore, in conclusion it appears that these two elements are most likely of the SINE family. However, it also poses the question as to the efficiency of the capture process to remove these repetitive elements.

CONCLUSIONS

The aim of this experiment was to use a modified subtractive hybridization method to isolate and identify sex specific DNA regions from the genome of the Australian carpet python; *Morelia spilota imbricate*. Subtractive hybridization through the physical separation of biotinylated 'driver' DNA is not a novel technique and has been used previously to isolate tissue specific genes and in the recovery of full-length open reading frames (Rodriguez and Chader 1992; Meyer *et al.* 2007). However, the use of the subtractive hybridization technique, described in this paper, to isolate sex specific DNA regions of a complex genome is a novel application. This subtractive hybridization technique was significantly modified in order to increase its efficiency of capture and to assist in monitoring the progress of the experiment.

One of the major downfalls of the technique was the temperamental nature of the subtractive hybridization process and the difficulty in monitoring it. We chose to design sex specific linker and use PCR as a means to investigate and track the hybridization efficiency of the process. However, PCR amplify DNA fragments bound to the streptavidin coated magnetic beads proved to be difficult and resulted in us trying to investigate varying intensities of smeared DNA. It has been suggested that the process required more controls, in the form of a male on male and a female on female subtractive hybridization procedure. In order to identify how much DNA was being removed from each cycle. This would have increased the cost of the project and due to the reduced sensitivity of the PCR we believe no further conclusions would have been made by adding these controls. Ultimately a more sensitive process was needed to track the removal of the female DNA. A possible solution would be to label the female DNA with a fluorescent or radioactive molecule, so it could be tracked through the system more accurately.

Although the identification of two repeating elements did not satisfy the initial aim of this experiment, the detection of these regions showed that the subtractive hybridization technique was working to capture DNA. Initially it was thought that repeating elements common to both male and female would be removed during the subtractive hybridization process. However we showed that after three subtractive hybridization cycles at 50˚C, repeating elements still persisted. Therefore to eliminate more repeating elements and increase the efficiency of the technique we suggest adding one or two more subtractive hybridization cycles to the procedure and increasing the hybridization temperature by 5-10˚C. Future research will be focused towards re-performing the dot blot screening process, using non-degraded fresh DNA. As we hope snake sex specific DNA regions might still be identified within the 102 isolates.

REFERENCES

Band M. and Ron M. (1996) *Anim. Genet.* **27:**243.
Buchanan F., Crawford A., Strobeck C., Palsboll P. and Plante Y. (2001) *Anim. Genet*. **30:**47.
Ezaz T., Rami S., Veyrunes F. and Graves J. (2006) *Curr. Biol*. **16:**R736.
Gallagher P.C., Lear T.L., Coogle L.D. and Bailey E. (1999) *Mamm. Genome.* **10:**140.
Matsubara K., Tarui H., Toriba M., Yamada K., Nishida-Umehara C., Agata K. and Matsuda Y. (2006) *PNAS*. **103:**18190.
Meyer Q.C., Burton S.G. and Cowan D.A. (2007) *Biotechnol. J.* **2:**36.
Rodriguez I.R. and Chader G.J. (1992) *Nucleic. Acids. Res.* **20:**3528.

MERINO EWES DIVERGENTLY SELECTED FOR CALM TEMPERAMENT HAVE A GREATER CONCENTRATION OF IMMUNOGLOBULIN G IN THEIR COLOSTRUM THAN NERVOUS EWES

K.W. Hart[1,2], C. Contou[2], M. Blackberry[2] and D. Blache[2]

[1]Department of Agriculture and Food Western Australia, Narrogin, WA 6312
[2]School of Animal Biology, University of Western Australia, Crawley, WA 6009

SUMMARY

At birth, the concentration of serum immunoglobulins (Ig) in the plasma of lambs is close to nil. A lamb must absorb Ig from its mother's colostrum to acquire immunity. There is considerable variation in the concentration of Ig in the colostrum of different ewes and the serum concentration of a lamb is influenced by the concentration in its mother's colostrum and the quantity of colostrum available. While Ig concentration in colostrum is influenced by environmental factors there is also evidence to show there is some genetic control. Temperament of the ewe has been shown to influence other aspects of colostrum quality. We hypothesised that temperament would also influence the concentration of Immunoglobulin-G (IgG), the most abundant Ig in ewe colostrum. Ewes from selection lines for calm and nervous temperament from the University of Western Australia's Merino flock were in individual pens and fed at maintenance for a single bearing late pregnant ewe. A 50 ml sample of colostrum was collected from each ewe shortly after she gave birth. IgG concentration was analysed using single radial immunodiffusion (IDRing® SHEEP IgG Test Plates: IDBiotech). The mean IgG concentration in the colostrum of ewes in the calm line (35.69 ± 2.47 mg/ml) was greater than ewes in the nervous line (30.26 ± 1.45 mg/ml; $P < 0.05$). Lambs of calm ewes can acquire increased immunity compared to lambs of nervous ewes.

INTRODUCTION

When a lamb is born its serum Ig concentration is close to nil (Parker and Nicol 1990). Neonatal lambs passively acquire immune competency by absorbing Ig that are present in colostrum (Halliday 1974). The total quantity of Ig produced by the ewe is correlated to the volume of colostrum produced (Shubber *et. al*. 1979a, 1979b). The Ig concentration in colostrum is greatest during the lambs first feed and it rapidly declines in subsequent meals (Shubber *et al*. 1979b) with the concentration by the fifth feed only being about 6% of the level available in the first meal. It declines even further over subsequent meals. This means there is a strong correlation between the volume of colostrum consumed by the lamb at its first meal and its serum Ig concentration 30 hours after that meal (Shubber *et al*. 1979b).

As this rapid decline in colostral Ig concentration is occurring there is a simultaneous rapid decline in the ability of the lamb to absorb Ig (Dominguez *et al.* 2001). As soon as 12 hours after it is born a lamb's ability to absorb Ig is only 40% as efficient as it is 30 minutes after birth.

The decline in both colostral Ig concentration and the lambs ability to absorb Ig means a peak in serum Ig concentration occurs about 24 hours after the first meal (Cabello and Levieux 1981; Parker and Nicol 1990). Once this peak is reached there is then a steady decline in the lambs serum Ig concentration until the lamb begins to develop the capacity to produce its own Ig. This process does not begin until the lamb is about four to six weeks of age (Ducker and Frazer 1976). It is crucial for a lamb to have access to a good quantity of colostrum with a high concentration of Ig.

Neonatal deaths are not due to low levels of colostral Ig. Lamb deaths related to inadequate Ig occur during the first 5 to 6 weeks of life. About 14% of lamb deaths occur between 2 - 10 days of age and a further 5% after day 10 (Khalaf *et al*. 1979). Serum IgG concentration at 24 hours of age

was much lower (< 18 mg/ml) in lambs that died compared to lambs that survived beyond 10 days of age (> 20 mg/ml; Khalaf *et al.* 1979). In addition McGuire *et al*. (1983) found 45% of lambs died within three weeks of birth where serum IgG concentration was low compared to only 5% where the serum IgG levels were adequate. In fact, serum IgG concentrations are lower in lambs whose mothers had less than 30 mg/ml of IgG in their colostrum compared to lambs with mothers that had more than 110 mg/ml (McGuire et al. 1983). While the volume of colostrum produced is an important factor for lambs to acquire adequate serum Ig, colostral Ig concentration also has an important role. There are also important environmental factors that influence colostral Ig concentration. These include nutrition of the ewe (Hall *et al*. 1992), age of the ewe (Halliday 1976) and litter size (Hall *et al.* 1992). There is also evidence of genetic control of Ig concentration in both colostrum and lamb serum. The performance of ewes in the transfer of Ig to their lambs over different years has a moderate repeatability (0.3; P< 0.001; Halliday 1974).

Recently we have shown ewe temperament is associated with differences in some components of colostrum, such as lactose, fat and viscosity, that relate to its quality (Hart *et al*. 2006). In addition, previous observations over two or more lambing opportunities have shown a difference in lamb mortality from day 3 to weaning, when the mortality rate of lambs born to calm ewes was lower (5.7%) than the lambs of nervous ewes (22.9%, P < 0.001; Murphy 1999). One of the factors underlying the difference in medium-term lamb mortality between calm and nervous ewes in Murphy's experiments may be differences in colostral Ig concentration. We hypothesised that ewes with calm temperament will have greater colostral IgG concentrations than nervous ewes.

MATERIALS AND METHODS

This project was approved by the University of Western Australia's Animal Ethics Committee (Approval number: RA/3/100/466). Two groups of single bearing ewe's were established using the temperament selection lines from the University of Western Australia's Merino flock. There were 18 ewes in the calm group and 23 in the nervous group. Ewe age ranged from 3 to 7 years and was not different between groups (Table 1). All ewes had previously reared at least one lamb and the mean weight of ewes was no different between groups (Table 1). The ewes were synchronised for artificial insemination using intravaginal sponges containing 40 mg of Fluogestone Acetate (Chrono-gest® 40, Intervet, Australia) for 14 to 15 days. At sponge withdrawal the ewes were injected with 200 IU of eCG (Folligon®, Intervet, Australia). The ewes were inseminated by laparoscopic intrauterine insemination about 50 hours after sponge removal.

Two weeks prior to lambing the ewes were moved indoors into single pens. Ewes were fed one kilogram of pellets per day (SPECIALTY FEEDS; late pregnant ewe cubes), which contained 12% protein and 10.5 MJ/kg M.E. Oaten hay and water were provided *ad lib.*

The ewes were observed 24 hours per day during lambing. Within 5 minutes of giving birth a sample of approximately 50 ml of colostrum was collected from each ewe. Each sample had 50 µl of potassium dichromate added and was then frozen until required for analysis.

Prior to analysis the frozen samples of colostrum were thawed slowly. The concentration of IgG was determined using single radial immunodiffusion (IDRing® SHEEP IgG Test Plates: IDBiotech).

Statistical analysis. A comparison of the means for IgG concentration between the treatment groups was made using a one-tail heteroscedastic t test.

RESULTS

The mean colostral IgG concentration is greater for calm ewes than for nervous ewes (P < 0.05; Table 1). Both groups showed a large variation between individuals in the IgG concentration in colostrum that was available at birth. Calm ewes ranged from 20.05 to 60.04 mg/ml, while the

range among nervous ewes was from 20.93 to 44.69 mg/ml. Mean birth weight of the lambs are presented in Table 1 and were not different between the selection lines.

Table 1: Shows the means ewe body weight, ewe age, lamb birth weight and the IgG concentration of colostrum available at birth from Calm and Nervous ewes (± S.E.)

Treatment group	No.	Ewe WT	Ewe age	Lamb BWT	IgG mg/m
Calm	18	57.7 ± 1.28 kg	4.28 ± 0.32	5.04 + 0.20 kg	35.69 ± 2.47^{a}
Nervous	23	56.1 ± 1.31 kg	4.13 ± 0.29	5.33 + 0.24 kg	30.26 ± 1.45^{b}

Values in a column with different superscripts are significantly different ($P < 0.05$).

DISCUSSION

Our hypothesis that ewes with calm temperament will have a greater concentration of colostral IgG than nervous ewes was supported. Given the narrow window of opportunity for a lamb to acquire Ig it is imperative that the concentration in colostrum is sufficient to acquire an adequate serum concentration. These results suggest that up until lambs begin to synthesise Ig the lambs of calm ewes have an increased chance of survival compared to the lambs of nervous ewes .

Colostrum produced by calm ewes has a greater concentration of lactose, which leads to lower viscosity (Hart *et al.* 2006). These factors mean the colostrum of calm ewes provides distinct advantages to their lambs during the critical neonatal period. It has more energy available to meet the lambs' immediate metabolic requirements and uses less energy as it is easier to suck. Having a greater concentration of colostral Ig provides the lambs of calm ewes with a greater level of immune competency until the lamb is able to synthesise its own Ig. These factors improve the chances of survival for the lambs of calm ewes compared to those with nervous mothers.

While the colostral IgG concentration is different between the selection lines the results of both groups are toward the lower end of the acceptable levels (~30 mg/ml) indicated by McGuire *et al*. (1983). Several environmental factors have been identified as influencing colostral Ig concentration. These include litter size (Hall *et al*. 1992), ewe age (Halliday 1976) and nutrition (Hall *et al*. 1992). Litter size is excluded as a factor in this experiment as all ewes produced singles. The ewes in this experiment are between 3 to 7 years old. The mean age (± se) of the ewes was 4.2 ± 0.2 years and was not different between the groups. Ewe age is excluded as a factor as Halliday (1976) found that colostral IgG concentration was only reduced in ewes that were 8 to 9 years old. This leaves nutrition as a possible explanation for low colostral IgG concentration in these ewes. If the relatively low IgG concentrations in colostrum are due to a nutritional challenge it may be postulated that the result in this experiment is due to calm ewes being better able to cope with that challenge.

The volume of colostrum is increased where particular nutritional supplements are given to ewes late in pregnancy (Banchero *et al*. 2004). Just as other factors associated with colostrum production by Merino ewes are associated with temperament of the ewe it may be that the response to such supplements would also be influenced by temperament. This is an important question as Banchero *et al*. (2004) found the response to be greater in twin bearing ewes. If temperament is also related to differences in response to supplements this may provide a useful strategy to reduce the mortality of lambs born as multiples.

Shubber *et al*. (1979a) found that the greater the total volume of colostrum the greater the amount of Ig produced and this was irrespective of litter size. However the total yield of colostrum is influenced by litter size. While twin bearing ewes produce more colostrum there is less available for each lamb (Shubber *et al.* 1979a, 1979b; Banchero *et al*. 2004). This suggests that the concentration of Ig in colostrum becomes an important factor for the survival of lambs born as multiples. There is a degree of compensation for multiples, as lambs that are less mature at birth

have an increased window of opportunity to absorb Ig (Cabello and Levieux 1981). While this provides some degree of compensation to multiple lambs it is confounded by lamb vigour. Shubber *et al*. (1979a, 1979b) found wide variation in the volume of colostrum produced by triplet bearing ewes. Ewes whose lambs were heavier at birth and had a greater suckling drive produced more colostrum than ewes with less vigorous lambs. These are important questions because lambs from multiple litters experience higher levels of mortality than lambs born as singles. It will be multiples that benefit most by having a mother that produces a greater quantity of better quality colostrum.

CONCLUSIONS

The differences in Ig concentration between calm and nervous ewes are in many respects similar to other differences in characteristics that define the quality in colostrum. Just as calm ewes have been shown to have a greater lactose concentration and lower viscosity than nervous ewes, calm ewes also have greater Ig concentrations. This suggests that the lambs of calm ewes have an improved chance of survival during the critical neonatal period compared to the lambs of nervous ewes, because they have access to colostrum that has more energy and is easier to suck . By also having greater concentrations of colostral Ig calm ewes improve their lambs' chances of survival beyond the neonatal period by providing an increased level of immunity to the lamb until it begins to develop its own immune system. There are two studies needed to follow up this research. The first is to determine whether there is a difference in the response to nutritional supplements given during late pregnancy between calm and nervous ewes. The second is to verify that these differences also apply to ewes that rear multiple lambs. This is because these factors are even more important to the survival of multiple lambs as each lamb in the litter has a reduced amount of colostrum available to it when compared to singles.

ACKNOWLEDGEMENTS

Ken Hart is supported by Meat and Livestock Australia. Thank you to Steve Gray, the manager of Allandale Farm for his assistance and management of the sheep. Thankyou to Celine Lenoury, Samantha Bickell, Pascal Poindron, Raymond Nowak and Aprille Chadwick for their assistance in collecting colostrum samples.

REFERENCES

Banchero, G.E., Quintans, G., Martin, G.B., Lindsay, D.R. and Milton, J.T.B. (2004) Reprod. Fertil. Dev. **16:** 634.

Cabello, G. and Levieux, D. (1981) Res. Vet. Sci. **31:**190.

Dominguez, E. Perez, M.D., Puyol, P., Sanchez, L. and Calvo, M. (2001) Res. Vet. Sci. **70:**275.

Ducker, M.J. and Frazer, J. (1976) Anim. Prod. **22:**411.

Hall, D.G., Holst, P.J. and Shutt, D.A. (1992) Aust. J. Agric. Res. **43:**325.

Halliday, R. (1974) Anim. Prod. **19:**301.

Halliday, R. (1976) Res. Vet. Sci. **21:**331.

Hart, K.W., Chadwick, A., Sebe, F., Poindron, P., Nowak, R. and Blache, D. (2006) Aust. J. Exp. Agric. **46:**827.

Khalaf, A.M., Doxey, D.L., Baxter, J.T., Black, W.J.M., FitzSimons, J. and Ferguson, J.A. (1979a) Anim. Prod. **29:**401.

McGuire, T.C., Regnier, J., Kellom, T. and Gates, N.L. (1983) Am. J. Vet. Res. **44:**1065.

Murphy, P.M. (1999) PhD thesis. University of Western Australia.

Parker, R.J. and Nicol, A.M. (1990) Proc. N. Z. Soc. Anim. Prod. **50**:275.

Shubber, A.H., Doxey, D.L., Black, W.J.M. and FitzSimons, J. (1979a) Res. Vet. Sci.. **27:**280.

Shubber, A.H., Doxey, D.L., Black, W.J.M. and FitzSimons, J. (1979b) Res. Vet. Sci.. **27:**283.

SURVIVAL OF ADULT SHEEP IS DRIVEN BY LONGEVITY GENES

S. Hatcher, K.D. Atkins and K.J. Thornberry

NSW DPI, Orange Agricultural Institute, Forest Rd Orange NSW 2800

SUMMARY

Genetic parameters for survival of adult Merino wethers were estimated using annual 'roll-calls' based on continued presence at the annual research sampling of a mixed bloodline flock run in the cereal belt of central NSW. Mortality rates were consistently 4 % annually resulting in an average survival of 88 % to 5 years of age. Phenotypic variation within age periods (1-2, 2-3, 3-4 and 4-5 years of age) was very small as was variation between sires and flocks. Phenotypic variation in cumulative survival was higher and along with variation between sires tended to increase with increasing age as did the heritability (0.00, 0.07, 0.10 and 0.13 at 2, 3, 4 and 5 years of age respectively) which suggests that longevity genes are driving survival in adult Merino sheep.

INTRODUCTION

One of the implications of widespread adoption of precision sheep production systems in Australia (Rowe and Atkins 2006) will be a change from age based culling to variable age culling. At present, selected individuals, based on a select/cull decision after their first shearing, are kept in the flock until culling at a set age (Atkins *et al.* 2006) commonly 6 years. The set culling age is a compromise between age related changes in wool production and quality traits (Hatcher *et al.* 2005) and the opportunity for more intense selection among young animals when selecting replacements (Atkins *et al.* 2006). Such trade-off's between wool traits, selection differential and generation length have been the main determinants of the optimum age structure of a flock (Turner and Young 1969). However recent work (Lee *et al.* 2009) identified potential gains to be made in lifetime reproductive rate by retaining high performing ewes longer in the breeding flock. Consequently a system of variable age culling, where an individual animal's level of superiority over others governs the length of time it remains in the flock (Atkins *et al.* 2006), will be a necessary tool to optimise wool production and quality, reproductive performance and genetic progress in both the current and future generations. Under this scenario survival or longevity of individuals within a flock, particularly high performing individuals, becomes a critical issue. Functional traits, including fitness and longevity, of dairy cows have recently been extensively studied due to their impact on herd profitability (Coelho and Barbosa 2006; Essl 1998; Sewalem *et al.* 2006) and in response to evidence that exclusive selection of production traits caused a correlated reduction in longevity (Essl 1998; Wall *et al.* 2006). Given the long history of age based culling in Australian sheep production it is not surprising that there is no published work relating to aspects of survival or longevity in adult Merino sheep. This paper reports a preliminary analysis of the genetic basis of survival of adult Merino sheep.

MATERIALS AND METHODS

The data used in this study were 'roll-calls' of adult Merino wethers run at Condobolin Agricultural Research and Advisory Station (ARAS) between 1992 and 2001. Condobolin ARAS is located in the centre of the NSW cereal belt just above the floodplain of the Lachlan River at an elevation of 195m. The average annual rainfall recorded at the station is 424 mm and non-seasonally distributed, although high summer evaporation rates (8-10 mm per day) render much of the summer rain ineffective. Temperatures range from 2.7°C in July with an average 23 days below 0°C to 33.5 °C in January with an average of more than 5 days above 40°C. Grazing

animals rely on naturalised medics (*Medicago* spp), some native grasses which grow during autumn and spring and are offered grain and hay based rations where appropriate to maintain body condition. The majority of wethers (80% as 20% remained at Armidale) born between 1991 and 1996 in the CSIRO Fine Wool Project flock at Armidale (Swan *et al.* 2000) were transferred to Condobolin ARAS following their hogget shearing at 10 months of age where they remained for 4 consecutive shearings.

A tag list of all wethers trucked to Condobolin ARAS established the initial population of each drop (1991, 1993, 1994 1995 & 1996) of wethers at the station. The subsequent appearance or otherwise of each wether at the annual midside sampling, shearing and off-shears liveweight activities, undertaken over 2-3 weeks in August, was used to code each wether as either 0 = dead or 1 = alive at 2, 3, 4 and 5 years of age. The 1,800 wethers represented 11 bloodlines (6 superfine, 3 fine and 2 medium) and each had an identified sire. The data set contained 315 sires with an average of 29 progeny (range 4 - 132). Cumulative survival at each age and survival between ages of those wethers that survived to the previous age were analysed separately with a univariate binary analysis that included the fixed effects of flock (11 levels), year (5 levels) and the flock x year interaction using ASReml (Gilmour *et al.* 2006). Random effects were estimated for sire (σ^2_s) and flock (σ^2_f) in separate univariate analyses that included the appropriate fixed effects from the binary analyses. The heritability and standard errors were estimated using ASReml as 4 times the sire variance divided by the phenotypic variance. The phenotypic variance was the sum of the sire and within animal components.

RESULTS AND DISCUSSION

The average survival of the Merino wethers at Condobolin to 5 years of age was 87.6%. Mortality rates were reasonably consistent across ages being about 4 % per year (Table 1). The phenotypic variance was low within each age interval (range 0.002 to 0.052) highlighting the difficulty in partitioning variation at this low level of mortality. Similarly, due to very low estimates (4.4×10^{-10} to 8.1×10^{-4} and 7.5×10^{-13} to 5.1×10^{-9} for sire and flock respectively) there was effectively no between sire or between flock variation in survival within ages. In contrast phenotypic variation in cumulative survival increased with increasing age from 0.00 at 2 years to 0.11 at 5 years. Variation between sires also increased with increasing age but to a lesser extent, from 0.000 at 2 years to 0.003 at 5 years. There was little evidence of between flock variation in cumulative survival; the largest estimate was 3.3×10^{-5} at 5 years of age.

Table 1. Predicted values, variance components (± s.e.) and heritability (± s.e.) of survival of adult wethers within ages and cumulative survival

	Survival within ages			
	1 - 2 years	2 - 3 years	3 - 4 years	4 - 5 years
Average	99.61	99.64	99.81	99.64
Between sires (σ^2_s)	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.001
Phenotypic variance (σ^2_p)	0.006 ± 0.000	0.004 ± 0.000	0.002 ± 0.000	0.052 ± 0.002
Heritability (h^2)	0.000 ± 0.000	0.000 ± 0.000	0.021 ± 0.054	0.063 ± 0.054
	Cumulative survival to			
	2 years	3 years	4 years	5 years
Average	99.44	95.56	92.59	87.59
Between sires (σ^2_s)	0.000 ± 0.000	0.001 ± 0.000	0.002 ± 0.001	0.003 ± 0.001
Phenotypic variance (σ^2_p)	0.006 ± 0.000	0.043 ± 0.001	0.069 ± 0.002	0.108 ± 0.003
Heritability (h^2)	0.000 ± 0.000	0.073 ± 0.049	0.096 ± 0.050	0.126 ± 0.054

For survival within ages, flock was significant at all time periods ($P<0.01$ for 2 - 3 years and $P<0.001$ for all other time periods). Year was significant ($P<0.001$) for all age periods except 4 - 5 years old. Flock and year were both significant sources of variation in cumulative survival to 2 years of age ($P<0.001$) but not at older ages. Despite being statistically significant, these effects represent differences between flocks and years of 2 or 3 wethers dying. As a result there were no clear trends in survival evident between flocks or years. The flock x year interaction was not significant for survival within any age period or for cumulative survival at any age.

The heritability of adult survival, both within ages and cumulative, was negligible at 2 years of age but tended to increase with increasing age (Table 1). Heritability estimates ranged from 0 between 1 and 2 years of age to 0.06 between 4 and 5 years with the cumulative survival estimates ranging from 0 at 2 years of age up to 0.13 at 5 years of age. The increasing heritability estimates for cumulative survival with increasing age suggest that longevity genes are driving survival in adult Merino sheep. Clearly more research is required in this area to increase our understanding of adult survival and longevity in Merino sheep. Genetic and phenotypic relationships between longevity and other traits including wool production, wool quality and reproduction are required. While there are no published reports concerning relationships between wool traits and longevity, ecological studies of undomesticated species of sheep suggest that individuals with high lifetime reproductive performance have greater longevity (Clutton-Brock *et al.* 1996; Hamel *et al.* 2009). However from a Darwinian point of view artificial selection for production traits should generally lead to a deterioration of longevity (Essl 1998). This has certainly been demonstrated in dairy cattle where selection for production traits including milk yield, live weight and growth rate resulted in negative correlated responses in fitness or longevity traits (Essl 1998; Wall *et al.* 2006). Similar effects may well occur in Merino sheep as Hatcher and Atkins (2007) found that phenotypic selection for high clean fleece weight leads to fewer progeny surviving to weaning. Lee *et al.* (2009) quantified heterogeneity of lifetime reproduction in Merino sheep. Similar analyses of lifetime performance in other traits of economic importance including wool production, wool quality and disease resistance are warranted together with their relationships with longevity. These relationships will allow the benefits of variable age culling to precision production systems for sheep to be accurately quantified. The challenge will be to identify existing datasets where individuals (ewes and wethers) were kept in the flock beyond 6 years of age that are adequate for survival analysis. It would also be interesting to compare genetic parameters for adult ewe survival with that of wethers to determine the effect of ewe selection for reproduction traits.

This analysis has demonstrated an annual death rate of 4% for adult Merino wethers grazing in a relatively benign environment. Indeed a mortality rate of 4% seems to be the 'acceptable norm' across the Australian sheep industry. However, in other production regions prevailing environmental conditions may well result in higher death rates, particularly for adult ewes that have to contend on an annual basis with the high physiological burden of pregnancy and lactation.

ACKNOWLEDGEMENTS

The contribution of numerous staff from CSIRO Livestock Industries (Armidale) and NSW DPI (Orange) who ably assisted with the collection of data during sampling and shearing for the 9 year during of the Condobolin Fine Wool flock made this and other analyses possible. Their assistance is gratefully acknowledged. We particularly would like to acknowledge the contribution of Laurie Barwick who was responsible for the day -to-day management of the flock at Condobolin and Heather Brewer (CSIRO Armidale) for suppling data from the CSIRO database.

REFERENCES

Atkins, K.D., Richards, J.S. and Semple, S.J. (2006) In '8th World Congress on Genetics Applied to Livestock Production'. Belo Horizonte, MG Brazil pp. 05-01.

Clutton-Brock, T.H., Stevenson, I.R., Marrow, P., MacColl, A.D., Houston, A.I. and McNamara, J.M. (1996) *J. Anim. Ecol.* **65:**675.

Coelho, J.G. and Barbosa, P.F. (2006) In '8th Wolrd Congress on Genetics Applied to Livestock Production'. Belo Horizonte, Minas Gerais, Brazil pp. 01-52.

Essl, A. (1998) *Livest. Prod. Sci.* **57:**79.

Gilmour AR, Gogel BJ, Cullis BR, Thompson R (2006) 'ASReml User Guide Release 2.0.' (VSN International Ltd, Hemel Hempstead, HP1 1ES UK).

Hamel, S., Cote, S.D., Gaillard, J.M. and Festa-Bianchet, M. (2009) *J. Anim. Ecol.* **78**:143.

Hatcher, S. and Atkins, K.D. (2007) *Proc. Assoc. Advmt. Anim. Breed. Genet.* **17**:260.

Hatcher, S., Atkins, K.D. and Thornberry, K.J. (2005) *Proc. Assoc. Advmt. Anim. Breed. Genet.* **16**:314.

Lee, G.J., Atkins, K.D. and Sladek, M.A. (2009) *Anim. Prod. Sci.* **49**, In press.

Rowe, J.B. and Atkins, K.D. (2006) In 'Australian Society of Animal Production 26th Biennial Conference 2006'. Perth, W.A. p. Short Communication number 33. (Australian Society of Animal Production).

Sewalem, A., Kistemaker, G.J., Miglior, F. and Van Doormaal, B.J. (2006) In '8th World Congress on Genetics Applied to Animal Production'. Belo Horizonte, Minas Gerais, Brazil pp. 01-22.

Swan, A.A., Purvis, I.W., Piper, L.R., Lamb, P.R. and Robinson, G.A. (2000) In 'Finewool 2000 Breeding for Customer Needs. ' Armidale NSW. (Eds AA Swan, LR Piper) pp. 65-73. (CSIRO Livestock Industries and The Woolmark Company).

Turner, H.N. and Young, S.S.Y. (1969) Quantitative genetics in sheep breeding. (Macmillan of Australia: Victoria).

Wall, E., Brotherstone, S. and Coffey, M.P. (2006) In '8th World Congress on Genetics Applied to Animal Production'. Belo, Horizonte, Minas Gerais, Brazil pp. 01-10.

GENETIC ASSOCIATION OF NET FEED INTAKE MEASURED AT TWO STAGES WITH INSULIN-LIKE GROWTH FACTOR –I, GROWTH AND ULTRASOUND SCANNED TRAITS IN ANGUS CATTLE.

M. G. Jeyaruban, D. J. Johnston and H.–U. Graser

Animal Genetics and Breeding Unit[1] , University of New England, Armidale, NSW 2351

SUMMARY

Net feed intake (NFI) is a measure of feed efficiency in beef cattle, calculated as the amount of feed eaten by an animal after adjusting for its growth rate and body weight. NFI was measured in weaned bulls and heifers of less than one year of age (NFI-P) and also in feedlot finishing steers mainly 18 months and older (NFI-F). Genetic parameters and their genetic relationship with insulin–like growth factor-I (IGF-I) around weaning age, growth traits and scanned traits were estimated for NFI-P and NFI-F. Estimated heritabilities for NFI-P and NFI-F were 0.41±0.05 and 0.34±0.09, respectively and the genetic correlation between the two was 0.65±0.14. Genetic correlations between NFI-P and IGF-I, ultrasound scanned subcutaneous fat depth at the rump (P8) in males and females and intramuscular fat percentage (IMF) in male and females were 0.18±0.11, 0.50±0.14, 0.49±0.10, 0.48±0.18 and 0.27±0.13, respectively. Genetic correlations between NFI-F and IGF-I, P8 in males and females and IMF in males and females were -0.14±0.18, 0.43±0.17, -0.13±0.14, 0.36±0.23 and -0.22±0.18 respectively. Both NFI-P and NFI-F had negative genetic correlation with weights at birth, 200, 400 and 600 days of age. This study showed that although the NFI measured at an early age and in the feedlot were moderately heritable, they were two different traits with varying genetic associations with growth and ultrasound scanned traits. Selecting for lower NFI at either stage will genetically decrease the ultrasound scanned fat traits in males. Selection for lower NFI-P can decrease fatness in heifers but this association was not evident with NFI-F. IGF-I, as recorded in this study, would have limited use as a genetic indicator trait for NFI in beef cattle due to genetic correlation differing in direction and of low magnitude with NFI-P and NFI-F.

INTRODUCTION

Feed cost constitutes the single largest expense in most beef cattle production systems. Any reduction in feed cost is expected to improve beef production profitability. The existence of variation for feed intake among animals of the same breed, sex and age class indicates that improvement for higher feed efficiency could be achieved through selection (Johnston *et al*. 2001). This individual variation in feed intake is being utilized in the form of net feed intake (NFI) to improve feed efficiency in beef cattle (Robinson and Oddy 2004). NFI is the difference between actual feed intake and the expected feed requirement for the growth rate and maintenance of body weight (Koch *et al*. 1963).

In Australia, NFI has been measured at two different stages of growth: in young post weaning bulls and heifers at around 300 days of age (NFI-P) and in feedlot finishing steers at around 18 months of age (NFI-F). As NFI-P and NFI-F are measured at two different stages of growth and maturity, they may not be genetically the same trait and could have different genetic associations with insulin–like growth factor I (IGF-I) and other production traits. Selection for lower NFI-P or NFI-F may therefore, produce different correlated responses in production traits. This study aimed to quantify the genetic associations of NFI-P and NFI-F with IGF-I, growth traits and ultrasound scanned fat traits.

[1] AGBU is a joint venture of the NSW Department of Primary Industries and University of New England

MATERIALS AND METHODS

Data used for this study were submitted to the Angus Society of Australia database until June 2006 for BREEDPLAN evaluation of NFI in Angus cattle. Feed efficiency test records for young bulls and heifers, fed at around 300 days of age with a ration containing an energy level of 10MJ/kg provided data for NFI-P. Steers fed at around 560 days of age, with a ration containing an energy level of 12MJ/kg provided data for NFI-F. Individual feed intakes for both groups of animals were measured over a 70 day feeding period, after an initial acclimatisation period. NFI-P and NFI-F were derived by adjusting the feed intake for the growth rate and metabolic mid test weight of individuals estimated for the two traits separately. Metabolic mid weight was calculated as the mid weight to the power of 0.73.

Plasma IGF-I levels were measured from blood samples obtained prior to weaning in a large number of seedstock herds. Testing age ranged from 150 to 250 days. Plasma IGF-I concentration was determined by using a commercially available Enzyme Linked Immunosorbent assay (ELISA). Growth traits analysed were birth weight (BWT), 200-day weight (200D), 400-day weight (400D) and 600-day weight (600D) with age at recording ranging from 80 to 300 days, 301 to 500 days and 501 to 700 days for 200D, 400D and 600D, respectively. For animals with multiple records for 200D, 400D and 600D, only their first record was used. Real time ultrasound scanned fat measurements included fat depth at the P8 (rump) site for bulls (BP8) and heifers (HP8), intramuscular fat percent at the $12/13^{th}$ rib for bulls (BIMF) and heifers (HIMF) and longissimus muscle area at the $12/13^{th}$ rib of bulls (BEMA) and heifers (HEMA) with age at recording for all traits ranging from 300 to 800 days. A relatively small number of real time ultrasound scanned carcass measurements of steers were grouped with heifer measurements.

For each trait, sires with recorded progeny for NFI-P, NFI-F and IGF-I were identified as common sires, and all data and pedigree from those herds containing progeny of common sires were extracted to build genetically linked data for each bivariate analysis. There were 242, 180 and 479 sires with progeny recorded for NFI-P, NFI-F and IGF-I respectively and 42 common sires with progeny recorded for NFI-P and NFI-F. The number of common sires between NFI-P, NFI-F and ultrasound scanned traits ranged from 130 to 159 and for growth traits ranged from 159 to 184. Number of records and descriptive statistics for all traits are given in Table 1.

Genetic variances, genetic correlations and variance ratios were estimated by restricted maximum likelihood (REML) using a series of bivariate animal model evaluations, with three generations of pedigree in ASReml (Gilmour *et al.* 2006). Records for growth and ultrasound scanned traits were pre adjusted for animal age and dam age using standard BREEDPLAN procedures (Graser *et al*. 2005). Models with fixed effect of contemporary group, and random additive genetic effect of animal were used for all traits. Contemporary group definitions for each growth and ultrasound scanned traits were as defined in BREEDPLAN (Graser *et al*. 2005) and those for NFI-P, NFI-F and IGF-I were defined by Moore *et al*. (2005). Additional random effects for BWT, 200D and 400D weights were maternal genetic and permanent environmental effects.

RESULTS AND DISCUSSION

Estimated heritabilities were moderate for all traits (Table 1). Heritability estimates of 0.41 and 0.34 for NFI-P and NFI-F were consistent with the value of 0.38 reported by Arthur *et al.* (2001) for NFI in Angus cattle. The moderate heritability estimates for NFI-P and NFI-F suggest that selecting for lower NFI at either stage will improve the feed efficiency in beef cattle.

The genetic correlation between NFI-P and NFI-F was high, but different from unity, indicating that these two feed efficiency measures are genetically different. This is further supported by the differences in genetic correlations of these traits with IGF-I, growth and ultrasound scanned traits. NFI-P had a low positive genetic correlation with IGF-I, while NFI-F

had a low negative genetic correlation with IGF-I. The positive genetic correlation between NFI-P and IGF-I is in agreement with other literature estimates, however, it is lower than the correlation of 0.42 reported by Moore *et al*. (2005) for Angus cattle in Australia. IGF-I samples used in this study included those of Moore *et al.* (2005) with additional samples collected by seedstock producers between 150 to 250 days of age, where actual weaning date was not known. Furthermore, NFI used in Moore *et al.* (2005) study was derived by pooling the NFI-P and NFI-F records.

Table 1. Number of records, the descriptive statistics of data , heritability (h^2) and the genetic correlation (using bivariate analysis) for postweaning net feed intake (NFI-P), feedlot net feed intake (NFI-F), insulin like growth factor-I (IGF-I) growth and ultrasound scanned traits (approximate standard errors in parenthesis).

Trait	No.	Mean	SD	Min	Max	h^2	Genetic correlation with NFI-P	Genetic correlation with NFI-F
NFI-P (kg/day)	2030	0.25	1.12	-3.9	3.6	0.41 (0.05)		0.65 (0.14)
NFI-F (kg/day)	1220	-1.09	2.21	-9.4	5.1	0.34 (0.09)	0.65 (0.14)	
IGF-I (ng/ml)	9216	393.56	190.52	27.0	1287.0	0.36 (0.04)	0.18 (0.11)	-0.14 (0.18)
Scanned traits								
HIMF (%)	22504	4.37	1.86	0.0	12.7	0.30 (0.02)	0.27 (0.13)	-0.22 (0.18)
BIMF (%)	14759	2.73	1.64	0.0	10.1	0.21 (0.02)	0.48 (0.18)	0.36 (0.23)
HEMA (cm^2)	28126	58.25	8.47	29.3	101.4	0.29 (0.01)	0.04 (0.12)	-0.01 (0.15)
BEMA (cm^2)	19579	75.91	10.11	33.0	113.8	0.26 (0.02)	-0.16 (0.17)	-0.01 (0.21)
HP8 (mm)	28424	6.20	2.93	0.0	41.9	0.44 (0.02)	0.49 (0.10)	-0.13 (0.14)
BP8 (mm)	19583	4.10	1.91	0.0	36.6	0.42 (0.02)	0.50 (0.14)	0.43 (0.17)
Growth traits								
BWT (kg)	74261	37.1	5.4	15.2	65.2	0.37 (0.01)	-0.04 (0.08)	-0.34 (0.13)
200D (kg)	89958	232.2	36.8	69.0	452.9	0.22 (0.01)	-0.05 (0.09)	-0.23 (0.13)
400D (kg)	64748	359.3	71.2	114.0	669.5	0.29 (0.01)	0.00 (0.08)	-0.16 (0.12)
600D (kg)	41106	496.9	94.4	217.6	886.4	0.40 (0.01)	-0.02 (0.10)	-0.25 (0.14)

Estimated genetic correlations of NFI-P and NFI-F with ultrasound scanned traits were variable in sign and magnitude. Ultrasound scanned fat traits had positive genetic correlations with NFI-P, with the estimates ranging from 0.27 for HIMF to 0.50 with BP8. Moderate to high genetic correlations between NFI-P and ultrasound scanned fat traits are in agreement with the correlations reported by Robinson and Oddy (2004). NFI-F had low negative genetic correlations with HIMF and HP8 in contrast to positive correlations with BIMF and BP8. The obvious differences between NFI-P and NFI-F were their correlation with P8 and IMF in females, where selection for lower NFI-P would decrease P8 but this association was not evident with NFI-F. Lower number of records in NFI-F increased the standard errors for the correlation of NFI-F with other traits. The HEMA and BEMA had low or no correlation with both NFIs. Both NFI-P and NFI-F had negative genetic correlations with growth traits, however the magnitudes of the correlations were different for the two feed efficiency measures. However, these estimates, combined with the non-unity correlation between NFI-P and NFI-F suggest that selection on the different net feed intake measures may lead to different correlated response on growth and ultrasound scanned traits.

CONCLUSIONS

Net feed intake measured in weaned young bulls and heifers and feedlot finished steers are moderately heritable. Genetic correlations less than unity and different genetic associations with growth and ultrasound scanned traits, however, indicate that these two feed efficiency measurements are genetically different. Improving feed efficiency at these two stages might yield different correlated responses in growth and ultrasound scanned traits. Selecting animals for lower NFI at either stage will genetically decrease ultrasound scanned fat traits in males, but selecting for lower NFI-F may increase P8 and IMF in females. Therefore, including both NFI-P and NFI-F in a genetic evaluation system is important to incorporate these changes. IGF-I, which had low and opposite different genetic correlations with NFI-P and NFI-F, is of limited use as a genetic indicator trait of NFI.

ACKNOWLEDGEMENT

The authors would like to thank the Meat and Livestock Australia (MLA) for their financial support and the Angus Society of Australia for providing data for this study. We also thank Primegro™ for analysing the blood samples to determine IGF-I concentration.

REFERENCES

Arthur, P.F., Archer, J.A., Johnston, D.J., Herd, R.M., Richardson, E.C. and Parnell, P.F. (2001) *J.Anim. Sci.* **79**:2805.

Gilmour, A.R., Cullis, B.R., Welham, S.J. and Thomson, R. (2006) "ASReml User Guide", Release 2.0 VSN international Ltd, Hemel Hempstead, HP1 1ES, UK.

Graser, H-U., Tier, B., Johnston, D.J. and Barwick, S.A. (2005) *Aust. J.Exp.Agric*. **45**:913.

Johnston, D.J., Herd, R.M., Reverter, A. and Oddy, V.H. (2001) *Proc. Assoc. Advmt. Anim. Breed Genet.* **14**:163.

Koch, R.M., Swiger, L.A., Chambers, D. and Gregory, K.E. (1963) *J. Anim. Sci*. **22**:486.

Robinson, D.L. and Oddy, V.H. (2004) *Livest. Prod. Sci*. **90**:255.

Moore, K.L., Johnston, D.J., Graser, H-U. and Herd, R. (2005) *Aust, J. Agri. Res*. **56**:211.

PREDICTING ENERGY BALANCE IN GROWING WETHERS AND ESTIMATION OF HERITABILITY FOR DERIVED PARAMETERS

E. Jonas, P.C. Thomson, K.J. Fullard, C.A. Cavanagh, and H.W. Raadsma

ReproGen, Faculty of Veterinary Science, University of Sydney, Camden, NSW 2570

SUMMARY

The aim of this study was to characterize the energy balance of growing sheep for three different models (CSIRO, CNCPS-s, AFRC) using estimates of feed intake and individual animal weight in an automatic feeder system. Data from 610 wethers of seven cohorts from an Awassi-Merino resource population were used during feed intake experiments. Aspects of data analysis and handling from computerized systems are described to calculate growth, feed intake, residual feed intake (RFI), predicted dry matter intake (DMI) and feed conversion efficiency (FCR). Three models were used to calculate energy balance of 610 progeny from seven half sib families and as a basis for phenotypic correlation and heritability estimates. Two different prediction equations for EB were derived from each model, a preliminary energy balance describing the difference between energy intake and energy for maintenance (EB I) and an approach describing the energy surplus as the difference between feed intake, energy for maintenance and growth (EB II). All three models gave similar predictions for EB I with phenotypic correlations >0.9, whereas greater differences between the EB II models with phenotypic correlations of -0.87, -0.47 and 0.52 were observed. Heritability was estimated between 0.26 and 0.37 for the EB I, 0.17 to 0.52 for EB II, 0.51 for body weight, 0.45 for feed intake, 0.16 for RFI, 0.34 for DMI, and 0.55 for FCR. We conclude that EB I and FCR under either of the three models should be appropriate for QTL analyses in growing sheep under a computerized feed intake system. We also conclude that an exact determination of the energy for body mass gain must be reconsidered before variables calculated using this trait can be used as phenotypes for further genetic analysis.

INTRODUCTION

Sheep production is a major contributor to global food production and sheep meat one of the few sources of meat with no cultural and religious restrictions in consumption. Feed input costs represent a significant cost factor to sheep production enterprises and selection for improved feed and energy utilization or generally termed feed conversion efficiency, is potentially an attractive but complex avenue for improving sheep profitability. A fundamental requirement for consideration in genetic improvement systems is a clear definition of the phenotype for both research parameter estimation and application in profit functions.

A number of methods have been published to describe the energy balance in cattle, especially dairy cattle examples as reviewed by Brosh (2007), and Van Knegsel *et al*. (2005), however only few methods have been adapted for sheep, even though similar approaches can be used in both species. Traditionally there are two ways to measure the energy balance using the input-output measures and the calculation of changes in body tissue composition. The most detailed input-output method is described by the modification of the Cornell Net Carbohydrate and Protein System (CNCPS), a mechanistic model that predicts nutrient requirements and biological values of feed for cattle, and has been adapted for the application in sheep (Cannas *et al.* 2004). This system accounts for differences in feeds of diverse characteristics fed at different levels of intake across a wide range of animal physiological states and environmental effects. Alternative systems to predict the energy balance are the application of residual feed intake (RFI) (Sainz and Paulino 2004). The aims of this study were as follows: 1) to evaluate the energy balance in growing sheep using three different input-output models, 2) to calculate input components of residual feed intake, and 3) to

calculate phenotypic correlations and heritability for energy balance estimates. In future, genome wide linkage and association studies for these traits will be performed using animals of the same population in an attempt to characterize the genetic architecture affecting these complex traits and identify genetic markers for marker assisted selection.

MATERIAL AND METHODS

The data used in this study were collected at the University Sydney research farm 'Mayfarm' at Camden, New South Wales, Australia between 2005 and 2009 using animals derived from an Awassi × Merino resource population (Raadsma *et al*. 2009). The information on body weight and feed intake were recorded using an automatic feeder system without any restriction of food available. The feeder experiments were performed during different growth phases at the time the animals entered the feeding system, ranging from 5 months till 2 years of age for a period of 70-90 days. To model the phenotype, the present study used observations from 610 animals. To stabilize the residual variances, estimates of the daily body weight and daily feed intake of each animal were summarized to weekly average values. Growth was calculated as the difference between two weekly averaged body weights.

Two different parameters describing energy balance were calculated based on the input-output model described by CSIRO (CSIRO 2007), the adapted formula of the CNCPS-S model (Cannas *et al.* 2004) and following the advisory manual prepared by the AFRC (AFRC 1993). The first main difference between the models is the calculation of the energy loss for maintenance, which is calculated as shrunken body weight in the model from CSIRO and as the sum of metabolism and activity energy using the CNCPS-S model. Difference between feed intake and energy loss for maintenance was further defined as energy balance I (EB I). The second greatest difference between the models is the estimation of energy requirements for growth, which is calculated using feed intake, full body weight and weight gain in the model of CSIRO, using body weight change, feed intake and maintenance energy in the model of CNCPS-s, and using body weight in the model of AFRC. Difference between feed intake and energy loss for maintenance and growth was further defined as energy balance II (EB II). The difference between the observed and predicted dry matter intake (DMI) was calculated following the description of CNCPS-S (Cannas *et al.* 2004). The residual feed intake (RFI) was calculated as the relation between feed intake and expected feed intake, which included here energy for growth (assuming 12.5 MJ ME required/kg weight gain) and energy for maintenance. Indices of feed conversion (FCR) were calculated as the amount of feed (kg) per kg body weight gain or loss and as additional feed (kg) required for each unit body weight.

All analyses were performed using the program R (version 2.6.0) and GenStat (10^{th} edition). Mixed models were applied to the data to estimate the repeatability of each trait using animal as a random factor in SAS (version 9.2). Further the mean of the values of each animal were used to estimate heritabilities in ASREML using a sire model.

RESULTS AND DISCUSSION

To compare the energy balance models, data from 610 animals over a period up to 17 weeks (average 47 days) were used. Animals had an average body weight of 57 kg. During the experiment animals gained in average 9.5 kg. The summary of the data from the feeder experiment are shown in Table 1. Using the estimates of residual feed intake, slight differences between the predicted and the measured feed intake were observed, whereas the difference between the predicted and the observed dry matter intake were smaller. Observed feed intake for growth was in the expected range with a mean of 8 kg feed intake/kg body weight gain.

Table 1. Body weight [kg] and feed intake [kg] energy balance I and II (CSIRO, AFRC and CNCPS-s) [MJ], and feed conversation rates, shown are number of animals (N), minimum (min), maximum (max), average (mean), standard deviation (std)

Trait	N	min	max	mean	std
Body weight beginning (BWT0)	8152	29.42	101.84	50.59	11.27
Body weight end (BWTe)	8142	35.38	103.08	60.03	11.32
BWTe-BWT0 (Growth)	8142	-9.56	23.34	9.45	6.03
Mean body weight (BWT)	6750	29.20	103.22	57.27	11.62
Weekly body weight change (Growth rate)	5299	-3.00	5.00	0.98	1.76
Feed Intake (FI)	6691	0.021	2.98	1.14	0.50
Feed conversion (feed intake / growth)	4958	0.01	49.86	7.90	8.24
Feed intake / body weight (gross efficiency)	6691	0.30	960.8	133.9	76.02
Energy available for growth (CSIRO) (EB I)	6691	-7.87	28.03	6.85	5.53
Energy surplus (CSIRO) (EB II)	6691	-34.45	4.83	-18.55	5.62
Energy available for growth (CNCPS-s) (EB I)	6691	-6.32	27.58	7.21	5.42
Energy surplus (CNCPS-s) (EB II)	6691	-34.92	-0.02	-9.24	5.53
Energy available for growth (AFRC) (EB I)	6691	-8.48	27.69	6.41	5.53
Energy for surplus (AFRC) (EB II)	5275	-95.54	54.68	3.37	9.09
Observed dry matter intake / Predicted dry matter intake	6691	0.01	2.76	0.85	0.37
Residual feed intake	5147	-17.90	19.98	2.16	2.42

Table 2. Heritability (h^2), repeatability (t), and phenotypic correlations between of body weight (BWT), growth (GR), feed intake (FI), feed conversion rate (FCR) and feed intake/kg body mass (FI/BWT), dry matter intake (DMI), residual feed intake (RFI), energy balance I and II (EB I and EB II), DMI, and RFI; ne: not estimated

Trait	h^2	t	BWT	GR	FI	FCR	FI/BWT	DMI	RFI
BWT - mean	0.51	0.87	-	0.01	0.15	0.09	-0.19	-0.17	-0.12
Growth -GR	0.14	ne	0.01	-	0.43	0.06	0.32	0.39	0.04
FI	0.45	0.2	0.15	0.43	-	0.26	0.80	0.94	0.25
FCR	0.55	0.002	0.09	0.06	0.26	-	0.25	0.23	0.13
FI/BWT	0.55	0.24	-0.19	0.32	0.80	0.25	-	0.87	0.29
DMI	0.35	0.14	-0.17	0.39	0.94	0.23	0.87	-	0.30
RFI	0.16	0.02	-0.12	0.04	0.25	0.13	0.29	0.30	-
EB I CSIRO	0.32	0.16	-0.01	0.43	0.99	0.25	0.84	0.98	0.27
EB I CNCPS	0.26	0.17	0.05	0.43	0.99	0.25	0.83	0.96	0.27
EB I AFRC	0.37	0.16	-0.03	0.43	0.98	0.24	0.85	0.98	0.27
EB II CSIRO	0.34	0.13	-0.19	0.39	0.94	0.23	0.86	0.99	0.29
EB II CNCPS	0.18	0.14	-0.04	-0.39	-0.90	-0.21	-0.77	-0.88	-0.24
EB II AFRC	0.52	0.007	-0.11	0.07	0.49	0.35	0.50	0.53	0.45

Only slight differences between the energy balance models of CSIRO, CNCPS-s and AFRC were found for EBI, resulting in very similar parameters for predicted energy balance (Table 1). Prediction of EB II on the other hand showed a major difference for estimates from all three models (Table 1) Phenotypic correlations between derived EB I from the different models was high (> 0.99). The EB II showed greater differences with high negative phenotypic correlations (-

0.87) between the CSIRO and CNCPS-s model, moderate correlation (0.52) between the CSIRO and AFRC model, and moderate negative correlated (-0.47) between the CNCPS-s and AFRC model.

The repeatability of all traits taken on a weekly basis using 610 animals and 8267 observations are shown in Table 2. The phenotypic correlations and heritability estimates for all main traits are also shown in Table 2. The heritability for EB I and EB II was low to moderate (0.18 to 0.52). Cammack *et al.* (2005) estimated a slightly lower heritability of RFI (0.11), compared to our study (0.16). In another study, total feed intake, and feed conversion ratio were moderate heritable (0.39 and 0.26) (Snowder and Van Vleck 2003). In our study the estimate for feed intake and FCR was higher (0.45 and 0.55).

CONCLUSION

The application of three different models to estimate the energy balance in growing wethers using data from an automatic computerized feed intake and body weight system predicted similar levels of energy balance in sheep allowing for maintenance. But greater differences were seen in Energy balance among the models allowing for maintenance and growth. The data could be used to predict residual feed intake and indicators of feed conversion efficiency in sheep, and allow phenotypic and genetic analyses. Low to moderate heritability was shown for energy balance, FCR, DMI, and RFI. Furthermore it should be possible to use the data from these models for QTL mapping.

REFERENCES

AFRC (1993) ″Energy and protein requirements of ruminants. An advisory manual prepared by the technical committee on response to nutrients″. CAB international, Wallingford, UK.

Brosh A. (2007) *J. Anim. Sci*.**85**:1213

Cammack K.M., Leymaster K.A., Jenkins T.G. and Nielsen M.K. (2005) *J. Anim. Sci*. **83**:777

Cannas A. (2004) *J. Anim. Sci*. **82**:149

CSIRO (2007) ″Nutrient requirements of domesticated ruminants″. CSIRO Publishing. Collingwood, Australia.

Raadsma H.W., Thomson P.C., Zenger K.R, Cavanagh C., Lam M.K., Jonas E., Jones M., Attard G., Palmer D. and Nicholas F.W. (2009) accepted in *Genet. Sel. Evol.*

Sainz R.D., Paulino P.V. (2004) eScholarship Repository, University of California. http://repositories.cdlib.org/anrrec/sfrec/2004 residual feed intake.

Snowder G.D. and Van Vleck L.D. (2003) *J. Anim. Sci*. **81**: 2704

Van Knegsel A.T.M., Van Den Brand, H., Dijkstra J., Tamminga S. and Kemp B. (2005) *Reprod. Nutr. Dev*. **45**:665

PREDICTING ENERGY BALANCE IN LACTATING EWES AS A BASIS FOR QTL ANALYSIS

E. Jonas, P.C. Thomson, K.J. Fullard, D. McGill and H.W. Raadsma

ReproGen, Faculty of Veterinary Science, University of Sydney, Camden, NSW 2570

SUMMARY

Four different models (AFRC, CSIRO, CNCPS-s and body reserve change) were used to validate predicted energy balance, Residual Feed Intake (RFI) and Feed Conversion Ratio (FCR) of dairy ewes in a computerised self feeding system. Estimates derived from all models were in strong agreement and allowed an estimate of energy balance during lactation. Low to moderate positive phenotypic correlations (range 0.17 to 0.67) were observed between energy balance with milk yield, protein, and lactose content, moderate negative correlations (range -0.47 to -0.48) with energy balance and fat content in the milk. The repeatability of the energy models (AFRF, CSIRO, CNCPS-s) was between 0.16 and 0.46. The system was thought to be sufficiently accurate to predict energy balance, RFI, FCR in individual ewes as input traits for genetic parameter and QTL studies. A preliminary linkage analysis among chromosome 1, 5 and 23 could identify significant linkage regions for energy balance on OAR 1 around 330 cM, which was in a comparative region as described by Sherman *et al*. (2009) in cattle and for RFI at 117 cM on the same chromosome.

INTRODUCTION

Energy balance is a complex trait with relevance to the study of human obesity, and maintenance energy requirements of livestock (Moody *et al*. 1999). Achieving a better understanding of the factors predisposing for differential mobilization of body reserves is an important step in the development of strategies to reduce health and associated reproduction problems in dairy cattle (Friggens *et al*. 2007). It was concluded in different studies that an increased mobilization of body reserves in early lactation is associated with increased health problems and a reduction in reproductive performance in dairy cattle (Hansen 2000; Pryce *et al*. 2001) termed Negative Energy Balance (NEB) syndrome. On the other hand it was proposed that the cow is adapted to energy mobilization in early lactation if this is genetically driven and therefore this does not affect the health and reproduction as environmentally driven mobilization does (Friggens *et al*. 2007). A better understanding of the complexities of energy mobilisation requires the recording of energy balance as a routine phenotype to further estimate genetic correlations between energy balance, milk production, health and fertility and as a target trait for genetic marker studies. Until now only a few studies have identified quantitative trait loci (QTL) for component characteristics of energy balance and feed intake in cattle, with evidence for a genetic component for both (Harder *et al*. 2006, Nkrumah *et al*. 2007, Sherman *et al*. 2009). The aim of this study was to predict energy balance in lactating ewes, to test the repeatability of the parameters and to perform a linkage analysis for detection of QTL on three candidate chromosomes.

MATERIALS AND METHODS

The data used in this analysis were collected at the University Sydney research farm 'Mayfarm' at Camden, New South Wales, Australia as a subset of studies carried out between 2005 and 2009 using dairy sheep. Animals used for the experiment derived from an Awassi × Merino resource population (Raadsma *et al*. 2009). Information was collected on body weight and feed intake using an automatic feeder system without restriction of food on offer. The feeder experiments were performed across three experimental cohorts using 349 ewes between 2005 and

2007. In the present study observations from AMM backcross (BC) and AM_AMM (DBC) double backcross ewes were used. Data were collected before lambing (week -10 until 0), included the lambing date (week 0) and until a maximum of week 26.

Feed conversion (FCR) was calculated as the amount of feed consumed (in kg) per kg body weight gain or loss. Feed intake (in kg) as a ratio of absolute body weight and the residual feed intake (RFI) as the relation between observed and expected feed intake were also calculated (Koch *et al.* 1963). Energy balance was calculated based on the Input-Output models described by CSIRO (CSIRO 2007), the adapted formula of the CNCPS-s model (Cannas *et al.* 2004) and the advisory manual prepared by the AFRC (AFRC 1993). Using these three models, the difference in energy between energy intake (feed) and energy loss (maintenance plus pregnancy or milk energy) was calculated, the resulting variables are further defined as Energy Balance (EB). The three models differ slightly in the calculation of energy required for maintenance. Additionally the change in body reserve was calculated based on body condition score and body weight (Friggens *et al*. 2007), a formula which incorporates constants (universal constants and literature-derived estimates) to describe energy mobilisation (Sanson *et al*. 1993, Tolkamp *et al*. 2007, Lewis and Emmans 2007).

The repeatability of the parameters was calculated on a weekly basis using animal as a random factor, and season, experiment, reproduction stage, sire and genotype as fixed effect terms. A linkage analysis was performed in 100 backcross ewes of one half sib family using 26 markers on chromosome 1, 5 and 23. QTL methodology (maximum likelihood method) and marker design are explained in detail in Raadsma *et al.* (2009). Phenotypes used were the mean values of EB from the three models, body reserve change and RFI. QTL were calculated for phenotypes within five phases: phase I pregnancy, phase II week 1 & 2, phase III week 3 to 10, and phase IV week 11 to 20 of lactation.

RESULTS AND DISCUSSION

Table 1. Results of residual feed intake, energy difference between input and output (CSIRO, AFRC and CNCPS-s) [MJ], body reserve change [MJ]; number of animals (N), average (mean), standard deviation (std), minimum (min), maximum (max) for one experiment

Trait	N	mean	std	min	max
Feed conversion (feed intake / growth)	62	13.84	4.28	6.12	21.98
Feed intake / body weight*100	62	1.93	0.56	0.82	2.99
Energy balance (CSIRO)	58	5.92	2.51	1.15	11.39
Energy balance (CNCPS-s)	58	6.14	2.68	1.38	11.03
Energy balance (AFRC)	58	4.93	2.66	0.53	9.88
Body reserve change	58	3.97	2.93	-1.63	15.15
Residual feed intake	62	1.46	0.32	0.26	2.15

The weekly feed intake, body weight and lactation performances of 62 ewes from the Awassi × Merino population were analysed in a first cohort. Predicted dry matter intake, RFI and FCR indices were derived from the summarized feeder data. Using the estimates of RFI, slight differences between the estimated and measured feed intake were observed. The EB models of CSIRO, CNCPS-s and AFRC resulted in very similar parameters (Table 1) and an average was of all three was taken for further analysis of QTL effects. High (0.73 to 0.79) correlations were also observed between EB (CSIRO, CNCPS-s, AFRC) and RFI. Correlations between energy balance and lactation performance are shown in Table 2.

Table 2. Phenotypic correlation between the estimates of the energy balance, feed conversion rate and residual feed intake with the milk yield and composition traits

Trait	Energy balance			Body reserve change	FCR	RFI
	AFRC	CNCPS-s	CSIRO			
Milk yield	0.17	0.14	0.17	0.19	-0.40	0.65
Protein content	0.53	0.52	0.53	-0.09	-0.36	0.55
Fat content	-0.48	-0.47	-0.48	-0.34	0.31	-0.61
Lactose content	0.66	0.67	0.66	-0.12	-0.21	0.21
Repeatability	0.34	0.32	0.16	ne	0.004	0.46

Body reserve change was negatively correlated to fat content. FCR and RFI were correlated with milk yield. EB was highly correlated (>0.93) using the three models (CSIRO, CNCPS-s, AFRC), but lowly correlated to Body reserve change (0.19 to 0.2).The repeatability estimates of parameters describing energy balance parameters were comparable (0.16 [CSIRO] to 0.3 [AFRC]) within the three approaches (CSIRO, CNCPS, AFRC). The repeatability of the RFI was high (0.46), moderate for feed intake/body weight (0.26) and low for feed conversion (0.004).

Table 3. Results of the QTL analysis for energy balance calculated under three models and RFI

Trait	Phase	OAR	Position	LOD	sign	Effect
Energy balance (CSIRO)	I to IV		329.5	1.77	*	-3.04
	I	1	335.5	3.81	***	-4.82
	II,III,IV		-	no significant QTL		-
Energy balance (CNCPS-s)	I to IV		331.5	1.49		-2.90
	I	1	336.5	2.72	**	-3.69
	II,III,IV		-	no significant QTL		-
Energy balance (AFRC)	I to IV		328.5	1.42		-2.85
	I	1	333.5	2.85	**	-3.77
	II,III,IV		-	no significant QTL		-
Residual feed intake	I to IV		117.5	2.27	**	0.41
	I	1	118.5	2.2	**	0.32
	II,III,IV		-	no significant QTL	-	
Energy balance and Residual Feed Intake	all	5	-	no significant QTL	-	
		23	-	no significant QTL	-	

* suggestive QTL (LOD>1.7); ** significant QTL (LOD>2.0); highly significant (LOD>3.0)

The linkage analysis using QTL-MLE showed one significant QTL for RFI on OAR 1, which was located in a comparative region to BTA 1 (Table 3). The QTL for all three energy balance models were located in the comparative region to BTA 3, where Sherman *et al*. (2009) had describe a QTL for RFI in cattle (Table 3). Additional studies will need to be done to predict the EB for total milk take-off based on milk yield and milk composition data as previously described in cattle (Friggens *et al*. 2007). The results showed that the data derived from the different feed intake models can be further used to study the genetic background of these traits.

CONCLUSION

We could show that the application of input-output models to estimate the energy balances in ewes before and during lactation using a computerised feeding system yielded very similar estimates of energy balance. The traits were moderately repeatable and amenable for further genetic analysis. A targeted linkage analysis revealed one significant QTL for RFI on OAR 1 and also significant linkage for energy balance on the same chromosome, which were located in the comparative region to the identified QTL for RFI in cattle.

REFERENCES

AFRC (1993) ″Energy and protein requirements of ruminants. An advisory manual prepared by the technical committee on response to nutrients″. CAB international, Wallingford, UK.

Cannas A. Tedeschi L.O., Fox D.G., Pell A.N. and Van Soest P.J. (2004) *J. Anim. Sci*. **82**:149

CSIRO (2007) "Nutrient requirements of domesticated ruminants" CSIRO Publishing. Collingwood, Australia.

Friggens, N.C., Ridder, C. and Lovendahl, P. (2007) *J. Dairy Sci.* **90**:5453

Hansen, L.B. (2000) *J. Dairy Sci*. **83**:1145

Harder, B., Bennewitz, J., Reinsch, N., Thaller, G., Thomsen, H., Kuhn, C., Schwerin, M., Erhardt, G., Forster, M., Reinhardt, F. and Kalm, E. (2006) *J. Anim. Breed. Genet*. **123**:89

Koch, R.M., Swiger, L.A., Chambers, D. and Gregory, K.E. (1963) *J. Anim. Sci*. **22**:486

Lewis, R.M. and Emmans, G.C. (2007) *Animal* **1**:1427

Moody, D.E. Pomp, D., Nielsen, M.K. and Van Vleck, L.D. (1999) *Genetics* **152**:699

Nkrumah, J.D., Sherman, E.L., Li, C., Marques, E., Crews, D.H. Jr., Bartusiak, R., Murdoch, B., Wang, Z., Basarab, J.A. and S.S. Moore (2007) *J. Anim. Sci.* **85**:3170

Pryce, J.E., Coffey, M.P. and Simm, G. (2001) *J. Dairy Sci.* **84**:1508

Raadsma, H.W., Thomson, P.C., Zenger, K.R., Lam, M.K., Jonas, E., Cavanagh, C., Jones, M., Attard, G., Palmer, D. and Nicholas, F.W. (2009) accepted in *Genet. Sel. Evol.*

Sanson, D.W., West, T.R., Tatman, W.R., Riley, M.L., Judkins, M.B. and Moss, G.E. (1993) *J. Anim. Sci*. **71**:1112

Sherman, E.L., Nkrumah, J.D., Li, C., Bartusiak, R., Murdoch, B. and Moore, S.S. (2009) *J. Anim. Sci.* **87**:37

Tolkamp, B.J., Yearsley, J.M., Gordon, I.J., Illius, A.W., Speakman, J.R. and Kyriazakis, I. (2007) *Br. J. Nutr*. **97**:1206

A COMPARISON BETWEEN SHEEP BRED FOR WORM RESISTANCE AND UNSELECTED CONTROLS WHEN EXPOSED TO LOW LARVAL CHALLENGE DURING SUMMER

K. E. Kemper[1], J.W.A. Larsen[2], S.C. Bishop[3], N. Anderson[2], M.E. Goddard[1], J.C. Greeff[4], R. Woodgate[4] and L.J.E. Karlsson[4]

[1]Melbourne School of Land and Environment, University of Melbourne; [2]The Mackinnon Project, Faculty of Veterinary Science, University of Melbourne; [3]The Roslin Institute & Royal (Dick) School of Veterinary Studies, University of Edinburgh; [4] Department of Agriculture and Food Western Australia

SUMMARY

Sheep in winter-rainfall regions typically show a rise in faecal worm egg count (WEC) over the summer-autumn period, when pasture larval challenge is low or near zero. This rise is important epidemiologically for worm control. Our study found that the peak WEC of sheep bred for worm-resistance was reduced by 43% compared to unselected control sheep during the summer-autumn period. Total worm counts confirm that the rise in WEC was not due to an increased worm burden. Explanations for the summer-autumn rise in WEC include an increased fecundity of existing adult worms or maturation of existing immature worms. However, at this stage the precise mechanisms behind the rise in WEC remain unclear. A marked increase in the WEC for both genotypes of sheep occurred when they were removed from pasture and housed for 3-4 weeks.

INTRODUCTION

Reducing the summer-autumn rise in WEC may be important to capture epidemiological benefits from sheep bred for low WEC in winter-rainfall environments. This is because eggs deposited during this period largely determine the peak larval challenge during winter (Anderson 1972), and hence the degree of parasitism and production losses (Niven *et al.* 2002). The summer-autumn period generally has low larval challenge and there is concern that sheep bred for low WEC, under high larval challenge conditions, may not suppress WEC during the summer-autumn period. This concern is based on the hypothesis that immunity is dependent on a threshold of ingested larval (Dineen 1963). The aim of this study was to investigate two hypotheses; (i) that sheep bred for low WEC (worm-resistant, R) will maintain lower WEC than unselected control (C) sheep throughout the summer-autumn period, and (ii) that the summer-autumn rise in WEC is caused by maturation of arrested *Teledorsagia* spp. rather than increased total worm burden. Hypothesis 2 is primarily concerned with the biological basis for the summer-autumn rise in WEC, rather than a comparison of the genotypes contained within this study.

MATERIALS AND METHODS

Animals. Animals were sourced from three epidemiologically independent replicates of Rylington Merino worm-resistant (R1, R2, R3) and unselected control (C1, C2, C3) mature ewes (n = 50/replicate) at the Mt Barker research station, Western Australia (Greeff & Karlsson 2006; Greeff *et al.* 2006). To test the first hypothesis we monitored a subset of 3-4 year old ewes from each replicate for WEC at four times during the 2007/2008 season (18th September, 11th December, 20th February, 28th March). Individuals were repeatedly measured, where possible, and a total of 473 WEC measures were made on 164 animals. Most animals (69%) had 3–4 WEC measures. Lactation status was recorded in early spring (September). The second hypothesis was tested by the necropsy of two ewes per replicate (2–7 years of age), before and after the summer-autumn rise in WEC. Animals were chosen from 6–9 cull animals per replicate. Animals for necropsy

were selected from culls based on WEC measurements at the above dates, either to reflect the observed WEC range (early summer) or on the basis of a demonstrated rise in WEC (late summer). After selection, animals were removed from their pasture contemporaries (11^{th} December or 25^{th} – 26^{th} February) and housed for 3–4 weeks on a mixed chaff and lupin ration. Housing prevented the ingestion of larvae so that worms counted in the early L_4 stage were truly arrested and were not confounded with recently ingested larvae. Total worm counts were performed over 7 (2^{nd} – 9^{th} January) or 2 days (18^{th} – 19^{th} March) and measurements included the count of adult worms in each genus, the count of immature L_5, delayed and early L_4, and WEC. Total worm burden was the sum of all counted worms.

Parasitology. WEC and necropsy measurements were made according to standard procedures at the Agriculture Western Australia laboratories, Albany. WEC was assessed using a modified McMaster method with 25 eggs per gram (epg) sensitivity. Larval cultures were conducted for each plot at each WEC measurement, or individually for necropsy animals; and a proportion of *H. contortus* was detected in early spring (C1, 0.28; R2, 0.12 & R3, 0.01), early summer (C2, 0.04; C3, 0.13) and late summer (C2, 0.22; R3, 0.70). To prevent potential bias, all WEC were adjusted for either the replicate or individual proportion of *H. contortus*. Necropsy procedures included the examination of the washings and digests from the abomasum and small intestine.

Statistical Analysis. WECs showed a skewed distribution and were transformed [$\log_{10}(x+10)$] for normality. Replicate geometric means and t-tests for each time-point are presented. To account for multiple measures over time a mixed, repeated measures linear model was fitted to log-transformed data. The model had the form:

$$\log_{10}(WEC_i+10) = \mu + G_j + T_k + G.T_{j,k} + L_l + R_{j,k,m} + ID_n + \varepsilon_i \quad (1)$$

where the fitted effects were the overall mean (μ) and fixed effects of genotype (G_j, j = R, C), time of measurement (T_k, k = 1…4), the interaction between genotype and time ($G.T_{j,k}$) and lactation status (L_l, l = wet, dry; k = 1 only); and random terms were replicate variance [replicate (m) x genotype x time interaction, $R_{j,k,m}$ where m = 1..3], between-animal variance (ID_n, for n^{th} animal) and within-animal variance (or residual error, ε_i). The linear model for necropsy data had the form:

$$\log_{10}(x_i+1) = \mu + G_j + T_k + G.T_{j,k} + \varepsilon_i \quad (2)$$

where i^{th} worm count was fitted to the overall mean (μ) and fixed effects of genotype (G_j), necropsy group (T_k, k = 1,2) and their interaction ($G.T_{j,k}$). Replicate structure could not be included due to small sample size per replicate. Necropsy WEC was $\log_{10}(x_i+10)$ transformed.

RESULTS AND DISCUSSION

Faecal worm egg count. WECs were low throughout the 07/08 season. However, temporal trends in mean WEC show a clear summer-autumn rise for both C and R genotype sheep (Table 1) and peak mean WEC was 43% lower in R animals. Approximately 50% of C and 30% of R genotype sheep had WEC > 200 epg in late summer (February) and early autumn (March). Unexpectedly, WECs for the necropsy animals were much higher than their genotype contemporaries on pasture at the time of slaughter. WECs increased markedly during 3–4 weeks of housed conditions. This rise makes inferences about the structure of parasite populations in pasture contemporaries impossible. The higher late summer WEC for R ewes in the late summer necropsy group was expected, as this group was chosen on the basis of a demonstrated rise in WEC (rather than to be representative of the genotype).

Geometric means show the variation between genotype replicates, as well as the high variation within each replicate (Table 2). The mixed model showed R animals to have significantly lower WEC than C animals when all measures are considered together (Table 3, $P < 0.05$), however the

time-point specific differences were never significant. The repeatability of WEC over the period was high (0.42), providing an upper-bound estimate of heritability for WEC over summer. Significant replicate and between-animal variance components show that (i) the variation between replicates was not simply due to the variation between animals and (ii) the variation between animals was not simply due to the variation between replicates ($P < 0.01$).

Table 1. Arithmetic mean faecal worm egg count (WEC; eggs per gram) for control (C) and resistant (R) genotype sheep from September 07 – March 08. Shown is mean for animals continuously grazed on pasture (mean of replicates) and the mean for the two necropsy groups removed from pasture in early (ES) or late (LS) summer. The mean for necropsy groups after removal from pasture is shown in italics

Animal group & genotype		E. Spring (September)	E. Summer (December)	*E. Summer necropsy*	L. Summer (February)	*L. Summer necropsy*	E. Autumn (March)
mean of replicates	C	142	56		346		324
"	R	53	31		198		206
ES necropsy gp	C	77	94	*2350*			
"	R	79	124	*845*			
LS necropsy gp	C	64	24		343	*1550*	
"	R	14	25		445	*1392*	

Table 2. Geometric mean faecal worm egg count (eggs per gram) for replicates 1-3 in control (C) and resistant (R) genotype sheep. Genotype geometric means are shown in bold. The range observed in each replicate is shown in parentheses

Replicate	Early spring (September)		Early summer (December)		Late summer (February)		Early autumn (March)	
C1	40	(0 – 485)	35	(0 – 300)	160	(0 – 950)	120	(0 – 975)
C2	65	(0 – 1875)	40	(0 – 360)	140	(0 – 1345)	140	(0 – 950)
C3	75	(0 – 600)	30	(0 – 740)	400	(75 – 1550)	400	(25 – 1325)
	60		**35**		**210**		**200**	
R1	60	(0 – 225)	35	(0 – 200)	140	(25 – 1600)	170	(25 – 1425)
R2	20	(0 – 25)	30	(0 – 150)	85	(0 – 475)	65	(0 – 180)
R3	50	(0 – 265)	30	(0 – 365)	130	(0 – 765)	210	(0 – 1050)
	40		**30**		**120**		**130**	
P- value^		*0.18*		*0.09*		*0.09*		*0.20*

^ one-sided t-test

Table 3. F-ratios and variance components for the mixed model of WEC

Effect	Df	F-ratio	Component	df	Variance
Lactation	1	10.1[a]	Replicate	24	2.07×10^{-2} [b]
Genotype	1	5.35 [a]	Between-animal	164	1.15×10^{-1} [b]
Time	1	37.9 [b]	Within-animal	463	1.60×10^{-1}
Genotype x Time	3	0.36			

[a] $P < 0.05$, [b] $P < 0.01$

Necropsy observations. Mean worm burdens were low and typical of what might be expected during summer (Table 4). A large range of worm burdens was observed, particularly in the early summer group. R animals tended to have lower total worm burden and WEC at necropsy. Fitted models generally had difficulty to predict the observed geometric means (not shown), indicating poor fit to the data, potentially due to the small sample sizes and high variability. WECs at necropsy were much higher than those observed on pasture, implying a change in the resident worm population under housed conditions or changes to faecal composition. Higher WEC may

have occurred due to increased egg production per worm or maturation of immature worms. Changes in WEC and a poor data fit by the models does not allow for valid testing of the second hypothesis. For future studies we recommend a shorter interval period of housing prior to necropsy (4 – 7 days) and the selection of animals for necropsy to be more representative of pasture contemporaries. These data suggest maturation of immature stages over summer, however increased worm fecundity cannot be excluded and these results would need to be reproduced. Worm burdens were stable over summer, potentially indicating low larval challenge and that an increased worm burden was not the cause of the summer rise in WEC for pasture contemporaries.

Table 4. Geometric mean worm counts from control (C) and resistant (R) genotype animals in early and late summer necropsy groups. The range of observed values is shown in parentheses.

	n	Total	*Teledorsagia*‡	*Trichostrongylus*‡	immature†,§	early L_4§	WEC
Early summer necropsy							
C	6	2950	70	120	80	1250	1750 α
		(450 – 11150)	(0 – 2450)	(0 – 6250)	(0 – 1100)	(150 – 3250)	(650 – 4650)
R	6	1950	470	90	130	180	610 β
		(400 – 5950)	(200 – 1300)	(0 – 1250)	(50 – 350)	(0 – 5200)	(100 – 1230)
Late summer necropsy							
C	6	3750	1120	1430	230	420	1250
		(1500 – 9100)	(350 – 3450)	(600 – 6750)	(50 – 1500)	(50 – 2000)	(600 – 3800)
R	6	1950	780	640	20	210	1070
		(600 – 3850)	(250 – 2450)	(250 – 2550)	(0 – 150)	(50 – 550)	(250 – 2800)
G x T		*Ns*	*ns*	*ns*	*0.06*	*ns*	*ns*
G.type		*0.19*	*0.33*	*0.61*	-	*0.17*	*0.25*
Time		*0.78*	*0.05*	*0.04*	-	*0.63*	*0.72*

α, β missing data; n = 3, 4

† contains immature L_5 and developing L_4; § all species; ‡ mostly adults, but some immatures were detected

CONCLUSIONS

Sheep selected for low WEC had a reduced mean WEC during the summer-autumn period compared to unselected sheep. This may partially explain reduced winter pasture contamination and epidemiologically-derived production benefits observed for R sheep in this environment (Greeff *et al.* 2006; Williams *et al.* 2006). The 43% reduction in peak WEC was less than reported under high larval challenge (89%, Greeff & Karlsson 2006). It seems sheep selected for low WEC may respond to larval challenge more readily than unselected controls. We could not test our second hypothesis as necropsy WECs were much higher than those on pasture. The increased WEC during housing maybe due to changes to faecal composition or indicate altered immunity; potentially due to stress, an absence of larval challenge or nutritional changes.

ACKNOWLEDGMENTS

We thank AWI, Department of Agriculture and Food Western Australia and BBSRC for financial support. Technical assistance from Jill Lyon is gratefully acknowledged.

REFERENCES

Anderson, N. (1972) *Aust. J. Ag. Res*. 23:1113.
Dineen, J.K. (1963) *Nature*, **197**:268.
Greeff, J.C. and Karlsson, L.J.E (2006) *Int. J. Sheep and Wool Sci.* **54**:8.
Greeff, J.C., Karlsson, L.J.E. and Underwood, N. (2006) *Proc. 8th WCGALP* **15**:468.
Niven, P. Anderson, N. and Vizard, A. (2002) *Aust. Vet. J.* **80**:559.
Williams, A.R., Greeff, J.C., Vercoe, P.E., and Karlsson, L.J.E. (2006) *Proc. Aust. Soc. Anim Prod*. **26**: SC42.

BREEDER PERSPECTIVES ON FAT AND FEMALE MANAGEMENT

S. J. Lee, I. K. Nuberg, W. S. Pitchford

Cooperative Research Centre for Beef Genetic Technologies
School of Agriculture, Food and Wine, The University of Adelaide, Roseworthy SA 5371

SUMMARY

Seedstock breeder's perspectives on topics associated with maternal productivity were investigated through the use of qualitative in-depth semi-structured interviews. Qualitative content analysis of interview data yielded a considerable divergence in attitudes to female management with regards grazing management, body condition score (BCS) fluctuation and the utilisation of body fat reserves. Variation in these attitudes were associated with divergence in perspectives on the importance of selecting for positive (> 0) rib and rump fat estimated breeding values (EBVs) for female fertility, or selecting for neutral or negative (≤ 0) fat EBVs to assist yield, particularly with regards commercial steer progeny. These results demonstrate that amongst seedstock breeders targeting similar end markets, substantial variation in animal selection and management exists that requires further characterisation to ensure breeding programs and animal management are optimal.

INTRODUCTION

Today's modern production animals have been selected on output traits and subsequently fitness traits and adaptability have generally declined, potentially leading to reduced overall productivity (Hohenboken *et al.* 2005). In response, the requirement of cattle breeders has shifted from simply selecting for increased output towards selecting for optimum conversion of available food resources to marketable product across the production system (Jenkins and Ferrell 2007). In forage based systems the quality and quantity of dry matter available varies greatly and periods of low feed availability may inhibit individual animal's ability to express their genetic potential for production traits. Jenkins and Ferrell (1994) effectively demonstrated the need for genetic potential to be aligned with production environment. If there is a high genetic potential but poor production environment (low or variable nutrition), it is likely biological production efficiency will be severely reduced due to factors such as reduced body condition and extended calving interval (Richards *et al.* 1986).

Body fatness at calving is recognised as a major factor determining duration of postpartum anoestrus and subsequent conception rates. BCS is a subjective measurement of the level of fat and muscle reserves that an animal displays and has been reported to be highly correlated to different measures of body fat. Wagner *et al.* (1988) estimated that BCS accounted for 80% of the phenotypic variation of carcass fat in 400kg beef cows. For dairy cattle BCS is heritable (0.17) and genetically correlated to fertility traits and thus it is possible to indirectly select for fertility traits which are lowly heritable by selecting for BCS (De Haas *et al.* 2007). In beef cattle, Johnston *et al.* (2003) estimated heritabilities of 0.16 for BCS and 0.38 for P8 fat depth, potentially indicating response to selection for body fatness may be better achieved through P8 rather than BCS selection.

Beef producers recognise the importance of BCS at calving for subsequent reproduction. However, there is considerable divergence in opinion about how to achieve optimal BCS at calving for subsequent rebreeding whilst producing beef in the most biologically efficient and economically profitable way. This paper aims to illustrate the divergence in breeder perspectives on the management of females and also on the perceived importance of fat EBVs in relation to female fertility, BCS and carcass yield.

MATERIALS AND METHODS

A qualitative research approach comprising 24 in-depth semi-structured interviews with Angus and Hereford seedstock breeders was conducted. The interviews focused on factors breeders perceived as important in contributing to maternal productivity in beef cattle. Sampling was purposeful with all breeders interviewed conducting the full range of Breedplan performance recording on both sexes. The interview schedule combined a structured approach to allow for systematic comparison of interviewees responses to a particular topic with unstructured spontaneous questioning to facilitate further understanding on topics of interest. The interview schedule was formed from existing research knowledge in beef production, before being internally examined by peers and formally assessed in a pilot study with 3 breeders.

Audio from the interviews was digitally recorded and transcribed in full. The transcripts were analysed using a content analysis approach as described by Miles and Huberman (1994) with the assistance of NVivo™ 8 software providing a database for the arrangement, retrieval and verification of data. The actual coding and arranging of data was conducted by the researcher. Content analysis is multi-staged and involves transcript familiarisation, data coding, and within and across interview analysis. Multiple readings of each transcript were undertaken prior to coding to understand fragments of text in context. Coding of data involved both data reduction; by arranging large texts into small analytic units (codes) which acted as labels for assigning meaning to descriptive information compiled from interviews; and analytical categorisation of the data (Neuman 2003). An initial coding structure was formed from prior research knowledge and complemented with concepts emerging from initial readings of the transcripts. Further codes were added to the coding structure when new information or relationships in data became apparent.

First level coding of each transcript in its entirety was conducted and involved placing segments of text to the appropriate code. Secondary coding was performed to cluster codes together and examine the relationships and interactions between initial codes and to form an initial conceptual order to indicate the nature of relationships between categories. Secondary coding was used to demonstrate interactions between codes, to show how concepts cluster together and how codes relate to each other under different circumstances.

Data within code for each interview was compared to ensure consistency of message on a particular topic, and identify underlying factors if content was not consistent. Data within code and between interviewees were then compared to determine if common themes or messages, or contrasting information was emerging for a particular code. Cross case analysis involved comparing comments on the theme, ‘fat EBVs and fertility’ across interviews and provided an insight into how applicable the findings of the project would be across multiple sites or similar circumstances (Miles and Huberman 1994).

RESULTS AND DISCUSSION

The outcomes of analysis of interview data demonstrated substantial variation in breeder attitudes towards expectations of the female that appeared independent of breed, calving season, average genetic potential and production environment (location). The vast majority of breeders interviewed had firm requirements about the need for females to calve at 24 months and subsequently calve on an annual cycle. However, considerable divergence existed in their attitude to female management with regards grazing management, provision of supplementary feed, BCS fluctuation and the utilisation of body reserves. Variation in these attitudes was associated with differing opinions on the importance of selecting for positive fat EBVs for female fertility, or selecting for neutral or negative fat EBVs to assist yield, particularly with regards commercial steer progeny. Figure 1 depicts the variation in attitudes held by many interviewees towards female management, associated attitude to fat EBVs, fertility and yield and subsequent perceived outcomes for the breeding female, commercial progeny, production efficiency and profitability.

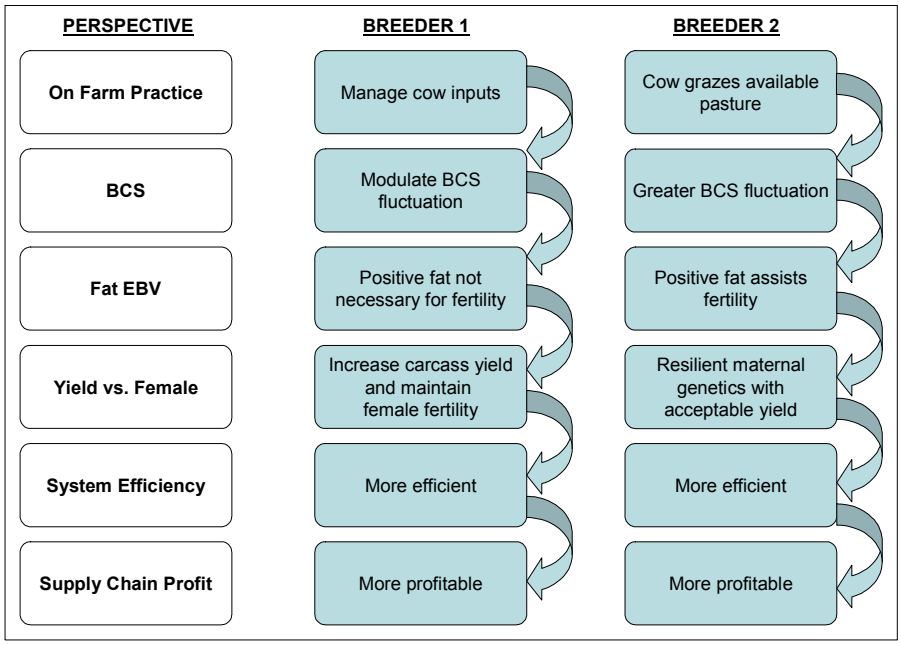


Figure 1. Contrasting breeder perspectives on animal management, selection and profit.

Breeder quotes outlining their perspective on animal management, attitude to fat and system efficiency demonstrate a considerable contrast. Quotes below are from breeders who believe that positive fat is not necessary for fertility (quote 1) and that large fluctuations in body condition are poor utilisation of available energy (quote 2) and that genetic gain and phenotypic production for fertility were better achieved through advancement in the days to calving (DTC) EBV (quote 3).

'If you couldn't do it any other way and we were dropping 10% conception rate, then you have to have that fat there. If you can have yield and have efficient cows, then to me the wrong thing for a seedstock business to do is to push positive fat.'

'I have a real problem with fat and I touched on this before in regard to cow condition…When you feed it (fat) *on and it costs you somewhere in the vicinity of 40% of every mega-joule you end up using in the process of putting it on and then re-metabolising it again, that's just crazy'*

'People that pick fat cows for fertility, they should concentrate on fertility, don't measure a correlated trait, measure the trait you want…we use our DTC EBV to make sure they are still fertile and staying in calf.'

The following quotes are from breeders who believe that positive fat is important for fertility and selection on the DTC EBV for fertility is not sufficient. Breeders were aware of the importance of carcass yield but believe that they cannot afford to select for negative fat due to the impact on female fertility (quote 1). Breeders believed managing cows to gain body condition during times of high feed availability and mobilise body condition when feed is limited is desirable and that positive fat cattle had greater ability to successfully fluctuate in BCS (quotes 2 and 3).

'I use the carcass traits more for maternal productivity than for carcass. It is huge, rib and rump, whenever I use an AI sire I have a really good look at his fat EBVs. When I use a negative fat bull the drop out rate in our maternal side is too high…I know the feedlot buyer wants high yielding steers but I can't afford that in the cow herd.'

'The ones that have stayed in the system have this ability where they can put weight on quickly in spring when there is compensatory growth and they can draw down on those reserves when things are tougher like now (March-April) *and then gain weight quickly when the feed is available. We see a real advantage in that.' 'They are positive fat cattle and pretty fertile cattle.'*

The quotes demonstrate a clear divergence on breeders perspectives on the optimal female, breeding and animal management strategy to maximise production efficiency and supply chain profitability. This result was somewhat surprising given that all breeders interviewed were aiming for high quality beef markets. The finding demonstrates that even though the breeders had similar end market goals, there were substantial differences in selection criteria and the way they interpreted and utilised fat EBVs. Had one group of breeders focused on a terminal end market and the other on a self replacing herd the divergence would have been less surprising.

Fluctuations in body weight and condition associated with synthesis and catabolism of protein and fat consumes substantial amounts of energy and is generally an inefficient use of metabolisable energy (ME). Ferrell *et al.* (1976) estimated the maternal energy gain in non pregnant heifers was 39% whilst Freetly *et al.* (2008) estimated the overall efficiency of maternal energy gain was 41%. However, Freetly *et al.* (2008) found when feeding pregnant females on a rising plane of nutrition following a period of low nutrition the efficiency of converting ME to retained energy ranged from 84-98%. The results of Freetly *et al.* (2008) demonstrate that the energy expenditure associated with changing body composition is negligible relative to total feed intake. This led the authors to suggest that, contrary to prior assumptions that allowing cows to lose and gain energy was an inefficient use of energy, it may indeed provide an efficient use of energy. By implementing feeding strategies focused on optimising body weight fluctuation to best use available feed resources it may be possible to increase the efficiency with which body tissue is synthesised and mobilised. This strategy would allow cattle breeders to determine how and when feed resources are utilised, namely females gaining body condition during times of high feed availability and mobilising stored energy during times of low feed availability. Further research on cow body condition is required to understand the effect fat genotype has on the efficiency of energy gain, storage and mobilisation, the nature of efficiency of body condition gain and the efficiency of accumulating and mobilising tissue reserves without impacting performance.

ACKNOWLEDGEMENTS

The support and enthusiasm of all interviewees in this research is gratefully acknowledged. Their insights and contribution to the research are invaluable.

REFERENCES

De Haas, Y., Janss, L. L. G. and Kadarmideen, H.N. (2007) *J. Anim. Br. And Gen.* **124**:277.

Ferrell, C. L., Garrett, W. N., Hinman, N. and Grichting, G. (1976) *J. Anim. Sci.* **42**:937.

Freetly, H.C., Nienaber, J.A. and Brown-Brandl, T. (2008) *J. Anim. Sci.* **86**:370.

Hohenboken, W., Jenkins, T., Pollak, J., Bullock, D. and Radakovich, S. (2005) *Proc. Beef Impr. Fed. 37th Annual Research Synposium* **37**:115.

Jenkins, T. G. and Ferrell C. L. (1994) *J. Anim Sci.* **72**:2787.

Jenkins, T. G. and Ferrell C. L. (2007) *J. Anim Sci.* **85**:1787.

Johnston, D. J., Reverter, A., Burrow, H.M., Oddy, V.H., and Robinson, D. L. (2003) *Aust. J. Agric. Res*. **54**:107.

Miles, M.B. & Huberman, A.M. (1994) "Qualitative Data Analysis" 2nd ed. Sage Publications, US.

Neuman, W.L. (2003) "Social Research Methods" 5th ed. Pearson Education, US.

Richards, M. W., Spitzer, J.C. and M. B. Warner (1986) *J. Anim Sci.* **62**:300.

Wagner, J.J., Lusby, K. S., Oltjen, J. W., Rakestraw, J., Wettemann, R. P. and Walters, L. E. (1988) *J. Anim. Sci.* **66**:603.

SOME CONSEQUENCES OF SELECTION FOR RESIDUAL FEED INTAKE IN BEEF CATTLE

D.S. Lines[1], M.L. Wolcott[2], W.S. Pitchford[1], C.D.K. Bottema[1], R.M. Herd[3] and V.H. Oddy[2]

Cooperative Research Centre for Beef Genetic Technologies

[1]Animal Science, The University of Adelaide, Roseworthy, SA, 5371.
[2]Division of Animal Science, University of New England, Armidale, NSW, 2351
[3]NSW DPI Beef Industry Centre, Armidale, NSW, 2351

SUMMARY

Body composition and energy expenditure were investigated in Angus heifers divergently selected for residual feed intake. Differences in fat deposition at rib and rump sites were observed between the lines but there was no difference in protein deposition, weight gain or energy expenditure. Most of the variation in energy expenditure could be accounted for by the metabolisable energy consumed by the animal. The implications of this observation on the biological consequences of selection for residual feed intake are discussed. These are preliminary observations and further work on the biological basis of the trait is required to provide definitive answers.

INTRODUCTION

In typical beef production systems, the cost of feed accounts for over half of the total cost of production. Accordingly, improvements in the efficiency of feed utilisation are a desired management objective. There is phenotypic variation in feed intake independent of variation in average weight and weight gain (termed residual feed intake, RFI) and is moderately heritable (Arthur *et al.* 2001a; Arthur and Herd 2008). Selection of beef cattle for and against RFI measured shortly after weaning has been underway for almost a decade (Arthur *et al.* 2001a; Arthur and Herd 2008). Here we used animals divergently selected for RFI for 3-4 generations to investigate possible biological mechanisms contributing to the trait.

In beef cattle divergently selected for RFI for 1 generation, up to 95% of the variation in RFI was attributed to differences in energy expenditure (as heat production) rather than to energy retained in body tissues (Richardson and Herd 2004). These authors suggested energy expenditure associated with whole body protein turnover, tissue metabolism and stress could contribute 37% of the variation in RFI. The energy costs of protein turnover contribute 20-25% of maintenance energy expenditure and 15-20% of basal metabolic rate across a range of species; protein synthesis alone may contributes up to 30% of heat production in cattle (reviewed by Richardson and Herd 2004). There is particular interest in the effects of selection for reduced RFI on the efficiency of energy use because it was originally hoped that selection would result in reduced maintenance energy and therefore, feed requirements. The aim of this study was to assess the contribution of protein synthesis to energy expenditure in young beef cattle selected for and against RFI. Here we report the data on energy expenditure, and briefly refer to preliminary results of protein synthesis measurements.

MATERIALS AND METHODS

Sixteen (16) Angus heifers from a research population divergently selected for approximately 3-4 generations for high RFI (low "efficiency"; n=8, average mid-point parental RFI EBV=0.64±0.07 kg/d) or low RFI (high "efficiency"; n=8, average mid-point parental RFI EBV=-0.78±0.26 kg/d) were used, where the parental EBV's were calculated by BREEDPLAN. They

were fed diets 105% maintenance requirements (MR) and 95% *ad libitum*, equal to approximately 180% MR, in a crossover design. The crossover design was implemented to reduce individual animal differences from the overall treatment effects, hence increasing the statistical power of the experiment resulting in the use of fewer animals. First; one half (4 high RFI and 4 low RFI) were fed at 180% MR and the other half (4 high RFI and 4 low RFI) fed 105% MR. After adaptation to this feeding level for 21 days, the following measurements were taken: Subcutaneous rump and rib fat depth, eye muscle area (EMA) and intramuscular fat content (IMF) were measured using an ultrasound scanner. Energy expenditure was estimated by infusion with 17.5 μmol/kgBM ^{13}C-sodium bicarbonate similar to the method of Li *et al.* (2008) and protein synthesis by continuous infusion with 10 μmol/kgBM$^{0.75}$/hr ^{13}C-leucine and analysed by the method of Calder and Smith (1988). Dietary treatments were switched and the measurements were taken again so that measurement duration was 35 and 42 days for periods 1 and 2. Each animal was measured twice, once at low and high feed intake. The animals were fed twice daily of a diet of 50% grain, 40% chopped sorghum hay, 9% Molofos® and 1% minerals. The diet was estimated to contain 11.5MJ metabolisable energy and 12.5g of crude protein/kg dry matter. Feed refusals were weighed twice daily. Data were analysed with a general linear model including factors for RFI line, diet, time of measurement, animal and all interactions between factors with initial body weight as a covariate. Tests of significance for main effects were based on type I sums of squares. Significance was defined as $P<0.05$.

RESULTS AND DISCUSSION

No interactions were observed between RFI line and feeding level for the body composition traits. Feeding level had the largest effect on the change in body composition traits during the measurement periods. Heifers fed at 180% MR grew faster and laid down more fat over the rump and ribs, and as IMF than heifers fed 105% maintenance (Table 1). Additionally, the low RFI heifers had lower rump and rib fat deposition ($P<0.05$), but not IMF deposition, than the high RFI heifers, regardless of feeding treatment.

Table 1. Main effects means and SEM for absolute changes over the treatment periods in weight and tissue depots for high and low RFI heifers fed at either 105% or 180% maintenance feeding levels

Main effects	Treatment	ADG (kg/d)	EMA (cm^2)	Rump fat (mm)	Rib fat (mm)	IMF (%)
RFI line	Low	0.66^a ±0.12	4.2^a ±0.61	1.0^a ±0.30	1.0^a ±0.27	0.9^a ±0.16
	High	0.70^a ±0.11	4.8^a ±1.07	1.9^b ±0.52	1.4^b ±0.38	0.9^a ±0.21
Feeding level	180% MR	0.88^a ±0.09	4.8^a ±0.85	2.4^a ±0.33	2.0^a ±0.24	1.3^a ±0.16
	105% MR	0.47^b ±0.09	4.3^a ±0.89	0.5^b ±0.41	0.4^b ±0.27	0.5^b ±0.15

Within main effects, means with different superscripts differ significantly ($P<0.05$). Interactions between main effects were not significant.

Diet affected energy expenditure (Figure 1) in that animals fed at 105% MR had lower energy expenditure than animals fed at 180% MR, but there was no significant difference between the RFI genotypes. However, most of the variation in energy expenditure or heat production can be accounted for by the energy consumed by the animal. Hence, it suggests that there is no detectable difference in the efficiency of energy utilisation between the RFI lines (Figure 1). Energy expenditure for muscle protein synthesis was highly correlated to whole body energy expenditure ($r=0.73$). However, there was no difference in muscle protein synthesis between RFI lines and therefore, appears not to contribute to the between RFI line variation in energy expenditure

observed in these animals.

Although there is substantial variation around the relationship, if correct, the analyses suggest that selection for RFI has resulted in no detectable change in the efficiency of utilisation of feed energy. However, RFI is heritable and the phenotype of the animals has changed as intended with selection (Arthur *et al.* 2001a; Arthur and Herd 2008). What are the implications if there has been no change in the relationship between energy expenditure and energy intake?

Kennedy *et al.* (1993) and Van der Werf (2004) elegantly demonstrated that selection of a trait such as RFI where FI = Weight + Production parameter + RFI is equivalent to selection on the component traits. So if selection for RFI has resulted in a reduction in feed intake, at constant weight and daily gain (as it has in this case, Arthur *et al.* 2001b) and no change in the relationship between energy intake and expenditure (this study), it follows that the energy content of gain (= fat content) must be less. This is exactly what has been observed in these animals (Table 1) and by Richardson *et al.* (2001) and inferred by the genetic and phenotypic correlations reported by Robinson and Oddy (2004).

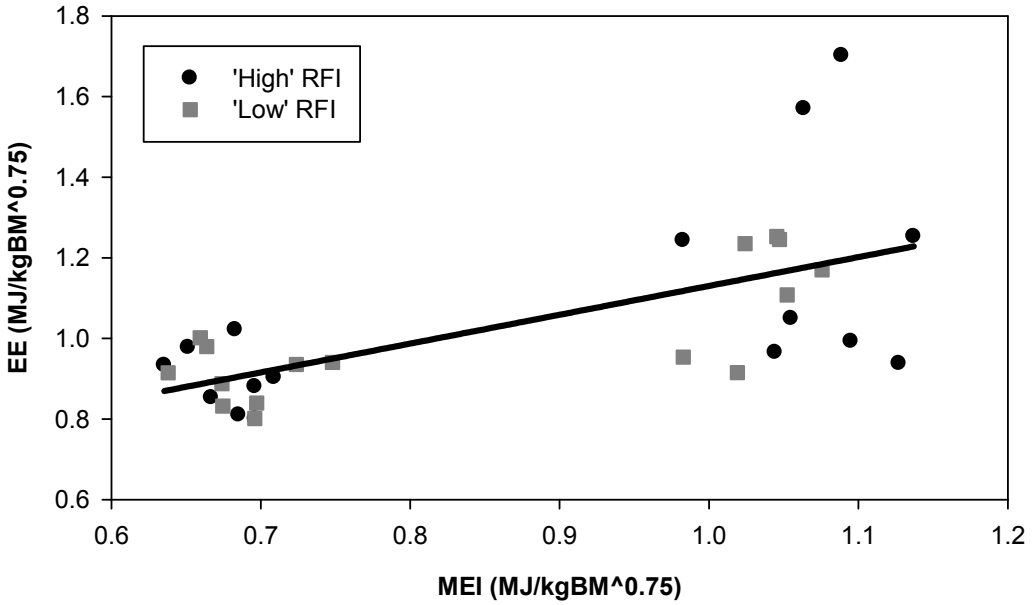


Figure 1: Energy expenditure of high and low RFI heifers against metabolisable energy-intake.

The current study was on a limited number of animals, and the measurements of energy expenditure were by indirect methods. What additional evidence is there that animals which vary in RFI do not vary unexpectedly in energy transactions? Basarab *et al.* (2003) calculated that RFI was related to the composition of liveweight gain and that some of the variation in RFI could be explained by variation in empty body fat gain. When RFI was adjusted for gain in ultrasound back fat thickness and marbling, they showed that animals with negative RFI values had lowered metabolisable energy intakes, lowered heat production (energy expenditure) and retained less

energy. They, therefore, concluded that a proportion of the metabolisable energy intakes of high RFI animals was accounted for by the differences in composition of gain. However, a much greater proportion could be attributed to differences in heat production. This relationship between metabolisable energy intake and heat production is exactly what would be expected from nutrition and energetic models.

CONCLUSION

In this study, most of the variation in energy expenditure could be accounted for by the amount of energy consumed. Energy expenditure per unit of metabolisable energy-intake did not differ between the selection lines. There was evidence for differences in fat deposition; the high-RFI animals retained more energy in fat. Modelling energy transactions suggested that there was no difference in efficiency of energy utilisation between the RFI lines. As this study and others show, any perceived differences in efficiency of the trait can be attributed to the amount of energy consumed by the animal and the divergence of fat deposition in the genotypes or phenotypes. However, as small numbers of animals were sampled in this trial, further physiological and biochemical evaluation is necessary before firm conclusions can be drawn.

Given the practical limitations for the measurement of the components of residual feed intake, (that is, actual feed intake, weight and average daily gain), then it is no surprise that selection pressure is extended on feed intake and composition of gain (relative proportion of fat and lean). Unfortunately, basal or underlying metabolic rate, the trait we desire to minimise appears at this stage to be unaltered. These concerns of selection pressure on feed intake and fat deposition associated with selection for residual feed intake should be addressed. Therefore, clearly more work needs to be undertaken to understand the full consequences of selection for RFI and before of the trait can be properly implemented within industry.

ACKNOWLEDGEMENTS

The authors thank the NSW Department of Primary Industries for access to the heifers.

REFERENCES

Arthur, P.F. Archer, J.A. Johnston, D.J. Herd, R.M. Richardson, E.C. and Parnell, P.F. (2001a) *J. Anim. Sci.* **79**:2085

Arthur, P.F. Archer, J.A. Herd, R.M. and Melville, G.J. (2001b) *Assoc. Advanc. Anim. Breed. Genetics (AAABG).* **14** :135

Arthur, P.F. and Herd, R.M. (2008) *Revista Brasileira de Zootecnia.* **37** (Suppl):269

Basarab, J.A. Price, M.A. Aalhus, J.L. Okine, E.K. Snelling, W.M. and Lyle, K.L. (2003) *Can. J. Anim. Sci.* **83**:189

Calder, A.G. and Smith, A. (1988) *Rapid Comm. Mass Spectrom.* **2**:14

Kennedy, B.W. van der Werf, J.H.J. and Meuwissen, T.H.E. (1993) *J. Anim. Sci.* **71**:3239

Li, L. Oddy, V.H. and Nolan, J.V. (2008) *Aust. J. Exp. Agric.* **48**:657

Richardson, E.C. Herd, R.M. Oddy, V.H. Thompson, J.M. Archer, J.A. and Arthur, P.F. (2001) *Aust. J. Exp. Agric.* **41**:1065

Richardson, E.C. and Herd, R.M. (2004) *Aust. J. Exp. Agric.* **44**:431

Robinson, D.L. and Oddy, V.H. (2004) *Livest. Prod. Sci.* **90**:255

van der Werf, J.H.J. (2004) *Aust. J. Exp. Agric.* **44**:405

SURVIVAL ANALYSIS FOR LENGTH OF PRODUCTIVE LIFE OF BEEF COWS

Karin Meyer

Animal Genetics and Breeding Unit*, University of New England, Armidale, NSW 2351

SUMMARY

Survival analysis is applied to records for length of productive life of Angus cows, fitting several proportional hazards models. It is shown that a piece-wise Weibull model with a time-dependent stage effect can model the distinct annual culling pattern reasonably well. Analyses fitting a sire model suggest some genetic variability, with heritability estimates on the logarithmic scale ranging from 6 to 10%. Issues of data quality and the need for whole herd inventory recording are discussed.

INTRODUCTION

Length of productive life (LPL) of cows is an important economic factor in beef cattle production. While genetic evaluation for 'survival' of dairy cows using so-called proportional hazards (PH) models is an integral part of many dairy improvement programmes, few attempts have been made to extend such analyses to beef cattle. PH models imply that each cow which has survived to a given time t has a certain risk to be culled at this time, the hazard $h(t)$. This is modelled as a baseline hazard function, $h_0(t)$, modified by certain risk factors ($\boldsymbol{\omega}$) which can be a combination of fixed and random effects, $h(t) = h_0(t)\exp(\mathbf{w}'\boldsymbol{\omega})$ (with $\mathbf{w}$ the vector of covariables). PH models are linear on the logarithmic scale, i.e. risk factors act in a multiplicative fashion. A major advantage of PH models is that censored records, i.e records for individuals still alive but likely to be culled in the future, can be included in the analysis. An introduction to survival analyses is given by Kachman (1999). This paper presents a first study applying this methodology to characterise survival of Angus cows in seedstock herds, comparing different models and examining the importance of various risk factors.

MATERIAL AND METHODS

Data. Raw data consisted of pedigree information for 1.66 million Angus cattle, extracted from the National Beef Recording Scheme data base. From these, birth dates of calves were collated for all cows occurring as dams. To accommodate embryo transfer (ET) and cross-fostering, calves were assigned to the rearing rather than the genetic dam. LPL was then defined as the number of months between the birth of the first and last calf on record. Any records less than 10 months were discarded, i.e. any cows without a second calf had LPL= 0 and were not considered. Further edits eliminated any cows with unknown sire or cows born prior to 1980 or after July 2005, and restricted age at birth of the first calf to 20 to 42 months. Records for cows with the last calf born after February 2007 were deemed to be 'censored'. For the analysis, only daughters of sires with at least 5 records were considered, and records for any herds with less than 50 records or less than 40% of cows by a sire used in more than one herd were disregarded. This yielded 98 804 records for daughters of 9 734 sires in 477 herds. Of these, 74.6% were uncensored with a mean LPL of 47.4 months and range of 10 to 174 months. To avoid problems with few records at the higher ages, 638 records above 144 months were set to this value and the records treated as censored. This increased the proportion of censored records to 26.0% while reducing the mean of uncensored observations to 46.5 months.

Analysis. Records were analysed fitting Cox or Weibull PH models, as implemented in the Survival Kit (Ducrocq and Sölkner 1998), by now in Version 5.1. All analyses fitted age of cow at birth of first calf, year of birth of cow, month of birth of cow, month of birth of first calf and month of

*AGBU is a joint venture of NSW Department of Primary Industries and University of New England

birth of last calf as discrete fixed effects. Other risk factors fitted throughout were the number of herds a cow's sire had progeny in ($1, 2-5, \geq 6$), the number of calves which subsequently became sires ($0, 1, \geq 2$), and whether or not a cow raised an ET calf, was an ET animal herself or changed herds. Herd-year of birth of first calf effects (HY, with 7330 levels) were fitted as either fixed (HY-fix), random (HY-rnd) or random within a fixed herd effect (H+HY-rnd). If treated as random, HY effects were assumed to have a log Gamma distribution with equal scale and shape parameter. For Weibull analyses, a time-dependent stage effect with up to 22 classes ($10-14, 15-22, 23-27, 28-34, 35-39, \ldots, 136-144$ months) was considered to model changes within and across years. Furthermore, analyses fitting up to 6 time-dependent strata, i.e. a piece-wise Weibull model (Ducrocq 2002), were carried out.

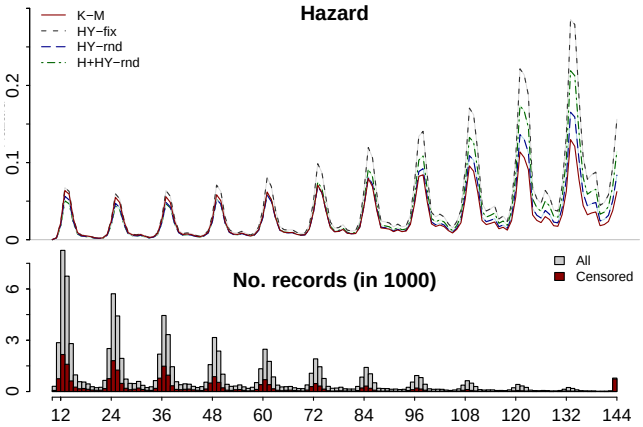


Figure 1. Hazard rates with distribution of records

Estimates of genetic variance were obtained fitting a sire model, with sire effects assumed to follow a multivariate normal distribution, accounting for pedigree information on sires and maternal grand-sires (12 681 sires in the pedigree). Pseudo-heritabilities on the logarithmic scale were computed as $h^2_{\log} = 4\sigma^2_S/(\sigma^2_S + \sigma^2_\gamma + \pi^2/6)$ (Ducrocq et al. 1988) with σ^2_S and σ^2_γ the variance components due to sires and HY effects, respectively. Values on the original scale were obtained as $h^2_{\text{org}} = \exp(\nu/\rho)^{-2} h^2_{\log}$, with $\nu \approx 0.5772$ (Euler constant) and ρ the shape parameter of the Weibull distribution, and the 'effective' heritability (Yazdi et al. 2002) was calculated as $h^2_{\text{eff}} = 4\sigma^2_S/(\sigma^2_S + \sigma^2_\gamma + 1)$.

RESULTS

Estimated hazard rates from non-parametric analyses together with the distribution of LPL records are shown in Figure 1. Both exhibit a clear annual pattern, with each year separable into 5 and 7 months periods of high and low risks and numbers of observations, respectively, and highest hazards at approximately 12 monthly intervals. Peak risks of being culled dropped slightly from 12 to 24 months and increased with age after an LPL of 48 months. Estimates of hazard from a Kaplan-Meier analysis (K-M), i.e. not fitting any risk factors, agreed well with those from Cox PH models initially, but were consistently lower at later ages. Fitting HY or herd effects as fixed inflated estimates for the highest ages substantially, suggesting some problems in the data structure due to limited numbers of records. Corresponding survival curves are displayed in Figure 2. Again, annual fluctuations are evident with periods of high and low hazards corresponding to steep and flat parts of the curves. Under the assumption of proportional hazards, a plot of log(− log(Survival)) against log(LPL) is expected to follow a straight line. Clearly this was not the case, with periods from 10 – 14 and 15 – 22 months distinctly different and linearity from about 35 months only approximately true.

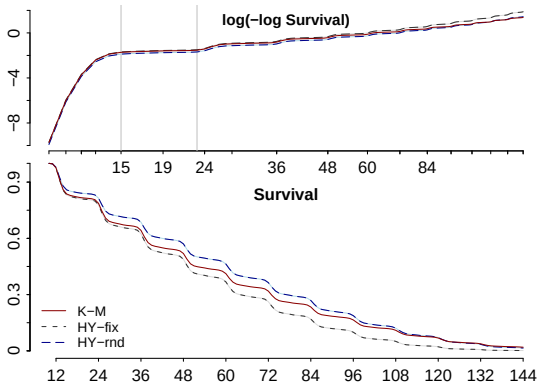


Figure 2. Non-parametric survival curves

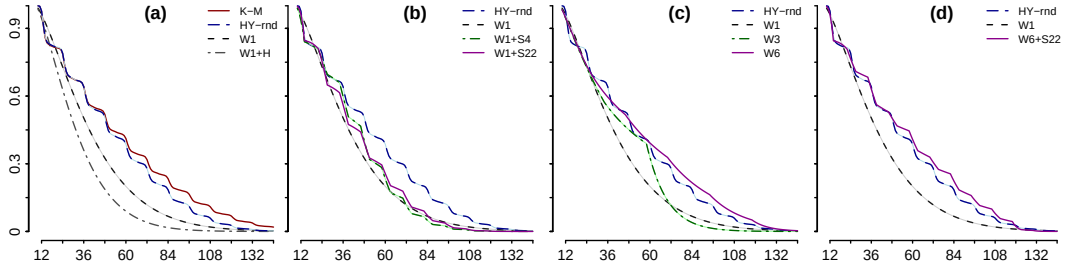


Figure 3. Estimated survival curves from Weibull analyses (see text for definitions).

Estimates of survival curves from different analyses fitting a Weibull hazard function are summarised in Figure 3. All analyses fit HY effects as random. Part (a) contrasts estimates from the non-parametric analyses (*c.f.* Figure 2) with those fitting a single Weibull curve without time-dependent covariables, either with (W1+H) or without (W1) fixed herd effects in the model of analysis. With 46% of records below 28 months, the single curve (W1) agreed reasonably well with the estimate from the Cox model, which does not make any assumptions about the shape of the baseline hazard, in this part but overestimated hazards later on, resulting in a survival curve which declined too quickly. Fitting herds in addition to HY effects exacerbated the deviations. As shown in part (b), allowing for a time-dependent stage effect modelled the wave pattern of the survival curve, but did not increase the quality of fit substantially. Whether this involved separate effects for all 22 stages (W1+S22) determined from the change-points in Figure 2, or whether the model was reduced to 4 stages (W1+S4) by merging odd/even stages from 23 months onwards into 2 effects had rather little impact. Allowing for different Weibull curves in different parts of cows' life, however, dramatically improved the fit of the Weibull model; see part (c). While the log(-log) plot suggested separate curves for records below 15 and between 15 and 23 months, such analyses proved unsuccessful, either due to numerical problems or with the resulting curves 'worse' than the single curve (W1). Fitting three strata (W3) with changes at 35 and 59 months yielded a good fit in the first two parts. However, 6 partial curves (with changes at 23, 47, 71, 95 and 119 months) were required to model the complete curve adequately. Further sub-stratification was attempted, but again failed. As shown in part (d), combining the 6-part Weibull curve with stage effects (W6+S22) yielded a model closely following the estimate from the non-parametric analyses.

Table 1 gives estimates of parameters, variance components and the resulting heritabilities (in %) from Weibull analyses in- and excluding sires, with γ the scale and shape parameter of the log Gamma distribution for HY effects and b_0 the 'intercept' of the curve, $b_0 = \rho \log(\lambda)$ where λ is the scale parameter of the Weibull function. For analyses fitting strata, $\hat{\rho}$, $\hat{b}_0$ and h^2_{org} given pertain to the first stratum. Estimates of the variance due to HY effects were large, ranging from 0.56 to 0.66 for analyses not fitting herds, corresponding to a parameter of the Gamma distribution around 2. Fitting fixed herd effects reduced both $\hat{\sigma}^2_\gamma$ and $\hat{\sigma}^2_S$ in proportion so that heritability estimates were little affected. For all analyses, fitting sires reduced the estimate of $\hat{\sigma}^2_\gamma$ slightly in spite of attempts to reduce confounding of HY and sire effects by restricting data to herds with at least 40% progeny of sires used in more than one herd. Standard errors of $\hat{\sigma}^2_S$ ranged from 0.0038 (W1+H) to 0.0051 (W1+S4). Fitting stages or a piece-wise baseline function increased estimates of σ^2_γ by up to 13% and of σ^2_S by up to 71%, resulting in heritability estimates about 50% higher than those from an analysis fitting a single Weibull curve without stages (W1). In contrast, Ducrocq (2005) found that fitting a more detailed model with a piece-wise curve reduced the estimate of the sire variance for survival of French dairy cows from 0.053 to 0.035. Results from analyses W1 were comparable to values of $\hat{\gamma} = 1.94$ and $\hat{\sigma}^2_S = 0.029$ reported by Forabosco et al. (2006) for Chianina cattle. However,

Table 1. Estimates of parameters from Weibull analyses (see text for definitions)

Model	Without sire effects				Fitting sire effects							
	$\hat{\gamma}$	$\hat{\rho}$	$\hat{b}_0$	$\hat{\sigma}^2_\gamma$	$\hat{\gamma}$	$\hat{\rho}$	$\hat{b}_0$	$\hat{\sigma}^2_\gamma$	$\hat{\sigma}^2_S$	$\hat{h}^2_{\log}$	$\hat{h}^2_{\text{org}}$	$\hat{h}^2_{\text{eff}}$
W1	2.160	1.352	-4.920	0.5861	2.250	1.362	-4.849	0.5569	0.0344	6.2	2.6	8.8
W1+H	4.872	1.367	-4.566	0.2278	5.154	1.375	-4.525	0.2141	0.0283	6.0	2.6	9.1
W1+S4	2.048	2.062	-7.293	0.6263	2.166	2.086	-7.232	0.5842	0.0590	10.3	5.9	14.4
W1+S22	2.058	3.231	-7.352	0.6224	2.164	3.243	-7.248	0.5846	0.0554	9.7	6.8	13.5
W3	2.037	1.098	-4.197	0.6266	2.155	1.104	-4.090	0.5877	0.0531	9.3	3.3	12.9
W6	1.953	1.102	-4.235	0.6645	2.050	1.107	-4.127	0.6252	0.0575	9.9	3.5	13.7
W6+S22	2.063	3.160	-7.226	0.6202	2.167	3.170	-7.116	0.5838	0.0534	9.4	6.5	13.0

Phocas and Ducrocq (2006) obtained estimates of sire variances of 0.04 to 0.05 for LPL of Charolais cattle, but variances due to HY effects of 0.04 to 0.06, differing substantially from our results.

DISCUSSION

Analyses have identified a distinctive, annual pattern for LPL of Australian beef cows with alternating stages of low and high risk, corresponding to periods of rearing a calf and the next calving. In part, this is inherent in our definition of LPL. Results have shown that survival analyses for LPL of beef cows under a Weibull model require careful modelling to adequately represent the baseline hazard function identified in non-parametric analyses, combining a piece-wise Weibull curve with time-dependent stage effects. There is clear evidence of some genetic variation. However, estimates of σ^2_S were small in comparison to environmental variation due to HY effects. Risk factor fitted, other than stage, explained relatively few differences in culling patterns. Disconcertingly, estimates of σ^2_S depended strongly on the model fitted and increased rather than decreased with more detailed modelling. It might be argued that an ill-fitting single curve (analysis W1) obscured differences between sires. Conversely, we rule out that the increase in variances was not, in part at least, due to other, unidentified systematic risk factors or problems inherent in the data structure.

The main limitation in this study has been the lack of 'proper' records for LPL, based on whole herd inventories, in particular recording of disposal dates and reasons. The measure of LPL used is a minimum and relies on registration of all calves to be informative. In addition, it does not account for culling based on failure to raise a calf until weaning. Use of weaning weights records to extract such information was disregarded as the ratio of numbers weaned to born was too low to assume complete, non-selective recording. If a genetic evaluation scheme for survival of beef cows is to be contemplated, it should be accompanied by instigating the appropriate recording scheme.

REFERENCES

Ducrocq, V. (2002) *CD-ROM Seventh World Congr. Genet. Appl. Livest. Prod.* **19**:20.
Ducrocq, V. (2005) *Anim. Sci.* **80**:249.
Ducrocq, V., Quaas, R. L., Pollak, E. J. and Casella, G. (1988) *J. Dairy Sci.* **71**:3071.
Ducrocq, V. and Sölkner, J. (1998) *Proc. Sixth World Congr. Genet. Appl. Livest. Prod.* **27**:447.
Forabosco, F., Bozzi, R., Filippini, F., Boettcher, P. J., Van Arendonk, J. A. M. and Bijma, P. (2006) *Livest. Prod. Sci.* **101**:191.
Kachman, S. D. (1999) *J. Anim. Sci.* **77**:147.
Phocas, F. and Ducrocq, V. (2006) *CD-ROM Eighth World Congr. Genet. Appl. Livest. Prod.* 03–13.
Yazdi, M. H., Visscher, P. M., Ducrocq, V. and Thompson, R. (2002) *J. Dairy Sci.* **85**:1563.

IMPUTATION OF MISSING GENOTYPES IN HIGH DENSITY SNP DATA

G. Moser[1], M.S Khatkar[2] and H.W. Raadsma[2]

The CRC for Innovative Dairy Products
[1] Bellbowrie, QLD, 4070, Australia
[2] ReproGen, Faculty of Veterinary Science, The University of Sydney, Camden, Australia

SUMMARY

The accuracy and computational complexity of five methods to impute missing genotypes in high density SNP data was investigated. The haplotype reconstruction package fastPHASE reached the highest accuracies (91% to 98%) for varying proportions (0.2% to 8%) of missing genotypes. Alternative methods based on principal component analysis were less accurate (67% to 94%), but their computational demand was an order of magnitude lower.

INTRODUCTION

Missing genotype information is a common feature of high density SNP datasets. Even if the missing rate is low, eliminating SNP markers with incomplete observations will result in considerable loss of information. Several methods exist to deal with missing genotypes, such as imputing missing values of row averages or medians, but accuracy can be significantly improved by exploiting the correlation between data. Methods using haplotype reconstruction informed by linkage disequilibrium between SNP are commonly used to infer missing genotypes. However, when the number of loci is large, these approaches are computationally demanding. Less accurate but faster methods exist, that use the global covariance structure of SNP markers on a chromosome. The purpose of this study was to compare the performance of haplotype-based with covariance-based imputation methods.

MATERIALS AND METHODS

Missing genotype imputation. Five methods were used to infer missing values (MV) in SNP data. The haplotype reconstruction package fastPHASE (Scheet and Stephens 2006) uses a Hidden Markov Model to describe the spatial distribution of clusters of haplotypes along the chromosomes. FastPHASE uses the EM algorithm to estimate genetic parameters and haplotype frequencies from which missing genotypes are inferred. FastPHASE requires the specification of 5 tuning parameters to control the algorithm, which could be inferred through cross-validation. We chose two values for the number of haplotype clusters (K parameter) viz. 10 (fastPHASE1) and 50 (fastPHASE2), the default values were used for the other parameters. For principal component analysis (PCA) based imputation, we employed four algorithms implemented in the freely available R package pcaMethods (R Development Core Team 2007, Stacklies *et al.* 2007); Probabilistic PCA (PPCA), Bayesian PCA (BPCA), svdImpute (SVDI) and Nipals PCA (NIPALS). The principle behind the first three approaches is that missing values are initially set to the row averages, and singular value decomposition of the SNP matrix is used to create orthogonal principal components. The principal components, which correspond to the largest eigenvalues are then used to reconstruct the missing SNP genotypes in the SNP matrix. Nipals PCA uses the NIPALS algorithm (non-linear estimation by iterative partial least squares, Wold *et al.* 1966) for finding the principal components of the SNP matrix. In a study using gene expression data, Brock *et al*. (2008) found that covariance-based imputation methods are highly competitive with each other, but that no method was uniformly superior across different data sets and that the optimal method depended on the correlation structure of the data. The optimal number of principal

components is the only tuning parameter required by the methods. We fitted models varying the number of principal components in steps of 20. The best model was chosen as the one that provided the smallest prediction error.

Assessment of performance. The normalised mean root square error of prediction (NRMSEP) was used as the metric to compare the performance between methods. The NRMSEP normalises the square difference between observed and estimated values for a certain SNP locus by the variance within this locus. For the PCA methods, estimates are on a continuous scale and can be less than 0 or larger than 2. An ad-hoc binning algorithm was used to assign the inferred values into distinct genotypes, where genotype 0 was assigned if the estimate was ≤ 0.5, 2 if the estimate was > 1.5, and 1 when the estimate was > 0.5 and ≤ 1.5. Accuracy was computed as the percentage of correctly imputed genotypes, the misclassification rate was calculated as the proportion of incorrectly inferred MV on the total number of genotypes in the sample.

Data. Two SNP datasets for bovine chromosome 6 (BTA6) generated by the CRC for Innovative Dairy Products were used. The original data contained 0.8% and 0.6% missing values. The first set (BTA6.1) consisted of a sample of 377 bulls genotyped for 1446 SNPs. Missing values were generated for a random selection of 25% of all loci. We randomly masked 1%, 5%, 10% and 25 % of the genotypes at the chosen loci. This corresponds to proportions of 0.25%, 1.52%, 3.14% and 7.97% newly generated MVs on the total number of genotypes. The second dataset (BTA6.2) comprised 1943 bulls genotyped for 325 SNPs. One locus was randomly selected in each segment of 50 consecutive SNPs and 10% of genotypes were assigned missing values at random, so that a total of 0.21% of all genotypes were generated as missing.

RESULTS AND DISCUSSION

Accuracy. Table 1 shows the summary statistic of the performance of the methods when we imputed an increasing proportion of genotypes for set BTA6.1. For all methods the prediction error increases with an increase in the proportion of missing genotypes. FastPHASE with tuning parameter setting 2 (fastPHASE2) was the most accurate method, with a prediction error about half the size compared to PCA methods. The number of correctly inferred genotypes was very high ranging from 98.3 % (MV=1385) to 97.3% (MV=43435). Of the PCA methods, NIPALS performed best for lower proportions of MVs, but PPCA and SVDI were more accurate if 8% of SNP genotypes were imputed. The choice of tuning parameters had a large impact on the accuracy of fastPHASE. Using fastPHASE with less optimal tuning parameters decreased the accuracy by about 6%. The number of misclassified SNP genotypes increased with the proportion of imputed genotypes. For low rates of missing genotypes (<3%) the percent of misclassified SNPs on the total number on genotypes was less than 0.5%, with the exception of BPCA. When 8% of genotypes were missing, fastPHASE2 had the lowest misclassification rate of 0.22%, which was 1% lower than the best PCA method (PPCA).

The methods were further compared for BTA6 in a second dataset (BTA6.2, Table 1) comprising a larger number of bulls (N=1943), but with a lower SNP density (N=325) in the chromosome. A total of 1333 missing values were generated, corresponding to a proportion of 0.21% of all genotypes. Again, fastPHASE2 was the most accurate method and assigned 95.8% of genotypes correctly. The best PCA methods were only slightly less accurate with NIPALS and PPCA imputing 94.05% and 93.15 of SNPs correctly. The chromosome-wide misclassification rate for NIPALS was less than 0.004% higher compared to fastPHASE2. The differences in accuracy between datasets may be an effect of the structure of the samples used and the density of

Table 1. Summary of performance of methods for missing genotype imputation

Data[1]	Missing SNP[2]	PPCA	BPCA	SVDI	NIPALS	fastPHASE1	fastPHASE2
				Prediction error (NRMSEP)			
	1385	0.546	0.756	0.555	0.501	0.513	0.206
	8304	0.571	0.817	0.581	0.536	0.527	0.288
	17103	0.579	0.833	0.588	0.554	0.523	0.288
	43435	0.606	0.887	0.618	0.636	0.531	0.286
				Accuracy (%)			
	1385	88.9	75.3	88.6	90.2	92.3	98.3
	8304	86.8	72.8	86.2	88.5	91.3	97.4
	17103	86.3	71.1	85.8	87.5	91.2	97.3
BTA6.1	43435	84.5	67.2	83.8	82.5	91.0	97.3
				Chromosome-wide misclassification rate(%)			
	1385	0.028	0.063	0.029	0.025	0.020	0.004
	8304	0.202	0.414	0.210	0.175	0.133	0.040
	17103	0.429	0.907	0.447	0.396	0.275	0.085
	43435	1.236	2.642	1.291	1.398	0.719	0.219
				Computing time			
	1385	0.58min	5.16min	4.25min	6.42min	1.15h	1.03d
	8304	1.21min	5.29min	4.19min	7.49min	1.18h	1.04d
	17103	2.59min	6.13min	4.18min	6.10min	1.00h	1.01d
	43435	6.42min	6.93min	4.21min	7.67min	1.06h	1.01d
				Prediction error (NRMSEP)			
	1333	0.145	0.190	0.156	0.145	0.162	0.123
				Accuracy (%)			
		93.2	91.7	86.7	94.1	91.5	95.8
BTA6.2				Chromosome wide misclassification rate (%)			
		0.015	0.028	0.018	0.013	0.018	0.008
				Computing time			
		7.09min	1.23min	5.08min	9.52min	1.11h	1.15d

[1] BTA6.1: 377 bulls x 1447 SNPs; BTA6.2: 1945 bulls x 325 SNPs. [2] Number of missing SNPs of 1436, 8304, 17102, 43435 and 1333 correspond to proportions of 0.25%, 1.52%, 3.14%, 7.97% and 0.21% missing values on the total number of genotypes.

SNPs on the chromosome. The data suggest that the gain in accuracy of haplotype-reconstruction methods is small for lower SNP densities. A reduction in gain in accuracy between fastPHASE and PCA methods was also found for shorter chromosomes (data not shown).

Computational complexity. The computational demand of the PCA methods is an order of magnitude lower than fastPHASE. For example imputation using fastPHASE with a tuning value of 50 haplotype clusters took more than 1 day to complete, whereas the all PCA based methods required less than 10min computing time. The estimates of PPCA, BPCA and SVDI are based on the routines implemented in the R package pcaMethods. Their speed depends largely on the computation of the singular value decomposition of the SNP matrix and implementation of these algorithms using faster programming languages will improve speed substantially. The NIPALS routine was implemented in FORTRAN and computation time was reduced by a factor of 25 compared to the pcaMethod implementation.

CONCLUSIONS

So far our analysis shows that haplotype reconstruction methods like fastPHASE, and presumably similar programs, provide the highest accuracy for inferring missing genotype data when compared to principal component based methods. The increase in accuracy might not be sufficiently large to justify its computational demand if the only purpose is to infer missing SNP genotypes. The use of fastPHASE might become prohibitive for bigger datasets, since the number of haplotype clusters normally increases with more animals in the sample and computation time increases quadratically with an increase in clusters. The accuracies of the methods depend on the choice of tuning parameters, and optimal values are usually found by cross-validation. Cross-validation adds considerably to the running time of fastPHASE. For PCA based methods the optimal number of principal components is the only tuning parameter required. The study should be extended to complete genomes and investigate the performance in relation to the accuracy of methods used to predict genomic breeding values.

ACKNOWLEDGMENTS

The genotype data was provided by the Co-operative Research Centre for Innovative Dairy products.

REFERENCES

Brock, G.N., Shaffer, J.R., Blakesley, R.E., Lotz, M.J. and Tseng, G.C. (2008) *BMC Bioinformatics* **9:**12.

R Development Core Team (2007) *Statistical Computing*, http://www.R-project.org.

Scheet, P. and Stephens, M. (2006) *Am J Hum Genet* **78**:629.

Stacklies, W., Redestig, H., Scholz, M., Walther, D. and Selbig, J. (2007) *Bioinformatics* **23**:1164.

Wold, H. (1966) In "Multivariate Analysis", p. 391, editor P.R. Krishnaiah, P.R., Academic Press, NY

PRELIMINARY RESULTS ON THE COMPARATIVE PERFORMANCE OF PRIMIPAROUS HOLSTEIN AND FLECKVIEH X HOLSTEIN DAIRY COWS

C.J.C. Muller[1], J.A. Botha[1], S.W.P. Cloete[1,2] and J.P Potgieter[3]

[1]Western Cape Department of Agriculture, IAP, Private bag X1, Elsenburg 7607, South Africa
[2]Department of Animal Sciences, University of Stellenbosch, Stellenbosch 7602, South Africa
[3]Simmentaler/Simbra Cattle Breeders' Society, PO Box 3868, Bloemfontein 9300, South Africa

SUMMARY

In this paper, preliminary results on the comparative performance of primiparous Holstein (H) and Fleckvieh x Holstein (FxH) cows are presented. Fleckvieh is a a dual-purpose breed derived from the Simmental breed in Germany. Twenty four H and 24 FxH heifers were sourced from a 1800 cow commercial dairy herd. Heifers were reared similarly to first calving. Cows were kept in feedlot style open camps and fed a total mixed ration twice a day. The 305-d milk yield, milk composition, live weight and reproduction parameters were recorded for cows of both breed types and compared. The milk, fat and protein yields of FxH cows did not differ ($P>0.05$) from their H contemporaries, although the fat and protein percentages of the milk of crossbreds was higher ($P<0.05$) than that of purebreds. The live weight of crossbreds was higher, although not significantly ($P>0.05$), at one week before and after calving. Reproduction parameters did not differ ($P>0.05$) between breeds. Further studies should be conducted to determine efficiencies of milk, fat and protein yields.

INTRODUCTION

Crossbreeding in dairy cows has become an important issue in the South African dairy industry. Farmers expect improved farm income from crossbreeding because of improved cow fertility, cow health and calf survival. These aspects seem to have become major problems in the Holstein breed (Funk 2006). Data recording schemes do not always record these so-called secondary traits while low heritabilities for these traits complicate their inclusion in traditional breeding programmes (Freyer *et al*. 2008). Dairy breeds used mostly in crossbreeding studies include Jerseys and Ayrshires. Heins *et al*. (2008), McAllister (2002) and Touchberry (1992) compared Jersey x Holstein, Ayrshire x Holstein and Guernsey x Holstein crossbreds to pure Holsteins, respectively. Dual-purpose breeds such as the Fleckvieh, a Simmental derived breed, have not been seriously considered in crossbreeding programmes. In the 1960's, Canadian Holsteins were included in a crossbreeding programme in Germany to produce a composite milk-emphasized, dual-purpose dairy breed (Schönmuth, 1963). Recently Walsh *et al*. (2008) reported on the differences among breeds such as Holstein-Friesian, Montbéliarde, Normande, Norwegian Red, Montbéliarde x Holstein-Friesian and Normande x Holstein-Friesian. Fleckvieh is a true dual-purpose breed having high milk yields and milk quality traits while in some countries it is primarily used for beef production (Grogan *et al*. 2005). In some parts of Germany, crossbreeding of German Holsteins are underway as crossbreeding is useful to improve fertility and productive life of dairy cows (Swalve, 2007). The aim of this study is to compare milk production parameters, live weight and some fertility parameters of primiparous Holstein (H) and Fleckvieh x Holstein (FxH) cows in a feedlot system.

MATERIALS AND METHODS

Location and Animals. Twenty four H and 24 FxH heifers were sourced from a 1800 cow commercial Holstein dairy herd at approximately one week of age. Heifers were progeny from two

Fleckvieh bulls (Hippo and Randy) and a Holstein bull (Jerome Red). Heifers were reared similarly to first calving at the Elsenburg Research Farm of the Western Cape Department of Agriculture. Elsenburg is situated approximately 50 km east of Cape Town in the winter rainfall region of South Africa. The area has a typical Mediterranean climate with short, cold, wet winters and long, dry summers. Production systems in this area consist of intensive feedlot style open camps or housing systems for dairy cows. Locally produced roughages are cereal crops such as oats, barley and triticale while lucerne hay is imported from the summer rainfall areas. Being a wheat producing area, wheat straw is also available and included in rations. These roughages generally have to be supplemented with large quantities of concentrates to sustain high milk yields. In this experiment, a typical total mixed ration (TMR) providing 17% CP and 11 MJ ME/kg DM consisting of lucerne hay, oat silage, wheat straw and a commercial concentrate mixture , was fed to cows in open camps with fence-line feeding troughs. Sufficient amounts of the TMR was fed twice a day to ensure an *ad libitum* feed intake. Fresh drinking water was freely available at all times. No protection was provided against summer heat and winter rain. Cows were machine-milked twice a day in a milking parlour approximately 500m from the open camps.

Data recording. The milk yield and milk composition of all pure- and crossbred cows were recorded according to standard milk recording procedures. This entailed the recording of the daily (evening and next morning's) milk yields of each cow every five weeks starting from five days after calving to drying up. Milk samples were collected at the afternoon milking for each cow and analysed for fat, protein and lactose concentrations using a Multi-Spec Infra-red Analyser. Following each milk recording event, the National Dairy Cattle Performance Testing Scheme of the Agricultural Research Council provide projected 305-d milk, fat, protein and lactose yields and a persistency value (the daily milk yield at 60-d divided by 280-d milk yield) for each cow using standard lactation curves for the South African Holstein breed (Mostert *et al*. 2003). Cows were weighed one week before and after calving and thereafter on a monthly basis until the end of the lactation period. Cows were observed for heat detection and inseminated from 30 days after calving. The reproductive performance of each cow was determined based on insemination dates and the result of rectal palpations by a veterinarian at least 45 days after the last insemination. The following reproduction parameters were determined: pregnancy rate (irrespective of number of times inseminated), the interval (number of days) from calving to first insemination, rate of first insemination within 80 days after calving, interval from calving to conception and pregnancy rate within 200 days after calving

Statistical analyses. The traits were compared by one-way analysis of variance with breeds as treatments and cows within breed as replicates. Breed means and probabilities of differences are provided. Monthly means for production parameters were calculated based on days in milk at each milk recording event throughout the lactation period. Linear regressions were fitted on monthly mean production parameters. Frequencies were compared in 2 X 2 tables, using Fisher's exact test.

RESULTS AND DISCUSSION

Calving started in August 2007 for both genotypes with the average calving date for the H and FxH cows being 5 and 13 December 2007, respectively. Age at first service was similar for both genotypes (*ca* 15.0 months) mainly because of management reasons. Mean (±SE) age at first calving was 25.7±0.6 and 26.4±0.6 months for FxH and H cows respectively ($P>0.05$). From the original 48 heifers, 18 H and 20 FxH had actual 305-d lactation records for milk, fat and protein yields as well as fat, protein and lactose percentages. Seven cows (4 H and 3 FxH) had short lactations (<240-days) because of health problems. The projected 305-d lactation records of cows were used for these individuals. The comparative milk traits of cows are presented in Table 1. The

milk yield of H and FxH cows varied from 4550 to 9319 and from 3222 to 9224 kg, with coefficients of variance of 19 and 23%, respectively. While average milk yield favoured H cows in absolute terms, this difference was not significant (P=0.30). With the exception of fat and protein percentages, other milk production traits did not differ between breeds (P>0.10). Because of higher (P<0.05) fat and protein percentages, the fat and protein yields of FxH cows were similar to H cows.

Table 1. The mean(±se) 305-d milk production and milk composition of first lactation Holstein and Fleckvieh x Holstein cows receiving a total mixed ration.

Parameters	Holstein	Fleckvieh x Holstein	P
Number of cows	22	23	-
Milk yield (kg)	6519±261	6109±289	0.30
Fat (%)	4.02±0.07	4.29±0.06	0.01
Fat yield (kg)	259±8	260±11	0.98
Protein (%)	3.32±0.05	3.49±0.04	0.01
Protein yield (kg)	215±8	213±9	0.82
Lactose (%)	5.58±0.24	5.16±0.17	0.17
Persistency (%)	105±4	102±4	0.64

The lactation curve for the daily milk yield of H cows showed a peak at the fourth month post calving. Generally the trend in milk yield over the lactation was linear, R^2=0.15 for H and R^2=0.03 for FxH cows. Both trends showed an increase towards the end of the lactation (Figure 1). No information is available in the literature on the lactation curves for Fleckvieh or similar type breeds. Mostert *et al*. (2003) showed that the lactation curve for milk yield of first lactation H cows in the South African national data set show an increase to 60 to 70 days post calving after which milk yield is reduced. First lactation Jersey cows show a downward trend in milk yield from the first test to the end of the lactation. Trends in fat and protein percentages did not differ (P>0.05) between breeds

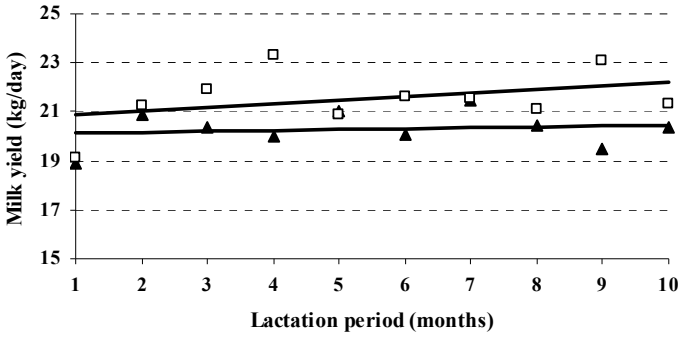


Figure 1. Daily milk yields of first lactation H (□) and FxH (▲) cows.

On average, cows from the two genotypes did not differ for weight at first calving, i.e. 631±10 kg for FxH vs. 608±13kg for H cows (P=0.17), although absolute figures favoured the FxH genotype. Holstein cows lost more weight (P<0.01) than crossbreds from calving to nadir, i.e. -50±10 vs -18±8kg while FxH cows gained weight at a faster rate (P<0.01) than H cows towards the end of the lactation .

With the exception of one cow in each genotype, all cows became pregnant. A similar proportion of FxH was inseminated within 80 days post calving (0.65 vs 0.55; Fisher's exact probability = 0.34). However, the interval from calving to first insemination tended (P=0.08) to be slightly shorter in FxH cows than in H contemporaries (74±4 vs 90±12 days). The interval from calving to conception was similar for both breeds, i.e. 132 and 126 days for H and FxH cows respectively. The proportion of cows confirmed pregnant by 200 days in milk was 0.83 for FxH and 0.64 for H cows (Fisher's exact probability = 0.47). Clearly more data is required to adequately assess the reproductive performance of the two genotypes.

Results obtained in the present study are similar to an ongoing study at Iden Experimental Station in Germany where H cows are crossed with Brown-Swiss (BS), a milk-emphasized dual-purpose breed in Germany. F1 (BSxH) cows reached almost the same milk yields accompanied by higher (P<0.05) fat and protein percentages (Swalve, 2007). Because of the small number of experimental animals in this study, results should be treated with caution. This is exploratory research for this area. A small number of local farmers have been crossbreedng using Fleckvieh sires. No crossbreeding studies have been done in this country because of a lack of facilities. Funds were also limited resulting in the small number of experimental animals. The project is ongoing and progeny from the original H and FxH cows will be included in the dataset. Farmers taking part in crossbreeding are encouraged to do milk recording to increase the number of crossbred animals. Further studies are invisaged to include reproduction and lifetime performance of crossbred and purebred cows.

CONCLUSIONS

The milk, fat and protein yields of FxH cows did not differ (P>0.05) from H cows while fat and protein percentages of the milk of FxH was higher (P<0.05). The difference in live weight at first calving between H and FxH cows was not significant, but H cows lost more weight (P<0.050) after calving and gained weight at a slower rate towards the end of the lactation. Reproduction parameters did not differ between breeds, possibly as a result of inadequate numbers. Further studies should be conducted to determine the efficiency of milk, fat and protein yields.

ACKNOWLEDGMENTS

This research is being supported by the Simmentaler/Simbra Cattle Breeders' Society of Southern Africa and the World Simmental and Fleckvieh Federation and funded through the Western Cape Animal Production Research Trust. The efforts of the management team at the Elsenburg dairy are also greatly appreciated.

REFERENCES

Freyer, G., König, S., Fischer, B., Bergveld, U. and Cassell, B.G. (2008) *J. Dairy Sci*. **91**: 3725.
Funk, D.A. (2006) *J. Dairy Sci*. **89**: 1362.
Grogan, A., Wickham, B., Olori, V. & Cromie, A. (2005) *Stočarstvo* **59**: 339.
Heins, B.J., Hansen, L.B., Seykora, A.J., Johnson, D.G., Linn, J.G., Romano, J.E. and Hazel, A.R. (2008) *J. Dairy Sci*. **91**: 1270.
McAllister, A.J. (2002) *J. Dairy Sci*. **85**: 2352.
Mostert, B.E., Theron, H.E. and Kanfer, F.H.J. (2003) *S. Afr J. Anim. Sci*. **33**: 70.
Schönmuth, G. (1963) *Arch. Anim. Breed*. **2**: 79.
Swalve, H.H. (2007) *Lohmann Inf*. **42**: 38.
Touchberry, R.W. (1992) *J. Dairy Sci*. **75**: 640.
Walsh, S., Buckley, F., Pierce, K., Byrne, N., Patton, J. and Dillon, P. (2008) *J. Dairy Sci*. **91**: 401.

Poster

ISOLATION AND CHARACTERIZATION OF ALPACA TETRANUCLEOTIDE MICROSATELLITE MARKERS

K.A. Munyard, J.M. Ledger, C.Y. Lee, C. Babra and D.M. Groth

Western Australian Biomedical Research Institute, Curtin Health Innovation Research Institute, School of Biomedical Sciences, Faculty of Health Sciences, Curtin University, GPO Box U1987, Perth Western Australia 6845.

SUMMARY

Hybridisation-capture was used to create 12 unique alpaca DNA libraries each enriched for a different tetranucleotide microsatellite motif. Two hundred and forty nine microsatellites were found, of which 26 were polymorphic (motifs GGAT, GTTT and GCAC). Nine markers were fully characterised on 45 samples. Allele numbers ranged from 6 (Locus P135) to 12 (loci P149 and PCTD17). There was no evidence of linkage disequilibrium ($p = 0.064 – 1$) or deviation from Hardy-Weinberg equilibrium ($p = 1$). Polymorphic information content ranged from 0.48 to 0.82. When combined, the markers had an exclusion probability of 97.7%. These markers will be useful for parentage determination (especially if combined into a multiplex) and will add to the pool of markers available for mapping of desirous or deleterious traits in alpacas.

INTRODUCTION

Alpacas are high-value animals because of the exceptional quality of their fibre. In their region of origin, South America, they are also used extensively as meat and pack animals, roles that are slowly becoming more important in Australia. Currently, there is very little research being conducted on Alpaca genetic traits. One particularly useful tool for studying genetic traits is a set of molecular markers. Availability of markers would facilitate the precise mapping of desirous or deleterious traits within a family, and can ultimately result in the discovery of gene(s) responsible for these traits.

Despite the recent rise in popularity of single nucleotide polymorphisms (SNP) for genetic studies, microsatellite markers are still considered useful, and for some applications, superior. Fewer than 150 characterised microsatellite markers are published for Alpacas (Lang *et al.* 1996, McPartlan *et al.* 1998, Obreque *et al.* 1998, 1999, Penedo *et al.* 1998, 1999a, b, Sarno *et al.* 2000, Reed and Chaves 2008). Reed and Chaves (2008) report an additional 1516 putative loci obtained through BLAST search of the 2x alpaca trace archive. However, even if all the putative microsatellites are able to be converted to usable markers, the number available limits the scope of genetic studies on the Alpaca. Many more markers are needed to create uniform coverage of the genome and thus facilitate accurate mapping of traits. For example, Ihara *et al.* (2004) placed almost 4000 microsatellite markers onto the bovine map to create a high density of genome coverage; and Watanabe *et al.* (1999) used over 5000 markers to form a radiation hybrid map of the mouse genome. A genome map will be an invaluable tool for future studies to develop genetic tests for disease as well as for coat colour and fibre quality.

Another limitation of the available alpaca microsatellite markers is that no tetranucleotide markers form part of this set, only di- and trinucleotides. Tetranucleotide microsatellites have the advantage that they have a much lower incidence of shadow bands, or stutter bands that occur during amplification compared with tri- and particularly dinucleotide microsatellite repeats. The observed polymerase error frequency for tetra-nucleotides is approximately 20-fold lower than for di-nucleotide repeats (Eckert *et al.* 2002), most likely due to the presence of fewer sites for misalignments per unit length of DNA (Katti *et al.* 2001). It is expected that fewer tetranucleotide microsatellites will be found in the alpaca genome compared with di- and trinucleotide microsatellites due to the relatively lower incidence of these motifs in mammals (Toth *et al.* 2000).

MATERIALS AND METHODS

Blood (approximately 5mL) was collected from 45 alpacas of both sexes, from herds situated in Western Australia, Victoria and New South Wales. Thirty nine of these animals were unrelated (3 generations). DNA was extracted using the salt precipitation method described by Miller *et al.* (1988). When the quantity of blood obtained was insufficient for the salt precipitation method (<1mL), the DNeasy blood and tissue DNA extraction kit (Qiagen) was used, according to the manufacturer's instructions.

DNA from 5 alpacas was pooled, and enrichment for various tetranucleotide microsatellite motifs (Table 1) was performed. The DNA hybridisation-capture method (Gardner *et al.* 1999, Hamilton *et al.* 1999, Zane *et al.* 2002) was used except that colonies produced from the capture were transferred to nylon membranes and subjected to a second selection process of hybridisation with radioactively labelled probes before being designated positive, and selected for sequencing.

Plasmid DNA was extracted using the AxyPrep Plasmid miniprep kit (Axygen) according to the manufacturer's instructions. Sequencing was performed using the ABI Big Dye Terminator® system and M13 sequencing primers. Products were separated on a 48-capillary ABI 3730 DNA analyser. Vector NTI software (Invitrogen) was used to visualise and analyse sequencing results.

Primers flanking tetranucleotide microsatellite regions were designed using Primer 3 (Rosen and Skaletsky 2000). Each 10µl polymerase chain reaction (PCR) contained genomic DNA (50-100ng), 2 µM of forward and reverse primer (Sigma Genosys), 1× Polymerisation buffer (Fisher biotec), 2mM $MgCl_2$ (Fisher biotec) and 0.75u BIOTAQ polymerase (Bioline). Amplification conditions were: 95°C for 2 min; 35 cycles of 95°C for 30 s, T_a (see Table 2) for 30s and 72°C for 1 min; then 72°C for 15 min. PCR products were visualised on 15% acrylamide gels (BioRad). For full characterisation of a marker (Table 2), each forward primer was labelled with one of three WellRed dyes (Table 2). No change in PCR conditions was required when using labelled primer as compared with unlabelled primer. Labelled PCR product was separated on a CEQ800 DNA Analyser (Beckman Coulter). Size standard 600 (Beckman) was included with each sample to allow identification of allele sizes.

GenePop 3.4 (Raymond and Rousset 1995) was used to calculate: the number of alleles, Hardy-Weinberg equilibrium, linkage disequilibrium, expected and observed heterozygosity for each locus. Bonferroni correction for multiple comparisons was applied (Rice 1989). PowerStats v1.2 (Tereba 1999) was used to calculate polymorphic information content and power of exclusion for each locus.

RESULTS

Twelve unique enriched alpaca DNA libraries were made, each enriched for a different tetranucleotide motif (Table 1). A total of 249 markers were found, with 142 (57%) being suitable for further analysis (Testable). Reasons for exclusion of markers at this stage of the research were: lack of flanking sequence, location within a SINE, repeat isolation of an already identified marker, or presence of an adjacent dinucleotide marker. Forty four (31%) of these 142 markers amplified cleanly in PCR and no non-specific amplification was evident (Specific). Of these 44 markers 26 (59%) were polymorphic when tested on 10 unrelated animals (Polymorphic). These 10 animals showed a minimum of two and a maximum of eight alleles for each marker. Fourteen of the most polymorphic markers were tested on two alpaca families (sire half-sib and dam half-sib). Two markers did not segregate in a Mendelian fashion, and these were removed from further analysis.

Nine markers were fully characterised on all 45 samples (Tables 2 and 3). Locus P135 had the lowest number of alleles (6) and the maximum observed alleles (12) was found in both loci P149 and PCTD17. Minor allele frequency varied from 0.011 to 0.014, and major allele frequency from 0.29 to 0.7. There was no evidence of linkage disequilibrium ($p = 0.064 – 1$) or deviation from Hardy-Weinberg equilibrium ($p = 1$). Polymorphic information content ranged from 0.48 to 0.82.

When combined, the makers had an exclusion probability of 97.7%.

Table 1 Summary of alpaca tetranucleotide loci isolated, the number that were suitable for further analysis (testable), the number that were able to be amplified using PCR (specific), and the number of polymorphic markers identified.

Motif	No. isolated	No. testable	No. specific	No. polymorphic
TCCC	0	0	0	0
GCTT	0	0	0	0
GCAC	41	23	3	2
TGCC	0	0	0	0
AAGG	0	0	0	0
GACA	3	1	0	0
GATA	0	0	0	0
GGAT	167	96	37	22
GAAA	0	0	0	0
GTTT	27	16	4	2
CATA	9	5	0	0
GCAT	2	1	0	0
Total	249	142	44	26

Table 2 primer sequences of the 9 fully characterised tetranucleotide microsatellite loci, each forward primer was labelled with the indicated WellRed label (Sigma Genosys).

Locus	Primer sequences	T_a (°C)	Primer label
P149	F:ATCAGGCTCCATTTTTGTGG	58	D4
	R:GTCCATCCTCCAGCACCTAA		
PCTD17	F:CCCTCTCACCTGTCTACTTG	62	D3
	R:GTATTCTGGCATTGGTTTGT		
P194	F:AGCAGGTGAAAAGCAGAATTGTGTG	59	D3
	R:AGTTTTTCCATTGCCGTTGTCAGAG		
P193	F:AAACCAATCCCCCATATATACAGAGG	57	D2
	R:AAAGAAACGAAGAACCTCCCCTGAC		
P147	F:TTAGCACCCAGCACCCTAAC	62	D2
	R:CAGGGTGTCTTTTTCCATCA		
P135	F:TGAATACAGAGGTTTCTGGCTCT	52	D3
	R:CACCTCCCTAAGGCCTCTTC		
P132	F:CAGAGGAGGGACCACTAATGCTGGC	63	D2
	R:GGGGCAAGTGAAGTGAGTGAAATGG		
P86	F:TTCCTTTCATTTGTCCACTC	56	D4
	R:TAGACCAGAAGTGTGGAAGG		
P57	F:CATGTCTTGTTGTAACCGCA	58	D2
	R:CTAAGTTCAAACCTCAGTGC		

DISCUSSION

These are the first reported tetranucleotide microsatellite markers available for alpacas. Although only 59% of the markers were polymorphic in the tested animals, given the probable founder effect in the Australian alpaca population, it is probable that more of the markers will be polymorphic in more diverse alpaca populations. The use of dinucleotide microsatellites has been replaced by tetranucleotide microsatellites in human genetic studies because the latter are technically more robust and less open to data misinterpretation (Ekert *et al.* 2002). If these markers can be combined into a multiplex, they are suitable to replace the existing panel of 10 or 12

dinucleotide markers that is currently being used worldwide for alpaca parentage testing. The low incidence of tetranucleotide microsatellites in alpacas compared with other species means that any whole genome analysis will need to rely mainly on di- and trinucleotide microsatellite markers.

Table 3 characteristics of 9 Alpaca tetranucleotide microsatellite loci. The loci were screened using 45 alpacas from herds in Western Australia, Victoria and New South Wales. Observed (H_O) and expected (H_E) heterozygosity, polymorphic information content (PIC) and paternity exclusion power (PE) are shown for all loci.

Locus	Repeat motif	Allele size range	No. alleles	H_O	H_E	PIC	PE
P149	(GGAT)n	204-358	12	0.72	0.632	0.65	0.326
PCTD17	(GGAT)n	89-217	12	0.793	0.86	0.74	0.357
P194	(GTTT)n	96-178	7	0.517	0.665	0.59	0.178
P193	(GTTT)n	147-261	8	0.28	0.728	0.65	0.167
P147	(GGAT)n	221-446	12	0.4	0.533	0.48	0.075
P135	(GGAT)n	212-236	6	0.625	0.792	0.63	0.220
P132	(GGAT)n	70-111	10	0.806	0.857	0.82	0.558
P86	(GGAT)n	185-335	8	0.625	0.698	0.56	0.435
P57	(GGAT)n	178-363	11	0.766	0.792	0.73	0.581

This research was funded by Morris Animal Foundation.

REFERENCES

Eckert, K.A., Mowery, A. and Hile, S.E. (2002) *Biochemistry* **41**:10490.
Gardner, M.G., Cooper, S.J.B., Bull, C.M., and Grant, W.N. (1999) *J. Hered* **2**:90.
Hamilton, M.B., Pincus, E.L., Di Fiore, A. and Fleischer, R.C. (1999) *BioTechniques* **27**:500.
Ihara, N., Takasuga, A., Mizoshita, K., Takeda, H., Sugimoto, M., Mizoguchi, Y., Hirano, T., Itoh, T., Watanabe, T., Reed, K.M., Snelling, W.M., Kappes, S.M., Beattie, C.W., Bennet, G.L. and Sugimoto, Y. (2004) *Gen. Res.* **14**:1987.
Katti, M.V., Ranjekar, P.K. and Gupta, V.S. (2001) *Molecular Biology and Evolution* **18**:1161.
Lang, J.D., Want, Y. and Plante, Y. (1996) *Anim. Genet*. **27**:293.
McPartlan, H.C., Matthews, M.E. and Robinson, N.A. (1998) *Anim. Genet*. **29**:158.
Miller, S.A., Dykes, D.D. and Polesky, H.F. (1988) *Nuc. Acids Res.,* **16:**1215.
Obreque, V., Mancilla, R., Garica-Huidobro, J., Gothran, E.G. and Hinrichsen, P. (1999) *Anim. Genet*. **30**:397.
Obreque, V., Coogle, L., Henney, P.J., Bailey, E., Mancilla, R., Garica-Huidobro, J., Hinrichsen, P. and Gothran, E.G. (1998) *Anim. Genet*. **29**:461.
Penedo, N.C., Caetano, A.R. and Cordova, K.I. (1998) *Anim. Genet*. **29**:411.
Penedo, N.C., Caetano, A.R. and Cordova, K.I. (1999a) *Anim. Genet*. **30**:399.
Penedo, N.C., Caetano, A.R. and Cordova, K.I. (1999b) *Anim. Genet*. **30**:166.
Raymond, M. and Rousset, F. (1995) *J. Hered.* **86**:248.
Reed, K.M. and Chaves, L.D. (2008) *Anim. Biotech.* **19**:243.
Rice, W.R. (1989) *Evol*. **43**:223.
Rozen, S. and Skaletsky, H.J. (2000) In: "*Bioinformatics Methods and Protocols: Methods in Molecular Biology",* p 365, editors S.S. Krawetz and Misener, Humana Press, Totowa, NJ.
Sarno, R.J., David, V.A., Franklin, W.L., O'Brien, S.J. and Johnson, W.E. (2000) *J. Mol. Ecol*. **9**:1919.
Tereba, A. (1999) *GenePrint* 14.
Toth, G., Gaspari, Z. and Jurka, J. (2000) *Gen. Res.* **10**:967.
Zane, L., Barglloni, L. and Patarnello, T. (2002) *Mol. Ecol.* **11**:1.

A DECADE OF SHEEP IMPROVEMENT LIMITED (SIL)

S-A.N. Newman[1], J.C. McEwan[1] and M.J. Young[2]

[1]AgResearch, Invermay Agricultural Centre, Private Bag 50034, Mosgiel 9053, New Zealand
[2]SIL, Meat & Wool New Zealand, PO Box 39-085, Harewood, Christchurch 8545, New Zealand

SUMMARY

In the decade since its commencement SIL has continued to develop as a performance recording service, expanding the range of traits able to be recorded and improving the range of analyses and selection indices available. There has been increased focus on identifying sources of superior genetics, with the SIL-ACE analysis including data from over 300 flocks and 6 million animals from a range of breeds. Ongoing research is reflected in the addition of new goal trait groups including disease traits such as facial eczema, resistance and resilience to internal parasites and dag score. On-farm production system changes have lead to the inclusion of new reproductive traits such as hogget lambing and twinning rates and processing industry developments such as yield grading of carcasses are also offering new opportunities. A major change has been the introduction of DNA technologies for defining pedigree and for gene tests of productive traits.

INTRODUCTION

The New Zealand sheep industry has had access to genetic evaluation systems for over 40 years (Callow 1985). However by the 1990s there were a number of schemes offering differing systems for analyzing and presenting information which caused confusion within the sheep breeding industry.

Geenty (2000) stated the aim of Sheep Improvement Limited (SIL) was to facilitate effective genetic improvement across the New Zealand sheep industry in cooperation with the previously existing sheep genetic improvement schemes. The database was also seen as a vehicle for transfer of research technology to the industry and to allow extensive breeding data to be used for further research.

The introduction of SIL provided breeders with routine BLUP analysis under a full animal model providing the ability to rank animals across sexes and ages using information from relatives, including ancestors and descendants. SIL provided on-demand access to across flock analyses, something previously only available in a limited manner to a small number of breeding groups.

USAGE

Use of SIL has grown dramatically to over 1000 active flocks, of which three-quarters are performance recording. There are close to 10 million animals on the SIL database, with about 400 thousand more added each year, making SIL the largest genetic database for sheep in the world. The median number of new animals added each birth year is 630 per flock for dual purpose sheep and 240 per flock for terminal sire sheep (Young and Wakelin 2009).

The number of animals included in genetic analyses has increased as depth of reliable data has increased, with 10 years data typical for analyses when SIL commenced, increasing to 18-20 years data in current analyses. The number of flocks in an analysis has also increased with more than one third of analyses having more than one flock in the analysis. There are more than 25 across-flock groups, with the number of flocks in the groups ranging from 5 to 70 and an average of 23 flocks. The number of analyses undertaken by across-flock groups is determined by the individual group and varies from weekly to annually, although most flocks run 3 or 4 official group analyses for reporting each year.

Increases in analysis sizes have been matched by improvements in processing capability to maintain rapid turnaround times for users. A typical BLUP evaluation for 2000 ewes over 12 years of data for Growth, Wool and Reproduction took about 25 minutes to run in 1999 but now takes only 5 minutes.

NEW GOAL TRAIT GROUPS

SIL has a unique role in delivering research outcomes to the wider sheep industry. Over the last decade SIL has added further genetic evaluation modules for traits of economic importance to the New Zealand sheep industry including resistance to facial eczema (Morris *et al.* 1994), resilience to internal parasites (Morris *et al*. 2001), dag score (McEwan 1992), hogget lambing and twinning rate (increased twinning rate at same lambing percentage) (Amer & Bodin 2006), as well as revision of existing modules to reflect industry, market and technical developments.

Implementation of these goal trait groups in SIL means they can be readily accessed by breeders. Some, such as facial eczema and resistance or resilience to internal parasites, require registration of the breeder where protocols must be followed to ensure valid data is collected for analysis.

Historically, genetic evaluations for meat production within SIL were designed to reflect the criteria on which payments were made to commercial farmers and as such focused on the weights of lean and fat in the whole carcass. Individual meat processors are now developing or purchasing carcass grading systems which not only give better estimates of the weight of fat and lean but also give indications of retail yield within individual cuts as well as overall, and these companies will reward for high yielding carcasses in their payment system. SIL has recently developed a new "MEAT YIELD" goal trait group which is designed to use data from carcass grading systems to estimate BVs for lean weight in the shoulder, loin and hind leg primal cuts, along with total carcass lean and fat (Jopson *et al.* 2009a, 2009b).

Currently SIL is collaborating in work across on a wide range of issues of current or developing importance to the sheep industry. These include revision of lamb survival (Kerslake *et al.* 2005), easy care fleeced sheep (bare breech and belly) (Scobie 1997), ewe longevity and efficiency (Sise *et al.* 2009) and reduction in green house gas emissions.

SIL-ACE

Although by the early 2000s there were a number of group breeding schemes routinely undertaking across flock analyses, there were few links between these groups even where groups were of the same breed. Genetic connections between different breeding groups were greatly improved with the establishment of the Meat & Wool New Zealand Central Progeny Test (formerly Alliance Central Progeny Test). One hundred and forty three rams from different breed groups have been used as sires at one or more of the three CPT sites since 2002, establishing good genetic connections between many groups.

SIL now routinely undertakes a large across-flock evaluation, known as SIL-ACE. This is an across-flock, across-breed genetic evaluation that produces reports on animals from genetically connected flocks. When SIL-ACE started in 2004, there were 151 flocks and 1.2 million animals in the evaluation. Now 335 flocks participate with over 3.2 million animals. The number of flocks connected for each major goal trait group are; Growth 269; Meat 197; Wool 151; Reproduction 232; WormFEC 58. More than 75 million breeding values are produced, making this the largest genetic evaluation of sheep in the world (Young & Newman 2009).

DNA TECHNOLOGIES

SIL has been involved in the delivery of DNA technologies to the sheep industry in recent years, including facilitating accurate recording of data and results to routine calculation of breeding values accounting for DNA parentage results.

Parentage determination in extensively farmed sheep is currently the most widely used test (Crawford *et al.* 2007). Increasingly these tests are being supplemented by the use of performance trait markers, either by themselves or in conjunction with parentage testing. Commercial single marker tests currently available include meat yield, prolificacy, parasite resistance and production, lamb survival, footrot and scrapie susceptibilities. The number is increasing by 1 to 2 additional tests per year (McEwan 2009).

Future developments in this area are likely to be some form of genome wide selection where the test explains a major part of the genetic variation in an animal for a range of traits (Dodds *et al.* 2007, Sise *et al.* 2008). This technology depends on the creation and use of high density SNP chips, something that has already been implemented in dairy cattle in several countries. A similar chip has been created for sheep and New Zealand animals are currently being genotyped (McEwan 2009). If successful the expectation is that industry release will occur in late 2009. Delivery to the New Zealand sheep industry will be in the form of breeding values using a blend of DNA and existing SIL information.

EXTENSION

The existence of genetic evaluation systems does not in itself ensure genetic progress. Information must be used when making selection decisions for genetic progress to be achieved. During the initial years of SIL much effort was devoted to extension with SIL breeder clients. This work aimed to show how SIL features could provide robust information on which to base valid selection decisions. Industry extension continues to be a major focus of SIL although increasingly effort is devoted to commercial ram buyers (Young & Wakelin 2009).

GENETIC GAINS

SIL's success in the decade since its commencement is clearly illustrated by increases in rates of genetic gain achieved (Young and Amer 2009). After the introduction of SIL in 1999, genetic gains almost doubled for both Dual Purpose and Terminal Sire flocks. A further lift occurred after 2004, when SIL-ACE was introduced, with rates of genetic gain close to three times that achieved prior to SIL.

Table 1. Annual rates of genetic progress achieved in the New Zealand ram breeding industry.

Period	Genetic Gain (Index $/year)	
	Terminal Sire	Dual Purpose
1990-1994	0.37	0.25
1995-1998	0.34	0.29
1999-2003	0.61	0.59
2004-2006	0.83	0.92

CONCLUSIONS

Increases in genetic gain since the introduction of SIL can be attributed to a number of factors with SIL responsible for or acting as the focus for these. SIL has enabled breeders to confidently identify superior genetics within farm, within breeding groups, within the national flock and even

between countries for some breeds. However a large number of flocks perform well below the rates of genetic progress achieved by other flocks, demonstrating there are substantial opportunities for further improvement.

SIL has an important role as a vehicle for delivering science to the sheep industry and works closely with researchers to enable scientific advances to be rapidly available. As ram breeders and commercial farmers face increasing pressure from changes in the market such as increased emphasis on product quality and on production efficiency, SIL's role will be critical in delivering appropriate solutions.

ACKNOWLEDGEMENTS

Meat & Wool New Zealand for the significant support it provides to maintenance and development of the SIL System. New Zealand sheep breeders for their support of SIL and their willingness to provide access to information from their flocks which provide the basis for on-going improvements to the SIL system.

REFERENCES

Amer, P.A. and Bodin, L. (2006) *Proc. N.Z. Soc. Anim. Prod*. **66**:429

Callow, C. (1985) *Proc. N.Z. Soc. Anim. Prod*. **45:**1

Crawford, A.M., Anderson, R.M. and McEwan, K.M. *In* McEwan, J.C. (2007) *Proc. N.Z. Soc. Anim. Prod*. **67:**168

Dodds, K.G., Amer, P.A., Spelman, R.J., Archer, J.A. and Auvray, B. *In* McEwan, J.C. (2007) *Proc. N.Z. Soc. Anim. Prod*. **67:**162

Geenty, K.G. (2000) *Proc. N.Z. Soc. Anim. Prod*. **60:**174.

Jopson, N.B., McEwan, J.C., Logan, C.M. and Muir, P.D. (2009) *Proc. N.Z. Soc. Anim. Prod*. **69:** 215

Jopson, N.B., Newman, S-A.N. and McEwan, J.C. *In* Newman, S-A.N. (2009) *Proc. N.Z. Soc. Anim. Prod*. **69:**161

Kerslake, J.I., Everett-Hincks, J.M. and Campbell, A.W. (2005) *Proc. N.Z. Soc. Anim. Prod*. **65:** 13

McEwan, J.C. *In* Newman, S-A.N. (2009) *Proc. N.Z. Soc. Anim. Prod*. **69:**165

McEwan, J.C., Mason, P., Baker, R.L., Clarke, J.N., Hickey, S.M. and Turner, K. (1992) *Proc. N.Z. Soc. Anim. Prod*. **52**:53

Morris, C.A., Bisset, S.A., Vlassoff, A., Mackay, A.D., Betteridge, K., Alderton, M.J., West, C.J. and Devantier, B.P. (2001) *Proc. N.Z. Soc. Anim. Prod*. **61:**92

Morris, C.A., Towers, N.R., Wesselink, C. and Wheeler, M. (1994) *Proc. N.Z. Soc. Anim. Prod*. **54:**263

Newman, S-A.N., Dodds, K.G., Clarke, J.N., Garrick, D.J. and McEwan, J.C. (2000) *Proc. N.Z. Soc. Anim. Prod*. **60:**195

Scobie, D.R., Bray, A.R. and O'Connell, D. (1997) *Proc. N.Z. Soc. Anim. Prod*. **57:**84

Sise, J.A., Auvray, B., Dodds, K.G. and Amer, P.A. (2008) *Proc. N.Z. Soc. Anim. Prod*. **68:**33

Sise, J.A., Shackell, G.H., Byrne, T.J., Glennie, S.F. and Amer, P.A. (2009) *Proc. N.Z. Soc. Anim. Prod*. **69**:223

Young, M.J. and Amer, P.A. (2009) *Proc. Assoc. Advmt. Anim. Breed. Genet.**18**:*422.

Young, M.J. and Newman, S-A.N. *In* Newman, S-A.N. (2009) *Proc. N.Z. Soc. Anim. Prod.* **69:** 153

Young, M.J. and Wakelin, R. *In* Newman, S-A.N. (2009) *Proc. N.Z. Soc. Anim. Prod*. **69:**148

MAPPING OF QUANTITATIVE TRAIT LOCI (QTL) FOR MUSCULARITY IN BEEF CATTLE

I. Novianti, W.S. Pitchford and C.D.K. Bottema

Cooperative Research Centre for Beef Genetic Technologies
School of Agriculture, Food and Wine, The University of Adelaide,
Roseworthy SA 5371 Australia

SUMMARY

Muscularity is a potential indicator for the selection of more productive cattle. Mapping quantitative trait loci (QTL) for traits related to muscularity is useful to identify the genomic regions where the genes affecting muscularity reside. QTL analysis from a Limousin-Jersey double backcross herd was conducted using QTL Express software with cohort and breed as the fixed effects. Nine QTL suggested to have an association with muscularity were identified on cattle chromosomes BTA 1, 2, 3, 4, 5, 8, 12, 14 and 17. The myostatin gene is located at the centromeric end of chromosome 2 and not surprisingly, the Limousin *myostatin* F94L variant accounted for the QTL on BTA2. However, when the *myostatin* F94L genotype was included as an additional fixed effect, the QTL on BTA17 was also no longer significant. This suggests that there may be gene(s) that have epistatic effects with *myostatin* located on cattle chromosome 17.

INTRODUCTION

Muscularity can be defined as "the thickness of muscle relative to the dimensions of the skeleton" (Boer *et al.,* 1974). To select for muscularity, estimated breeding values for loin eye muscle area (EMA) adjusted to a 300kg carcass can be used (Graser *et al*. 2005). Eye muscle area is also used in calculating retail beef yield (RBY). EMA and RBY are relatively highly heritable (Koots *et al*. 1994) and estimated breeding values for these traits can be applied in selection programs to breed cattle for specific market requirements. However, many carcass traits that are commonly selected, such as hot standard carcass weight, are not sufficient to describe the ability of the animals to produce meat since these traits are also affected by the overall size of the animal. Therefore, other muscularity traits, such as meat percentage (retail beef yield) and meat to bone ratio, would be better descriptors. Unfortunately, such traits are not usually recorded and only a limited number of quantitative trait loci (QTL) for these muscularity traits have been mapped. Finding QTL is necessary in order to identify the regions of the genome that may contain genes affecting the traits of interest.

One gene known to have a significant role in muscle development is *myostatin.* McPherron *et al*. (1997) determined the biological function of *myostatin* by knocking out the gene in mice and demonstrating that the mutant mice were larger than the wild type mice as a result of increased muscle mass. The results proved that *myostatin* has an important role in skeletal muscle development by inhibiting muscle overgrowth.

Studies have also reported that there are many other proteins involved in the *myostatin* regulation pathway of muscle development (McPherron *et al.* 1997; Hill *et al.* 2002; Hill *et al.* 2003; Lee 2004; Dominique and Gerard 2006). Therefore, there are likely to be other genes that interact with *myostatin*. The objectives of this project were to identify QTL for muscularity and related carcass traits and to determine if there are QTL that may be epistatic with the *myostatin* gene.

MATERIALS AND METHODS

Materials. Genotype and phenotype data from the JS Davies cattle gene mapping project were used for this study. Two breeds of cattle were used for this project, Limousin and Jersey. The two breeds (Jersey and Limousin) were chosen in the project to maximise the trait variation in the progeny from their crosses. Limousin is a beef breed of a moderate to large frame, while Jersey is a small frame dairy breed. Limousin cattle have the F94L *myostatin* genotype which affects retail beef yield (Sellick *et al.* 2007).

The first phase on this study was conducted in 1993 by mating 280 purebred Jersey and Limousin cows to produce the first cross progeny, namely Limousin x Jersey F_1, which were born in 1994 and 1995. In the second phase, three Limousin x Jersey F_1 sires were mated to the pure Jersey and Limousin dams in Australia and New Zealand (NZ) to produce double backcross animals, called Limousin cross progeny and Jersey cross progeny herein (Sellick *et al.* 2007). There were 161 Limousin cross progeny and 205 Jersey cross progeny born in Australia.

The phenotypic traits that were used for this study were hot standard carcass weight (HSCW), meat weight, meat percentage, bone weight, bone percentage and meat to bone ratio. All traits except HSCW were estimated using regression equations from previous bone-out trials based on HSCW, fat depth, loin eye muscle area, the weight of 2-3 cuts and 2-3 bones with the protocol differing slightly for each cohort as described in Esmailizadeh *et al. (*2008*)*. This study used the genotype data from 150 microsatellite markers in the 3 F_1 sires and their progeny.

Mapping QTL. QTL Express software (http://qtl.cap.ed.ac.uk/) was used to map the QTL by regression analysis of phenotypes (HSCW, meat weight, meat percentage, meat to bone ratio) and genotypes obtained from all the backcross progeny. The software is suitable for half-sib outbred populations and F_2 populations (both inbred and outbred crosses) (Seaton *et al.* 2002). A multiple marker approach for interval mapping in the half sib families was used as described by Knott *et al.* (1996) and completed at 1 cM intervals along the chromosome. Based on Knott *et al*. (1996), three steps were applied. Firstly, informative marker alleles from the sires (361, 368 and 398) were identified to determine which allele the progeny inherited (there were 366 progeny in total) so that the sire gametes for the markers could be re-formed. On average, the sires were informative for 189 loci (Esmailizadeh 2006). Secondly, probabilities of the individual progeny inheriting either allele 1 or 2 from the sires were calculated. Then, these probabilities were combined and provided coefficients on which the phenotypic data can be regressed. Cohort (six levels), breed (Limousin cross and Jersey cross), with and without *myostatin* F94L genotype (CC, CA, AA) were included as fixed effects and were nested within the sire. Three covariates were used: HSCW as a covariate for meat weight, bone weight as a covariate for meat weight and bone percentage as a covariate for meat percentage. Significant QTL were defined by selecting the QTL maxima with F-values greater than 4 as the threshold for the 3 sire families (Lander and Kruglyak 1995). F-values greater than 4 represent $P<0.05$ with 3 degree of freedom (for the 3 sire families).

RESULTS AND DISCUSSION

QTL for all the traits (HSCW, meat weight, meat percentage and meat to bone ratio) were detected on BTA 1, 2, 3 ,4, 5, 8, 12, 14 and 17 (Table 1). There were 4 QTL for HSCW, 3 QTL for meat to bone ratio, 4 QTL for meat weight with HSCW as a covariate, 3 QTL for meat weight with bone weight as a covariate, and 3 QTL for meat percentage with bone percentage as a covariate. Of these, 1 QTL was in common for all the traits on BTA 17. All traits except HSCW also had major QTL on BTA 2. The QTL for meat percent and meat-to-bone ratio are of particular interest as they may represent genes that specifically control muscle mass rather than just increased growth.

Table 1. Significant QTL for muscularity related carcass traits with cohort and breed as fixed effects

BTA	Traits	F-value				QTL Location (cM)			
		Nocov	HSCW cov	Bone wt cov	Bone % cov	Nocov	HSCW cov	Bone wt cov	Bone % cov
1	HSCW	4.6				87			
1	MeatWt	4.28				98			
2	MeatWt	6.08	17.27	10.96		6	6	8	
2	Meat%	20.2			17.31	6			5
2	Mttobn	9.24				8			
3	MeatWt		4.06			-	100		
3	Meat%	4.12			4.45	100			100
4	MeatWt		4.28				37		
5	HSCW	6.08				41			
5	MeatWt	4.4				32			
8	MeatWt	5.11		4.19		57		17	
12	Mttobn	4.23				31			
14	HSCW	6.74				36			
14	MeatWt	5.39				35			
17	HSCW	4.09				85			
17	MeatWt		4.87	4.84			37	82	
17	Meat%	6.07			5.11	38			38
17	Mttobn	4.42				82			

Nocov = no covariate, Hscwcov=hot standard carcass weight as covariate, Bone wt cov=bone weight as covariate, Bone % cov=bone percentage as covariate, Hscw=hot standard carcass weight, Meatwt=meat weight, Meat%=meat percentage, Mttobn=meat to bone ratio. Column with shade represent traits that were not analysed using specified covariate. Only significant results are noted

In order to confirm the identified QTL, a second QTL analysis was conducted which included the Limousin *myostatin* F94L genotype as a fixed effect. This QTL analysis could thus identify other chromosomal regions that might contain gene(s) that interact with *myostatin*. The QTL on BTA 1, 3, 5 and 14 were not affected by the inclusion of *myostatin* F94L genotype as a fixed effect. Since the level of significance and the location of the QTL did not change, this suggests that there are genes in these regions which control muscularity but act independently of *myostatin.* There were minor effects for the QTL on BTA 4 and 8 as the F-value slightly decreased (Table 2).

On the other hand, there were major effects of the *myostatin* genotype detected for the QTL on BTA 2 and 17. The results for BTA 2 verified that the *myostatin* F94L genotype accounted for the QTL on BTA 2. Interestingly, the meat percent QTL on BTA 17 also disappeared with the inclusion of the *myostatin* F94L genotype in the model. The F-values for the other QTL on BTA17 also decreased, although not as dramatically.

A test of co-linearity between genotypes on BTA17 and myostatin was conducted to test whether this disappearance was a random effect. Probabilities of 0.59 for the overall alleles and 0.99 for the sire alleles were found. Thus, the *myostatin* allele and marker alleles were not correlated, implying that the QTL disappearance on BTA 17 is likely to be a consequence of an epistatic effect with *myostatin.*

Table 2. Changes in the QTL level of significance with *myostatin* F94L genotype fitted as a fixed effect with cohort and breed.

BTA	Traits	F-value			
		Nocov	Hscwcov	Bnwtcov	Bn%cov
2	MeatWt	3.07	2.31	2.39	
2	Meat%	2.72			2.44
2	Mttobn	1.44			
4	MeatWt		3.49		
8	MeatWt	3.66		3.92	
17	Hscw	3.66			
17	MeatWt		2.05	3.92	
17	Meat%	2.92			2.65
17	Mttobn	3.63			

Nocov = no covariate, Hscwcov=hot standard carcass weight as covariate, Bnwt=bone weight as covariate, Bn%cov=bone percentage as covariate, Hscw=hot standard carcass weight, Meatwt=meat weight, Meat%=meat percentage, Mttobn=meat to bone ratio. Column with shade represent traits that were not analysed using specified covariate

CONCLUSION

QTL for carcass traits related to muscularity were detected on chromosome 1, 2, 3, 4, 5, 8, 12, 14 and 17. The QTL found on BTA 2 and 17 affected the most traits of interest. The QTL on BTA 2 and 17 were no longer significant when the *myostatin* F94L genotype was included in the model. The QTL affected by the *myostatin* genotype on BTA2 were for meat weight, meat percent and meat-to-bone ratio, while the main QTL on BTA17 affected by *myostatin* was for meat percent. The results for BTA 2 verified that the *myostatin* F94L genotype accounted for the QTL on BTA 2, while the results on BTA 17 suggest that there may be gene(s) that interact or have an epistatic effect with *myostatin* for muscling on this chromosome.

REFERENCES

Boer, H.D., Dumont, B.L., Pomeroy, R.W. and Weniger, J.H. (1974). *Livest. Prod. Sci.* **1**:151.

Dominique, J.E. and Gerard, C. (2006) *Exp. Cell.* Res **312**:2401.

Esmailizadeh, A.K (2006) PhD Thesis. University of Adelaide.

Esmailizadeh, A.K., Bottema, C.D.K., Sellick, G.S., Verbyla, A.P., Morris, C.A., Cullen, N.G. and Pitchford, W.S. (2008). *J. Anim. Sci.* **86**:1038

Graser, H.U., Tier, B., Johnston, D.J. and Barwick, S.A. (2005) *Aust.J.Exp.Agric.* **45**:913

Hill, J. J., Davies, M.V., Pearson, J. H. Wang,J.H., Hewick,R.M., Wolfman, N.M. and Qiu, Y.C. (2002) *J. Biol. Chem*. **277**:40735.

Hill, J. J., Qiu, Y.C, Hewick,R.M. and Wolfman, N.M. (2003) *Mol. Endocrinol* **17**:1144.

Knott, S.A., Elsen, J.M. and Haley, C.S. (1996) *Theor. Appl. Genet.* **93:**71

Koots, K. R., Gibson, J.P., Smith, C. and Wilton, J.W. (1994) *Animal Breeding Abstracts* **62**:309.

Lander, E.S. and Kruglyak L. (1995) *Nat. Genet*.. **11:**241.

Lee, S. J. (2004) *Annu. Rev. Cell Dev. Biol*. **20**:61.

McPherron, A. C., Lawler, A.M. and Lee, S. J. (1997). *Nature* **387**(6628):83.

Seaton G., Haley C.S, Knott S.A., Kearsey M. and Visscher P.M. (2002). *Bioinf* **18:**339.

Sellick, G. S., Pitchford, W. S., Morris, C. A., Cullen, N. G., Crawford, A. M., Raadsma, H. W. and Bottema, C. D. K. (2007). *Anim. Genet.* **38**:440.

GENETIC RELATIONSHIPS AMONG LAMB SURVIVAL, BIRTH COAT SCORE, BIRTH WEIGHT AND 42-DAY BODY WEIGHT IN A SOUTH AFRICAN FINE WOOL MERINO STUD

W.J. Olivier[1,2], S.W.P. Cloete[2,3] and A.C. Greyling[4]

[1]Grootfontein ADI, Private Bag X529, Middelburg (EC) 5900, South Africa
[2]Department of Animal Sciences, University of Stellenbosch, Stellenbosch 7602, South Africa
[3]Institute for Animal Production: Elsenburg, Private Bag X1, Elsenburg 7607, South Africa
[4]Cradock Experimental Station, PO Box 284, Cradock 5880, South Africa

SUMMARY

Lamb and mutton are the most important sources of income for South African Merino farmers. The possibility of genetically improving survival of a Merino lambs was assessed with data from 5769 lambs that were born alive in the Cradock fine wool stud from 1998 to 2003. Three traits can possibly be linked to lamb survival and can therefore be included in selection programs, namely birth coat score, birth weight and 42-day body weight. The aim of this study is therefore to quantify the genetic relationship among lamb survival, birth coat score, birth weight and 42-day body weight in a South African fine wool Merino stud. Direct h^2 estimates were 0.19 for birth weight, 0.19 for birth coat score, 0.06 for 42-day body weight and 0.09 for lamb survival. Corresponding estimates for m^2 were 0.34, 0.18, 0.37 and 0.26 respectively, while c^2 was estimated at respectively 0.06, 0.03, 0.02 and 0.01. Genetic correlations among treats were generally below double the corresponding standard error, and therefore not significant. It was concluded from the results of this study that it would be possible to improve lamb survival genetically, by selection for a related trait, like ewe rearing ability. The scope for successful indirect selection using birth weight, birth coat score or 14-days body weight seems to be limited.

INTRODUCTION

The production of meat is the most important source of income for South African Merino farmers. An increase in meat production can be obtained through selection for increased growth rate or an increase in the number of lambs that survived until weaning as well as a reduction in slaughter age. Increasing the number of lambs that survive until weaning can have a quick and immediate effect on the efficiency of the enterprise.

The survival of Merino lambs is linked directly to its genetic makeup and management factors. Management problems can be rectified with immediate effect, however, the genetic makeup of the lambs is also of utmost importance. Changing the genetic makeup of the lambs will lead to more viable lambs being born.

Three traits can possibly be linked to the postnatal survival of Merino lambs and can therefore be included in selection programs, namely birth coat score (Alexander 1964), birth weight (Morris *et al.* 2000) and 42-day body weight (as an indication of milk production; Brand and Franck 2000). The genetic relationship among these traits and lamb survival must be quantified before they can be included in selection programs. The aim of this study is therefore to quantify the genetic relationship among lamb survival, birth coat score, birth weight and 42-day body weight in a South African fine wool Merino stud.

MATERIALS AND METHODS

Data. The Cradock Fine Wool Merino Stud was established in 1988 as described by Olivier *et al.* (2006). Ewes were bought from 30 Merino farmers with the finest clips throughout South Africa

and four fine wool rams were imported from Australia as sires. Since then, another seven rams were introduced into the stud to be used as sires. Data collected on 5769 ram and ewe lambs that were born alive within this stud from 1988 to 2003 were used for the analyses.

The traits included in the analysis were birth weight, birth coat score, 42-day body weight and lamb survival. Lamb survival was defined as the number of lambs born alive that survived until weaning. This trait was coded as a binary trait with two categories, namely lambs born alive that died before weaning (coded as 1) and lambs that survived until weaning (coded as 2).

Birth coat scores were recorded since 1992 and assessed on a scale from 1 to 4 with 1 being woolly and 4 being hairy. All birth weights were recorded within 24 h of birth. The 42-day body weight was measured at an average age (± s.d.) of 47 ± 6 days.

Data analysis. The means, standard deviations, coefficients of variations, minima and maxima for the respective traits were obtained with the PROC MEANS-procedure of SAS, and significance levels for the fixed effects were obtained with the PDIFF-option under the PROC GLM-procedure of SAS (Littell et al., 2002). The effects tested included year of birth, sex, age of dam in years and birth status. The age of the animals (linear regression) at 42 days of age was also tested for significance. Only effects that had a significant effect were included in the final model for each trait.

The estimation of variance components and genetic parameters was done with THRGIBBSF90 (Misztal *et al.* 2002). This software can be used to estimate variance components and genetic parameters in linear-threshold mixed animal models for any combination of categorical and continuous traits (Lee *et al.* 2002). POSTGIBBSF90 was used for Post Gibbs analysis to obtain solutions for the random effects (Misztal *et al.* 2002). A single chain of 150 000 cycles were run and the first 50 000 cycles used as the burn-in period. Every 10^{th} sample after the burn-in period was stored, giving a total of 10,000 samples for the computation of posterior means and posterior standard deviations.

RESULTS AND DISCUSSION

Descriptive statistics for the respective traits are summarised in Table 1. It is evident from these that the average birth weight of the lambs that were born alive was 4.47 kg, with a range from 1.10 kg to 8.00 kg. The average 42-day body weight was 16.09 kg and it ranged from 4.40 kg to 30.20 kg. The average birth coat scores (1 being more woolly and 4 being more hairy) and lamb survival (1 for lambs that died before weaning and 2 for lambs that survived until weaning) on the underlying scale were 1.92 and 1.91 respectively. This can be translated into the fact that more than 90 % (5247 lambs) of the lambs born alive survived until weaning.

Table 1. Descriptive statistics for birth weight, birth coat score, 42-day body weight and lamb survival

	Birth weight (kg)	Birth coat score	42-day body weight (kg)	Lamb survival
Number of records	5769	4456	5352	5769
Mean	4.47	1.92	16.09	1.91
Standard deviation	0.87	0.81	3.69	0.29
Coefficient of variation	19.50	42.12	22.94	15.01
Minimum	1.10	1	4.40	1
Maximum	8.00	4	30.20	2

The point estimates for direct heritability (h^2), maternal heritability (m^2), maternal permanent environment effect (c^2) and the correlation between the direct and maternal genetic effects (r_{am}) on the underlying scale for birth coat score and lamb survival, as well as for birth weight and 42-day body weight are presented in Table 2.

The heritability for lamb survival estimated in this study (0.09) falls within the range of the values cited by Safari *et al.* (2005) that ranged from 0.00 (Olivier *et al.* 1998 – threshold model) to 0.11 (Hall *et al.* 1995 – threshold model). However, Cloete *el al.* (2009) reported a higher heritability of 0.27 for lamb survival obtained with THRGIBBSF90 (Misztal *et al*. 2002). The maternal heritability (0.26) estimated in this study is higher than the estimate (0.14) obtained by Cloete *et al.* (2009), as well as the range of values reported by Safari *et al.* (2005) from the literature. The heritability, maternal heritability, maternal permanent environment effect and the correlation between direct and maternal genetic effect for birth weight falls within the range of values cited for Merino sheep by Safari *et al.* (2005) from the literature. The values estimated in this study for heritability, maternal permanent environment effect and the correlation between direct and maternal genetic effect for 42-body weight is lower than the values reported by Safari *et al.* (2005) from the literature, while the maternal heritability is higher than the reported values (Safari *et al.* 2005). The heritability for birth coat score is lower than values reported in the literature that range from 0.65 (Kemper *at al.* 2003) to 0.70 (Cloete *et al.* 2003). The permanent maternal environment effect estimated in this study is in the same order as the value reported by Cloete *et al.* (2003).

Table 2. The direct heritability (h^2), maternal heritability (m^2), maternal permanent environment effect (c^2) and correlation between direct and maternal genetic effects (r_{am}) for the different traits (± s.e.)

	h^2	c^2	m^2	r_{am}
Birth weight	0.19 ± 0.05	0.06 ± 0.02	0.34 ± 0.08	-0.26 ± 0.12
Birth coat score	0.19 ± 0.02	0.03 ± 0.01	0.18 ± 0.07	-
42-day body weight	0.06 ± 0.06	0.02 ± 0.01	0.37 ± 0.13	-0.04 ± 0.06
Lamb Survival	0.09 ± 0.05	0.01 ± 0.00	0.26 ± 0.16	-0.14 ± 0.10

Genetic and maternal correlations among the different traits are summarised in Table 3. It is evident that all the correlations were low to moderate among the respective traits, with only the maternal correlation between birth weight and 42-day body weight reaching a level of double the corresponding standard error. Cloete *et al*. (2009) correspondingly reported a low negative genetic correlation between birth weight and lamb survival. Sawalha *et al.* (2007) also reported an unfavourable genetic correlation between lamb viability (coded as 0 for survivors and 1 for animals that had died) and birth weight of 0.21. The relationship between lamb survival and birth weight is complicated by the non-linear relationship that exists between these two traits. This relationship suggest that it would be better to produce lambs with intermediate birth weights, as the extreme to both sides, i.e. too low or too heavy birth weights, will reduce lamb survival. Small and ill thrifty lambs will most probably die due to starvation and hypothermia, whereas dystocia is considered as the biggest problem in lambs that are too big and heavy.

The low genetic correlation between 42-day body weight and lamb survival suggest the improvement in the growth rate of lambs would not have a marked effect on the number of lambs that survived until weaning. The effect of the ewe is bigger than the lamb's own performance at this stage of its life. It might be a more viable option to select ewes with better mothering ability at this stage, because lambs still largely depend on the milk production of the ewes. The genetic

correlation between lamb survival and birth coat score in this study, support findings in the literature that lamb survival is not highly related to the birth coat score of Merino lambs (Ponzoni *et al.* 1996; Cloete *et al.* 2003).

The low heritabilities for lamb survival in the literature (Olivier *et al.* 1998; Snyman *et al.* 1998; Morris *et al.* 2000) suggested that it would not be possible to improve lamb survival genetically. In contrast, the present study, as well as that of Cloete *et al*. (2009), suggests that it would be feasible to improve lamb survival genetically.

Table 3. The genetic (above diagonal) and maternal (below diagonal) correlations among the different traits (± s.e.)

	Birth weight	Birth coat score	42-day body weight	Lamb survival
Genetic correlations				
Birth weight		0.20 ± 0.26	0.44 ± 0.33	-0.31 ± 0.19
Birth coat score	0.04 ± 0.07		0.35 ± 0.76	0.19 ± 0.38
42-day body weight	0.17 ± 0.07	0.02 ± 0.05		0.11 ± 0.23
Lamb survival	0.03 ± 0.05	0.01 ± 0.05	0.05 ± 0.05	

CONCLUSIONS

It can be concluded that it would be possible to improve lamb survival genetically. This can be achieved by selection for traits directly related to lamb survival in ewes (i.e. rearing ability or multiple rearing ability) by culling of unproductive ewes that failed to rear lambs. The scope for indirectly selecting for lamb survival by considering birth weight, birth coat score or 42-days weight appears to be limited.

REFERENCES

Alexander, G. (1964) *Proc. Aust. Soc. Anim. Prod.* **5**:113.

Brand, T.S. and Franck, F. (2000) *Small Rumin. Res.* **37**:85.

Cloete, S.W.P., Misztal, I. and Olivier, J.J. (2009) *J. Anim. Sci.* **87**:2196

Cloete, S.W.P., Olivier, J.J., Van Wyk, J.B., Erasmus, G.J. and Schoeman, S.J. (2003) *S. Afr. J. Anim. Sci.* **33**:248.

Hall, D.G., Fogarty, N.M. and Gilmour, A.R. (1995) *Aust. J. Exp. Agric.* **35**:63

Kemper, K.E., Smith, J.L. and Purvis, I.W. (2003) *Proc. Assoc. Advmnt Anim. Breed. Genet.* **15**:139.

Lee, D., Misztal, I., Bertrand, J.K. and Rekaya, R. (2002) *J. Appl. Gen.* **43**:209.

Littell, R.C., Freud, R.J. and Struop, W.W. (2002) SAS-system for linear models, 4th Ed. SAS Institute. Inc. Cary, N.C., USA.

Misztal, I., Tsuruta, S., Strabel, T., Auvray, B., Druet, T. and Lee, D. H. (2002) *Proc. 7th World Congr. Gen. Appl. Livest. Prod.* **33**: 743. Montpellier, France.

Morris, C.A., Hickey, S.M. and Clarke, J.N. (2000) *N. Z. J. Agric. Res.* **43**:515.

Ponzoni, R.W., Grimson, R.J., Jaensch, K.S., Smith, D.H. and Hynd, P.I. (1996) SARDI Research Report Series, No 11. pp. 44.

Olivier, W.J., Olivier, J.J., Cloete, S.W.P. and Van Wyk, J.B. (2006). Proc. 8th *World Congr. Gen. Appl. Livest. Prod.*, Belo Horizonte, 13-18 Augustus, 84.

Olivier, W.J., Snyman, M.A., Van Wyk, J.B. and Erasmus, G.J. (1998) *Livest. Prod. Sci.* **56**:71.

Safari, E., Fogarty, N.M. and Gilmour, A.R. (2005) *Livest. Prod. Sci.* **92**:271

Sawahla, R.M., Conington, J., Brothersone, S. and Villanueva, B. (2007) *Anim.* **1**:151.

Snyman, M.A., Erasmus, G.J. and Van Wyk, J.B. (1998) *S. Afr. J. Anim. Sci.* **28**:120.

SHEEP SELECTED FOR RESISTANCE TO FACIAL ECZEMA DISEASE ALSO SHOW HIGHER TOLERANCE TO ACETAMINOPHEN CHALLENGE

S. H. Phua[1], P. Johnstone[1], H. Henry[1], A. Findlay[1] and C. A. Morris[2]

[1]AgResearch Invermay, Private Bag 50034, Mosgiel 9053, New Zealand
[2]AgResearch Ruakura, Private Bag 3123, Hamilton 3240, New Zealand

SUMMARY

Facial eczema disease (FE) is a hepato-mycotoxicosis caused by sporidesmin, leading to secondary photosensitisation in severely affected sheep and cattle. Two genetic lines of Romney are maintained by selection for resistance and susceptibility to FE. Five groups of 8 animals, each group composed of 4 animals from each line, were artificially challenged with different dose rates (25, 50, 75, 150 & 250 mg/kg liveweight) of acetaminophen (or paracetamol). Weekly blood samples were collected from each animal for 5 weeks after dosing and were used for liver function tests (glutamate dehydrogenase (GDH), gamma-glutamyl transferase (GGT), bilirubin and albumin/globulin ratio) to measure their liver injury. The results showed that FE resistant-line animals were more tolerant of acetaminophen than the susceptible-line animals, and that they also recovered faster after the drug challenge.

INTRODUCTION

There are many genes contributing to FE resistance in sheep (Phua *et al*. 2009). An obvious group of candidate genes would be those encoding the enzymes involved in the liver detoxification pathways. It was therefore not surprising to find that FE-resistant sheep also tended to be resistant to ryegrass staggers disease (Morris *et al*. 1995a); the latter is a neuromuscular in-coordination caused by the neurotoxin lolitrem B which is produced by a ryegrass endophyte. Similar observations were also made in rodents, where genetic lines of mice selected for resistance and susceptibility to tall fescue toxicosis were also found to be associated with greater and lesser resistance to sporidesmin (Hohenboken *et al*. 2000). We extended similar study here to include a potential hepato-toxin in human. Acetaminophen, also commonly known as paracetamol, is a non-prescriptive drug widely used in humans for the relief of fever and pain. Overdose of this drug causes acute liver failure. But the modes of action and target tissues within liver are different between acetaminophen and sporidesmin. In this experiment, we wanted to assess the acetaminophen-induced liver responses of sheep selected for resistance and susceptibility to sporidesmin.

MATERIALS AND METHODS

Animals. The Romney FE resistant (R) and susceptible (S) selection lines were established in 1975. In the early years, selection was made based on liver injury scores, later changed to blood GGT levels and finally to the present log(GGT) breeding values. In general, about 6 rams and 120 breeding ewes are maintained in each line (Morris *et al*. 1995b). The experimental animals used in this study were 2001-born progeny.

Acetaminophen trial. This trial received an ethical approval from the AgResearch-Invermay Animal Ethics Committee. Twenty R and 20 S 14-month ewes were obtained from the FE selection lines; their respective weight ranges were 46-58kg (average 51kg) and 43-59kg (average 50kg). The animals were randomly distributed into 5 groups of 8, having 4 R and 4 S animals per group. Each group was dosed with a different rate of acetaminophen purchased from a hospital

pharmacy: the dose rates were 25, 50, 75, 150 and 250 mg/kg liveweight (LWT). Before dosing (i.e. Week-0), blood samples were collected from the animals to determine their background levels of serum glutamate dehydrogenase (GDH, i.u./l), gamma-glutamyl transferase (GGT, i.u./l), bilirubin (µmol/l), albumin (g/l) and globulin (g/l). After dosing, blood samples were collected weekly for 5 weeks (i.e. Weeks 1-5) for the same assays above. Two animals in the highest dose-rate group died from unknown causes: 1 R animal died 1 week after dosing and 1 S animal died 3 weeks post-dosing.

Statistical analyses. A mixed model was fitted to the natural logarithm of the GGT and GDH data. The fixed effects were the selection lines, dose rates of acetaminophen and the time (i.e. weeks) of observation. The random model included the experimental animals' sires. The errors were assumed to be autoregressive order 1 (AR1). The model was fitted using the REML directive in GenStat 10th Edition (Payne *et al*. 2007).

RESULTS AND DISCUSSION

Overall the dose rates of acetaminophen used in this study did not cause permanent liver damage in sheep. This was shown by the calculated albumin/globulin ratios which were within normal range for all animals during the 5-week trial period. Corollary was that there was no elevation of bilirubin in the blood throughout the trial. However, the dose rates were sufficient to elicit some definitive responses in terms of GGT and GDH data.

A mechanism of sporidesmin toxicity is believed to be through the production of reactive oxygen species, and the target tissue is the biliary tracts. In histological sections of FE-affected livers, the foci of pathological lesions are in the portal triad regions (Ozmen *et al*. 2008). The levels of serum GGT are used to measure the degrees of biliary tract damage (Towers & Stratton 1978). When the selection-line sheep were challenged with acetaminophen, the S animals showed significantly higher GGT levels than the R animals ($P < 0.001$), indicating that they were more susceptible to the drug (Figure 1).

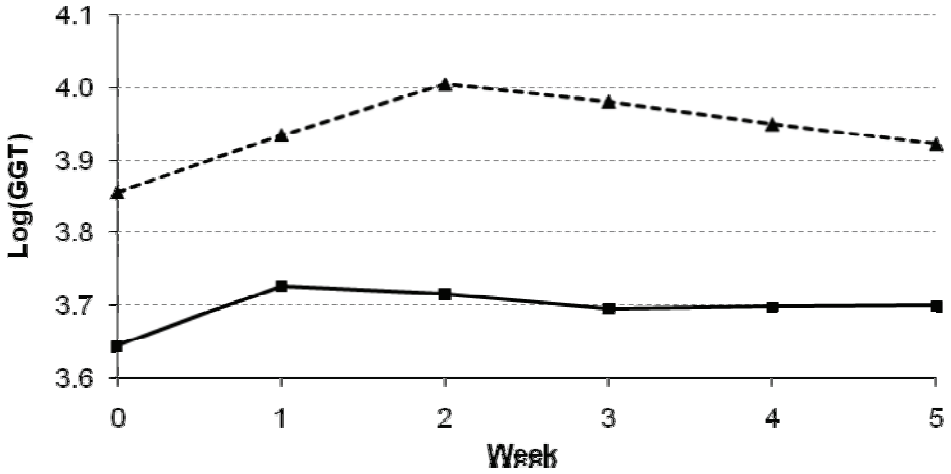


Figure 1. GGT responses of R (squares) and S (triangles) animals over time. Week-0 is the week before acetaminophen dosing, and Weeks 1-5 are 1-5 weeks after dosing. GGT values are averages of all R and S animals from all dose rates. The average standard error of difference (SED) is 0.06.

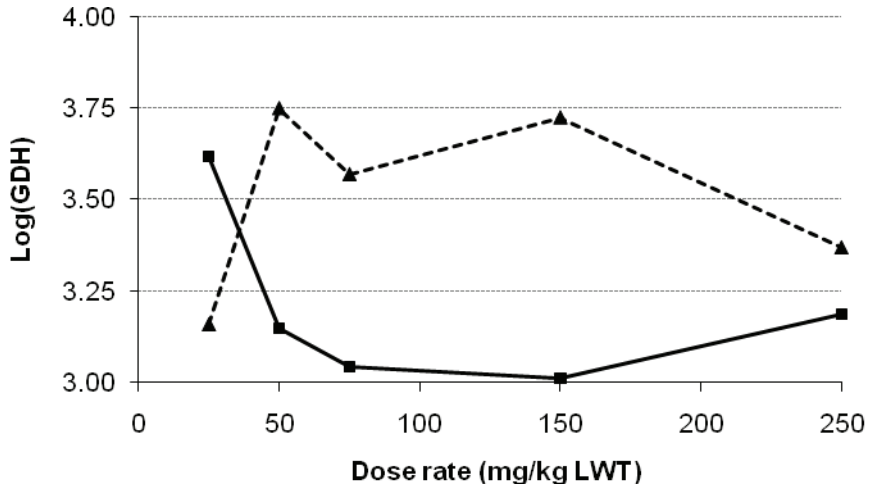


Figure 2. GDH responses of R (squares) and S (triangles) animals under different acetaminophen dose rates. GDH values are averages of R and S animals from all weeks. The average SED is 0.38.

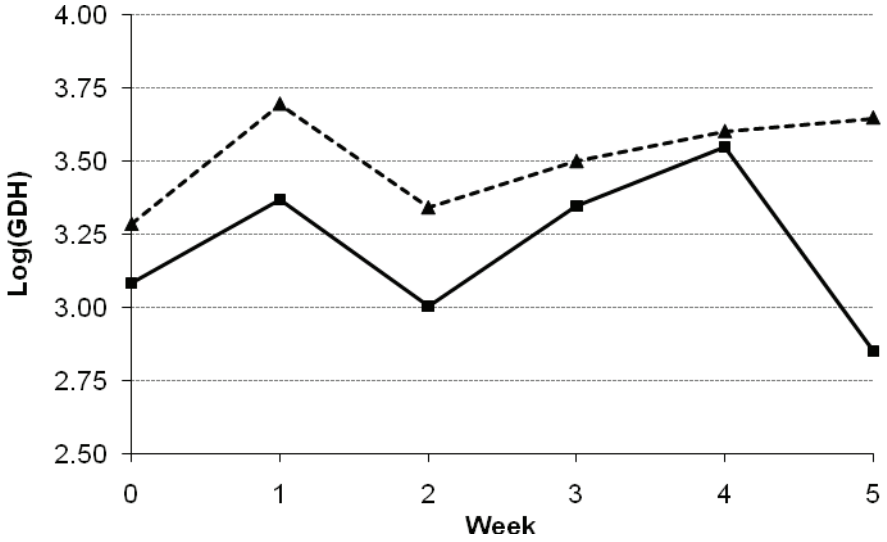


Figure 3. GDH responses of R (squares) and S (triangles) animals over time. Week-0 is the week before acetaminophen dosing, and Weeks 1-5 are 1-5 weeks after dosing. GDH values are averages of all R and S animals from all dose rates. The average SED is 0.20.

Acetaminophen is used in humans as an analgesic and antipyretic drug. At therapeutic dose rates, the compound is glucuronidated and sulphated in the liver for elimination from the body. Its toxicity arises when an overdose of the compound is given: under such conditions, the glucuronidation and sulphation pathways are saturated and the compound is diverted into the

glutathione-conjugation pathway for elimination. In the latter pathway an intermediate metabolite is produced which is toxic to hepatocytes and causes lesions in the centrilobular regions of the liver (Prescott 1996). GDH levels in the blood are used to measure such liver damage. It is the GDH data which showed the most Interesting differences between the R and S sheep.

Forty percent (8 out of 20) of S animals, compared with 10.5% (2 out of 19) of R, showed at least 3 weeks of elevated serum GDH levels. A chi-square test of this comparison showed that the susceptible selection line had a significantly higher proportion of animals with elevated GDH levels than the resistant line ($P < 0.05$). There was a pronounced difference in GDH profiles between the genetic lines with increasing dose rates of acetaminophen challenge (Figure 2): the S animals had lower GDH than the R at 25 mg/kg, and this was reversed at higher dose rates. As depicted in Figure 3, S animals showed significantly higher GDH than the R animals ($P < 0.05$), indicating that they are more susceptible to acetaminophen. Most interestingly, the graph also showed that the R animals recovered faster than the S after the drug challenge.

CONCLUSIONS

The acetaminophen dose rates used in this trial were low enough to not cause any permanent liver damage in sheep, but they were high enough to elicit responses which reflected the genetic make-up of the animals. The FE resistant-line sheep consistently showed significantly lower blood GGT and GDH than the susceptible-line animals under acetaminophen challenge. An inference is that sheep selected for resistance to sporidesmin were also resistant to acetaminophen. An extrapolation from this finding is that FE resistant animals could well show tolerance to many other pasture toxins and drugs. Further, it appeared that in the selection for resistance to FE, we also selected resilient animals which could recover quickly after a xenobiotic challenge.

ACKNOWLEDGMENTS

Funding for this work was provided by the New Zealand Foundation for Research, Science and Technology.

REFERENCES

Hohenboken, W.D., Robertson, J.L., Blodgett, D.J., Morris, C.A. and Towers, N.R. (2000) *Journal of Animal Science* **78**:2157.

Morris, C.A., Towers, N.R., Wheeler, M. and Amyes, N.C. (1995a) *N. Z. Journal of Agricultural Research* **38**:367.

Morris, C.A., Towers, N.R., Wheeler, M. and Wesselink, C. (1995b) *N. Z. Journal of Agricultural Research* **38**:211.

Ozmen, O., Sahinduran, S., Haligur, M. and Albay, M.K. (2008) *Trop. Anim. Health Prod.* **40**:545.

Payne, R.W., Murray, D.A., Harding, S.A., Baird, D.B. and Soutar, D.M. (2007). "GenStat 10th Edition for Windows" VSN International, Hemel Hempstead.

Phua, S.H., Dodds, K.G., Morris, C.A., Henry, H.M., Beattie, A.E., Garmonsway, H.G., Towers, N.R. and Crawford, A.M. (2009) *Animal Genetics* **40**:73.

Prescott, L.F. (1996) In "Paracetamol (Acetaminophen): A critical bibliographic review", p. 285, Taylor & Francis, Florence, Kentucky, U.S.A.

Towers, N.R. and Stratton, G.C. (1978) *N. Z. Veterinary Journal* **26**:109.

MAPPING QTL FOR EARLY GROWTH AND MATERNAL PERFORMANCE IN SHEEP

H.W. Raadsma, E. Jonas, K.R. Zenger, C.A. Cavanagh, M.K. Lam and P.C. Thomson

ReproGen, Faculty of Veterinary Science, The University of Sydney, Camden, Australia

SUMMARY

This paper describes QTL for bodyweight, and growth rate between week 2 after birth up to 43 weeks of age, and maternal performance based on milk yield and milk energy content in Awassi-Merino backcross progeny. QTL were mapped using a population specific framework map with 50 QTL (LOD>2) mapped for body weight and growth rate, and six QTL for milk yield and milk energy content. The majority of QTL for bodyweight and growth showed a significant sex by QTL interaction with wethers showing stronger QTL effects than ewes. Fewer QTL were detected for body weight or growth in the period from 2 to 25 weeks (range 1-4 QTL) whereas for the period from 25-43 weeks and over the entire growth period a greater number of QTL were observed (range 5-8). Implementation of multiple QTL for early growth and maternal performance will be difficult in sheep with need for validation and tracking in each population.

INTRODUCTION

In sheep, growth rate and body mass represent economically important traits, which are under moderate genetic control and respond to directional selection. Furthermore in intensive lamb production, maternal performance is considered a major contributor to early growth of lambs. Despite extensive background information in terms of population specific genetic parameters, relatively few QTL studies have been reported for growth and maternal performance in sheep. In addition many QTL studies have been restricted to partial genome scans, limiting the discovery of and reports on new QTL. In this study we report on QTL for growth from birth till ten months of age, and milk production in ewes as one of the major traits for maternal performance.

MATERIALS AND METHODS

Animals. As described by Raadsma *et al*. (2009), animals from an Awassi x Merino backcross population were generated specifically for QTL mapping across a broad range of phenotypes. In brief, four Awassi (A) sires were crossed with medium and superfine Merino (M) ewes to produce 16 AM F1 families. Four F1 ram lambs were selected to represent each of the founder families, and these were backcrossed to superfine and medium Merino ewes, producing families with approximately 611, 202, 141, and 186 AMM backcross progeny respectively. Additionally 1391 double-backcross and intercross progeny were generated from all F1 sires by mating them to backcross females. Final family sizes were 987, 703, 274 and 567 for the four families. Results are specifically detailed for Family 1 in AMM backcross progeny only

Framework map for QTL mapping. A population specific framework genetic linkage map was constructed after screening a panel of 558 pre-selected polymorphic micro-satellite markers for a genome wide scan, of which a total of 204 markers were informative in the AMM flock using Carthagene software (http://www.inra.fr/bia/T/CartaGene/index.html) with a minimum significance of LOD > 2 for the next best map order (average LOD: 6.92). The average adjacent marker density across the genome was 22.2 cM and predicted genome coverage was approximately 95% for each chromosome.

Traits and Analyses. Non-fasted body-weight measurements were taken at weeks 2, 15, 25, 32, 37, and 43 for up to 510 backcross progeny (246 ewes and 264 wethers) of family 1. All sheep were weaned at 15 weeks of age, and maintained as separate sex cohorts on pasture. Growth rates were divided into four growth phases: week 2 to week 15 (GR 2-15), week 15 to week 25 (GR15-25), week 25 to week 37 (GR25-37), and week 37 to week 43 (GR37-43). In addition growth rate was analysed for each period adjusting for initial body weight at the start of the growth period (GR-adj). Lactation data were obtained from 590 ewes over different lactations. Ewes were milked once daily and milk yields recorded on average every 2-3 days. Total milk yield throughout the lactation was determined by fitting the Wood lactation curve model (Wood 1968) to the milk yield data and standardizing the milk yield till day 100 making adjustments for average age, parity, and milking frequency. Predicted energy take off through the milk was predicted from yield and composition analyses based on fat, protein and lactose composition recorded from weekly samples. From these ewes, genotypes of 170 ewes from family 1 were available among the group of animals with recorded body weight.

Genetic and statistical analyses. QTL-MLE was used as a maximum likelihood based approach specifically for half sib designs in non-inbred strains (Raadsma *et al*. 2009). A chromosome-wide threshold for statistical significance was calculated for each chromosome, based on a LOD score of 2. The model fitted to the phenotypes (y) had the form $y = \mu + Sex + QTL + Sex.QTL + \varepsilon$, where inclusion of the Sex.QTL interaction term facilitated sex-specific QTL to be fitted. A likelihood ratio test was used to assess statistical significance of the interaction term.

RESULTS

Average body weights and growth rates at the four growth phases are shown in Table I as well as the mean lactation performance of ewes milked. Significant sex effects were observed for all measures where applicable with ewes being smaller and showing lower growth rates.

Table 1. Average body weight at weeks 2, 15, 25, 37 and 43 (BWT, kg) with corresponding growth rate (GR, kg/week), and milk yield till day 100 (MY, L) with corresponding energy yield until day 100 (EY, MJ) for N progeny in the QTL mapping data set

Trait	BWT WK 2	BWT WK 15	BWT WK 25	BWT WK 37	BWT WK 43	GR WK 02-15	GR WK 15-25	GR WK 25-37	GR WK 37-43	MY day 100	EY day 100
N	514	406	409	385	385	385	400	382	382	156	139
Mean	6.22	11.52	17.31	19.24	29.07	4.69	5.79	2.14	9.79	73.30	326.9
SD	1.33	2.61	2.86	2.68	3.51	1.60	1.67	1.74	2.00	19.44	87.5
CV	0.21	0.23	0.17	0.14	0.12	0.34	0.29	0.81	0.20	0.27	0.27
Min	2.50	4.80	9.50	10.00	16.50	-0.40	-0.50	-4.00	0.00	26.04	136.2
Max	10.30	24.00	26.00	28.00	40.50	10.00	11.80	13.00	17.50	126.32	545.5

Average body weights at all sequential times showed positive growth. The period from pre weaning (week 2-15) showed similar growth to the period immediately post weaning (week 15-25), but the period between week 25 and 37 was characterised by relatively slow growth coinciding with the winter season, and a large number of lambs showing weight loss. The period of greatest growth was observed from week 37-43 coinciding with spring season and supplementary feeding late winder early spring. The lactation performance of the 170 genotyped ewes milked once a day showed an average (Wood model based) cumulative milk yield till day

100 of 73 L with a range of 26 to 126 L. The predicted energy output corresponding to the lactation was 328 MJ ranging from 136 to 545 MJ over the 100 day lactation period. The energy calculation was based on the milk yield and the milk composition which was on average 5.5% protein, 7.1% fat and 3.6% lactose.

Table 2 shows the QTL from the single QTL analyses for body weights at weeks 2, 15, 25, 37, and 43 and corresponding growth phases, as well as milk yield traits. Four chromosomes showed significant QTL (LOD>3) for body weight and growth on OAR 3, 11, 21 and 25. In addition chromosomes 1, 3, 5, 6, 7, 8, 9, 11, 12, 16, 19, 21, 23, 24, and 25 showed QTL with LOD >2. The majority of QTL for BWT and GR showed a significant sex by QTL interaction. Average QTL effect sizes were 0.35, 0.43, 0.58, 0.48 and 0.33 SD for BWT, and 0.60, 0.01, 0.14, 0.20 SD for growth rate in wethers, whereas corresponding QTL in ewes were 0.1, -0.06, 0.05, 0.13 and 0.08 SD for BWT and -0.01, -0.06, 0.17 and 0.23 SD for GR, respectively. Fewer QTL were detected for body weight or growth in the period from 2 to 25 weeks (range 1-4 QTL) whereas for the period from 25-43 weeks and over the entire growth period a greater number of QTL were observed (range 5-8). QTL for cumulative milk yield till day 100 were detected on OAR 2, 3, 20 and 25. Matching QTL for energy output were only detected for OAR 3 and 20 for corresponding locations to cumulative milk yield on the same chromosomes.

Table 2: Chromosomes (OAR) with significant QTL for bodyweight (BWT), growth (GR) and growth adjusted for starting body weight (GR adj), milk yield and milk energy. Location of QTL position (cM) indicated for each chromosome, QTL in italic suggest significance of LOD scores <1.75<LOD<2.0, whereas those in bold were LOD >3.0 Number (n) of significant QTL (LOD>2.0) for each trait shown in last column.

OAR	1	2	3	4	5	6	7	8	9	11	12	15	16	19	20	21	23	24	25	n
BW2	-	265	-	26	-	-	-	52	-	-	-	*79*	-	-	*99*	-	-	-	-	3
BW15	238	-	49	-	-	-	44	-	*146*	-	-	-	-	-	-	-	-	-	-	3
BW25	-	-	50	-	121	-	-	-	140	-	-	-	-	-	-	-	*25*	85	-	4
BW37	*111*	*265*	51	-	119	65	-	-	144	**58**	-	-	-	-	-	78	41	**87**	-	8
BW43	-	-	**47**	-	-	68	16	75	145	55	-	-	-	-	-	**25**	-	85	-	8
GR2-15	238	-	-	-	-	-	-	-	*146*	-	-	-	-	-	-	-	-	-	-	1
GR15-25	-	-	-	-	-	-	-	-	127	-	-	-	*20*	*20*	*60*	-	-	105	-	2
GR25-37	*310*	-	-	-	-	-	-	-	-	71	-	-	*120*	**60**	-	-	-	-	-	2
GR37-43	209	-	*139*	-	-	-	116	*81*	*154*	-	-	-	-	67	-	**58**	-	-	85	5
GR2-43	118	-	28	-	-	62	-	77	-	55	-	-	-	-	-	26	-	86	-	7
GR2-43 adj	357	-	-	-	-	43	-	-	-	56	104	-	98	-	-	19	-	91	-	7
Milk yield	-	214	96	-	-	-	-	-	-	-	-	-	-	-	-	52	-	-	105	4
Milk Energy	-	-	**94**	-	-	-	-	-	-	-	-	-	-	-	-	50	-	-	-	2

DISCUSSION AND CONCLUSIONS

In this study, we report 50 significant QTL across 17 chromosomes related to body weight and growth rate for animals up to 10 months of age as part of a subset of observations not reported in the study described by Raadsma *et al.* (2009). In addition we report on six QTL for maternal performance as expressed by milk yield and composition. The pattern of growth in this flock is consistent for sheep maintained on semi-improved pasture in a temperate Australian tablelands climate characterized by low pasture availability in the colder winter months, and abundant pasture

growth in spring with a residual pasture carry over in summer and autumn. This was reflected in the growth curves seen here, with a period of arrested growth over winter, followed by a period of rapid, almost compensatory growth in spring. The sheep were weighed immediately off pasture so some of the fluctuations in body weight and growth rates may reflect gut content. Despite these limitations the overall growth as shown across multiple measurements could be estimated as shown over the period shortly after birth till 43 weeks. From this analysis, relatively few QTL were observed for body weight and growth pre weaning, whereas growth post weaning almost revealed twice as many QTL. This study had the advantage of positioning the QTL on a resource specific framework map and covers a full autosomal genome scan as distinct from the previous partial genome scans. The QTL on OAR 1, 2, 8, 20 support previous findings by others who report QTL for similar traits in sheep (Walling *et al*. 2004; McRae *et al*. 2005; Beraldi *et al*. 2007; Hadjipavlou and Bishop 2008). No QTL have previously been reported for milk energy content, but QTL for milk yield on OAR 3 reported in this study supports previous findings reported by Carta *et al*. (2003) and Barillet *et al*. (2005).

Consistent with the extensive genetic variation which has been reported for growth, body size and to a lesser extent maternal performance and milk production, the presence of QTL with relatively large effect were detected. The effects of the QTL detected in this study are relatively large (in the order of 0.4 to 0.7 phenotypic standard deviations) and originated predominantly from the larger-framed Awassi grandsire, as expected with the favourable allele for growth and body size. However, surprisingly for most traits, significant sex by QTL interactions were observed, with most QTL expressed in males but not females. No immediate or obvious explanation can be given for this, since the ewes and wethers were managed together for most of the period.

Utilization of multiple QTL for early growth and maternal performance would be difficult with at least 14 QTL of potential interest which would have to be mapped and tracked simultaneously in commercial populations. The problems of applying a multi QTL selection programme have been detailed by Boichard *et al*. (2006) for dairy cattle. A more amenable approach may be to combine high density marker information in genomic selection strategies.

ACKNOWLEDGMENTS

The authors acknowledge in particular Mr D Palmer, Mrs G Attard and M Jones.

REFERENCES

Barillet, F., Arranz, J.J. and Carta, A. (2005) *Genet. Sel. Evol.* **37**:S109.

Beraldi, D., McRae, A.F., Gratten, J., Slate, J., Visscher, P.M. and Pemberton, J.M. (2007) *Evolution* **61**:1403.

Boichard, D., Fritz, S., Rossignol, M.N., Guillaume, F., Colleau, J.J. and Druet, T. (2006) *Proc. 8th Wrld. Congr. Genet. Appl. Livest. Prod. Brazil* **22**:11.

Carta, A., Barillet, F., Casu, S., Cribiu, E.P., Elsen, J.M., Fraghi, A., Mura, L. and Schibler, L. (2003) *Ital. J. Anim. Sc*. **2**:31.

Hadjipavlou, G. and Bishop, S.C. (2008) *Anim.Genet*. **40**:165.

McRae, A.F., Bishop, S.C., Walling, G.A., Wilson, A.D. and Visscher, P.M. (2005) *Anim. Sci*. **80**:135.

Raadsma, H.W., Thomson, P.C., Zenger, Lam, M.K., Jonas, E., Cavanagh, C., Jones, M., Attard, G., Palmer, D. and Nicholas, F.W. (2009) *Genet. Sel. Evol*. **41**:34.

Walling, G.A., Visscher, P.M., Wilson, A.D, McTeir, B.L., Simm, G. and Bishop, S.C. (2004) *J. Anim. Sci*. **82**:2234.

Wood, P.D.P. (1968) *Nature*. **218**:894.

IMPROVING THE ACCURACY OF SELECTING ANIMALS FOR REDUCED METHANE EMISSIONS

D.L. Robinson

NSW Department of Primary Industries & CRC for Sheep Industry Innovation,
Primary Industries Innovation Centre, University of New England, Armidale, NSW 2351

SUMMARY

Enteric methane emissions of livestock represent 10.3% of Australia's greenhouse gas emissions, so it is important to identify low-emitting animals in order to study the mechanisms that lead to low emissions relative to production. This will require field testing of large numbers of animals to identify those with the lowest emissions, and at the same time generate useful data from which heritabilities, and also genetic and phenotypic correlations with production traits, can be estimated.

The repeatability of methane emissions from a 1-2 hour field test are expected to be much lower than for daily methane production measured in a respiration chamber. Use of multi-stage selection procedures is therefore recommended to increase the accuracy of identifying animals with low methane emissions. For example, if field tests have a repeatability of 0.25 and resources are available for 1,000 tests, the lowest 50 emitters will have a mean of -1.03 standard deviations below the population average if 1,000 animals are tested once, compared to -1.31 using a multi-stage selection system testing 725 animals, re-testing the best 180 and then the best 95. Other multi-stage schemes to select individuals, as well as sires or families, are evaluated and discussed.

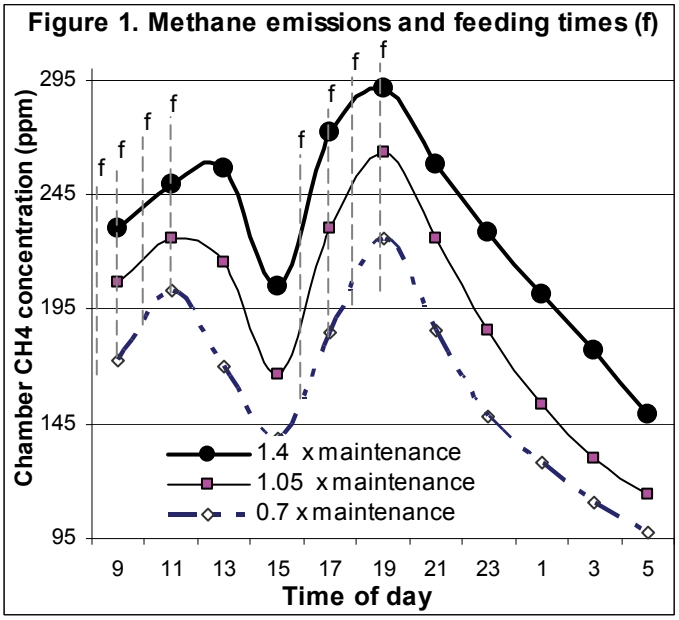
INTRODUCTION

Greenhouse gas emissions from Australian livestock in 2006 amounted to 62.8 million tonnes of CO_2-equivalent. This represents 69.7% of the agricultural sector's emissions and 10.9% of net national emissions (DCC 2008). The vast majority of this (59.3 million tonnes) was enteric methane emissions. It is therefore important to find ways of reducing methane emissions from livestock. A promising research strategy is to identify the lowest emitting individuals, in order to study the mechanisms and identify physiological indicators associated with their low level of emissions relative to their production value. This will require large numbers of animals to be tested in the field, with the data being used to identify extreme animals, as well as provide useful information on phenotypic and genetic correlations between methane emissions and production traits, needed in order to include methane emissions in future breeding objectives.

A review of between- and within- animal variation in daily methane production from respiration chamber measurements suggests that the trait is moderately repeatable for some groups of animals (Robinson, in preparation). However, field measurements of emissions taken over 1 or 2 hours will be much less repeatable (Goopy *et al.* 2009). This paper evaluates the potential increase in accuracy from using sequential selection systems to screen large numbers of animals for traits with low repeatability, using repeat tests on promising individuals to achieve the best possible accuracy of selection.

REPEATABILITY OF FIELD MEASUREMENTS

For methane emissions, a respiration chamber experiment (Goopy *et al.* 2009) with sheep on 3 feeding levels (0.7, 1.05 and 1.4 x maintenance) showed that, as long as animals have been synchronised to the same feeding pattern, 2-hour measurements provide useful estimates of daily methane production. Emissions were strongly dependent on feeding level, with a daily pattern related to time of feeding; in this case four morning (8, 9, 10, 11 am) and four afternoon feeds (4, 5, 6, 7 pm) of equal size (Fig 1).



The results demonstrate the need for an effective protocol to synchronise or adjust for feeding patterns and/or time of day. If this effect is ignored, an animal measured at 7 pm on the lowest feeding level would appear to have higher emissions than when on the highest feeding level, if measured at 3 pm, implying that repeatability of field measurements could be very low.

A range of possible field measurement protocols are currently under consideration, from simple schemes to measure animals in batches of 10-20 and statistically adjust for batch (which will include time of day and feeding effects) to more complicated schemes controlling feeding times. For example, animals could fast overnight until 2 hours before testing the following day, then graze for 1 hour, fast for a second hour, with methane emissions then measured for 1-2 hours in a transportable chamber. It is hoped that the chosen protocol will allow adjustment for feeding and time of day effects, and so achieve repeatabilities between 0.25 and 0.5.

As well as estimating the heritability of methane emissions and genetic correlations with production traits, the aim is to identify animals (and perhaps also sire groups) with low emissions relative to their level of production. Once identified, the extreme animals will be transported to the research facilities to confirm their level of emissions in respiration chambers, and study the mechanisms responsible for favourably low emissions.

Because of the expense and difficulty of identifying these animals, and the expectation that 1-2 hour field measurements will have low repeatability, use of multi-stage selection procedures was considered to improve the accuracy of selecting extreme animals and provide an estimate of the repeatability of field measurements. An additional benefit is that increased accuracy of identifying low methane emitters should also improve knowledge about the phenotypic and genetic relationships between low methane emissions and production traits.

STATISTICAL METHODS

Repeatability is defined as the correlation between one measurement and another recorded under the same conditions. In statistical terms, if **x** is the vector of true measurement (CH_4 emissions of each animal) and **e** the error, observed measurements, **m**, can be expressed as:

m = **x** + **e**, with repeatability, r = Var(**x**)/(Var(**x**)+Var(**e**)) where Var = variance.

In multi-stage selection procedures (Robinson 1984) the best estimate of an animal's phenotypic performance is calculated from all available measurements. In the case of a simple trait with constant Var(**e**), this is the mean of all prior measurements (including the current one). If more extreme than a specified value, the animal is either re-tested or, at the final stage, selected.

The methods of Robinson (1984) were used to evaluate various one, two and three-stage selection strategies for repeatabilities of 0.5, 0.33 and 0.25, assuming that errors, **e**, and true methane measurements, **x**, are both normally distributed.

RESULTS

Table 1 presents results of four strategies to select the lowest emitting 50 animals using 1,000 individual test results. The first strategy is to measure 1,000 animals once. The second is to test half as many animals twice, increasing the accuracy of assessing each individual animal, at the expense of being able to test fewer animals. The first strategy is more accurate for repeatability of 0.5, resulting in a mean of the selected animals of -1.46 s.d. from the population mean (Table 1). In contrast, for less accurate tests (repeatabilities of 0.33 and 0.25) taking repeat measurements on a random sample of 500 animals results in more accurate selection and a lower mean of the selected animals. This, in fact, is the better strategy whenever repeatability is less than 0.45.

An even better strategy is to use multi-stage selection, carrying out a first test, then a second (and if desired a third) test on animals that perform well in previous tests. The improvement from use of multi-stage selection can be substantial. For example, the mean of selecting the lowest 50 from 1,000 animals using a three-stage strategy for a test with a repeatability of 0.33 is -1.45, almost as good as the mean of -1.46 from a single stage selection strategy from a test with only a half the error variance (repeatability of 0.5, Table 1).

Table 1. Mean of 50 selected animals (expressed in standard deviations of the true values, x, from the population mean), based on 1,000 individual test results

	Mean of the 50 selected animals		
Repeatability	0.5	0.33	0.25
1,000 animals tested once	-1.46	-1.19	-1.03
500 animals tested twice	-1.43	-1.24	-1.11
Test 850, then lowest 150	-1.61	-1.38	-1.23
Test 725, then lowest 180, then lowest 95	-1.65	-1. 45	-1.31

Results from strategies to select 100 animals using 1,000 tests (Table 2) show a similar pattern, although the mean of selected animals is not as negative, because more animals are selected. In this scenario, there is little or no benefit from repeat testing all animals, even for low repeatabilities. For repeatability of 0.25, repeat testing 500 animals produces a gain of -0.89, only marginally better than the gain of -0.88 from testing 1,000 animals once. At repeatabilities of 0.33 and 0.5, testing all animals once is better than repeat testing 500 animals. However, there are still gains from multi-stage selection. For repeatabilities of 0.33 and 0.25, the mean of the 100 lowest emitters from a two-stage strategy testing 800 then 200 animals (Table 2) are 12% to 14% lower than for the one-stage strategy. In fact, the means (-1.13 and -1.00 for repeatabilities of 0.33 and 0.25) of the lowest 100 from the two-stage strategy (Table 2) are almost as low as those for the lowest 50 (-1.19 and -1.03) testing 1,000 animals only once (Table 1).

Table 2. Mean of 100 selected animals (expressed in standard deviations of the true values, x, from the population mean), based on 1,000 individual test results

	Mean of the 100 selected animals		
Repeatability	0.5	0.33	0.25
1,000 animals tested once	-1.24	-1.01	-0.88
500 animals tested twice	-1.14	-0.99	-0.89
Test 800, then lowest 200	-1.32	-1.13	-1.00

For breeding purposes, there is also a need to identify families as well as individuals. This could be done as part of the multi-stage selection procedure, by using genetic analyses to rank sire groups prior to each stage of testing. However, even when the main focus is to select individuals, the increased accuracy of selection should also mean that superior families will have a higher proportion of individuals tested at subsequent stages, leading to increased accuracy of selecting families as well as individuals.

Tables 1 and 2 show than even simple strategies, such as re-testing the most extreme 15% or 20% of animals based on the initial test results, are substantially better than one-stage testing. Such strategies also provide information on repeatability and assist with the selection of both genetically and phenotypically extreme individuals.

DISCUSSION

Identifying low emitting animals for further study is an important part of the strategy to reduce methane emissions from livestock. Although the technology to measure field emissions based on a 1 or 2 hour test looks promising, its accuracy will be limited by the ability to synchronise animals or statistically adjust for time of day effects, the latter perhaps being more practical for field tests with non-research flocks.

Given the difficulties and expense of field testing, other methods (such as testing for the presence or absence of indicator organisms in the rumen) might in the long run become the preferred strategy. Critical to the development of effective, new strategies is our ability to identify and study exceptional animals, either extremely low emitters, or high and low extremes, so that we identify and understand what causes the differences. As well as facilitating the development of alternative tests, understanding the mechanisms that lead to high or low emissions could aid the development of other possible ways of reducing emissions (e.g. dietary additives).

If atmospheric greenhouse gases are to stabilise, it may be necessary to reduce global emissions by 50% by 2020, a very ambitious target that is unlikely to be achieved for livestock solely by breeding strategies. Nonetheless, breeding from the most extreme animals (or extreme families) will provide useful information about the trait and whether responses at the extremes are linear.

Strategies to identify low or extreme emitting animals are therefore seen as at least as important to the research program as estimating genetic parameters. Multi-stage selection strategies are practical only if the results of the first test are available in time to select animals for the next stage of testing, which might be scheduled immediately after the end of the first stage. As long as this is feasible, multi-stage testing is recommended to improve the accuracy of selecting the most extreme animals, as well as provide an estimate of repeatability, without re-testing substantial numbers of animals, allowing the bulk of effort to be concentrated on testing animals at least once to maximise the amount of useful information collected on phenotypic and genetic variability.

REFERENCES

DCC (2008). "National Greenhouse Gas Inventory" Australian Government Department of Climate Change. Available at: http://www.climatechange.gov.au/inventory/2006/index.html

Goopy, J.P, Hegarty, R.S. and Robinson, D.L. (2009). Two hour chamber measurement provides useful estimate of daily methane production in Sheep. Proc 11th International Symposium on Ruminant Physiology, Clermont-Ferrand, France, September 6-9, 2009 (*in press*).

Robinson, D.L. (1984). *J. Agric. Sci. Camb.* **102**:119.

HERITABILITY OF PLASMA CONCENTRATIONS OF IGF1 AND ITS CORRELATION WITH REPRODUCTIVE PERFORMANCE IN HOLSTEIN COWS IN VICTORIAN HERDS

T.E. Stirling [1] C.R. Stockdale[2] and K.L. Macmillan[1]

[1]The University of Melbourne, Department of Veterinary Science, Werribee, Victoria 3030
[2]Department of Primary Industries, Tatura, Victoria.

SUMMARY

Compromised cow fertility is a significant cause of productivity loss in the dairy industry. It is hard to measure accurately, and has low heritability, but can be improved through selection of correlated traits with greater heritability. Plasma concentration of insulin-like growth factor-1 (IGF1) has recently been identified as a possible indicator trait for fertility. To investigate this further, the heritability of plasma IGF1 and its correlation with reproductive performance were measured in a group of 3700 Holstein cows from 22 commercial dairy herds in Victoria. Heritability was moderate for IGF1 (0.23 ± 0.05) but low for reproduction traits (0.06 – 0.10) even though between-cow variation was high. Genetic correlations between IGF1 and reproductive outcomes could not be estimated. Phenotypic correlations between IGF1 and reproductive outcomes were low and positive, but significant. Pearson's correlation coefficients between sire Australian Breeding Values for fertility had moderate positive correlations with sire best linear unbiased prediction solutions for IGF1, suggesting a moderate genetic correlation exists. Analysis of Variance between IGF1 and reproductive outcomes also showed a positive relationship between IGF1 and oestrus and conception. These results suggest that in the pasture-based dairying systems of Victoria, plasma concentrations of IGF1 could be gradually altered by genetic selection. Any associated affects on fertility would be positive, but minor.

INTRODUCTION

Continual breeding for high milk production in dairy cows has lead to physiological changes associated with declining fertility. Compromised fertility is a significant area of productivity loss. Improving it through genetic selection is difficult as fertility is hard to measure accurately and has low heritability. Plasma concentrations of insulin-like growth factor-1 (IGF1) could be used as a tool to improve cow fertility (Moyes 2004). Plasma IGF1 is related to the reproductive performance of the dairy cow due to its association with the partitioning of nutrients among biological functions (Bauman and Currie 1980) and its direct stimulatory affects on the ovaries (Spicer *et al.* 1993). IGF1 stimulates ovarian granulosa cell proliferation and mitogenesis, enhancing steroidogenesis by granulosa cells, and stimulating progesterone production (Spicer *et al*. 1993). Plasma concentrations of IGF1 are sensitive to nutrition, and the negative energy balance that dairy cows experience in early lactation is associated with low plasma IGF1 concentrations (Sharma *et al*. 1994). Therefore, IGF1 is considered to be a mediator of the effects of nutrition on reproduction, and has been associated with the resumption of cyclic activity in early lactation both overseas (Lucy *et al*. 1992; Beam and Butler 1998) as well as in Australia (Obese 2003; Moyes 2004). Previous studies involving IGF1 in dairy cows have only investigated phenotypic relationships. Moderate negative genetic and phenotypic correlations between IGF1 and milk production have been demonstrated for the group of cows in this study (Stirling *et al*. 2008). The aim of this study was to derive heritability estimates for plasma IGF1 in Holsteins in Victorian herds as well as genetic and phenotypic associations between IGF1 and reproductive performance. The findings could provide information to allow the industry to balance the need to maintain its commercial competitiveness by increasing milk yield without sacrificing fertility.

MATERIALS AND METHODS

Herds. Twenty-two commercial dairy herds throughout Victoria were involved in the study. Herd selection criteria included: predominantly Holstein genetics, consistent use of semen from the progeny testing program of Genetics Australia; good records and safe facilities. Herds were classified into one of two "systems" referring to calving management i.e. "seasonal" (single calving period each year) and "split" (multiple calving periods each year). Herds were also classified by calving "season" referring to the time of year and year and in which that herd calved, for example "Spring 2005" and "Autumn 2006".

Reproduction. Each herd bred cows during defined "breeding periods". Seasonal herds had one breeding period per year whereas split calving herds had two or three breeding periods per year. Mating start date (MSD) marked the beginning of each breeding period and usually occurred approximately 12 weeks after the start of the corresponding calving period. During the breeding period cows were artificially inseminated (AI'd) when they were detected in oestrus. Breeding periods varied in length between herds (typically ~9-12 weeks). Some herds introduced a bull into the herd at the end of the AI breeding period. Pregnancy testing typically occurred ~6 weeks after the end of AI. Reproductive performance was estimated using calving dates, AI dates and pregnancy test results. Parameters calculated included the interval (in days) between MSD to first AI and conception (MSD1stserv and MSDConc, respectively) for cows with positive reproductive outcomes. Binary outcomes for 3 and 6 week AI submission rates and 6 and 21 week in-calf rates were calculated as an indicator of oestrus and conception, respectively, at a given number of weeks after MSD (for all cows; both positive and negative reproductive outcomes).

Sampling and assays. Blood samples were collected from the tail vein of each enrolled cow during the dry period (Dry) and around MSD. Samples were placed on ice immediately, centrifuged at 1800 x g for 10 min, and plasma frozen at -20ºC until assays were undertaken. Concentrations of total plasma IGF1 was measured using the DSL-10-2800ACTIVETM Non-extraction IGF1 ELISA (Diagnostic Systems Laboratories, Webster, Texas, USA) in the laboratories of Primegro Ltd, Adelaide, using the method of Obese *et al*. (2008). Inter-assay variation (CVs) ranged from 4.3% to 6.9%, and mean intra-assay variation from 3.0% to 4.5%.

Data analyses. Heritabilities and genetic correlations of IGF1 and reproductive performance were calculated using ASReml (Gilmour *et al*. 2006). An animal model was used, utilising pedigree information from each cow's dam, sire and maternal grand-sire. Approximately 3700 cows of $\geq$ second parity were included in the analyses. Univariate analyses were performed to estimate heritabilities and confirm the suitabilities of models used for each trait. Bivariate analyses were performed to obtain estimates of genetic correlations. Contemporary groups of cows were constructed from the concatenation of calving system, herd and recording date as these were not cross-classified effects. Linear covariates for the intervals between drying off and sampling and sampling to calving were fitted for dry IGF1. A linear covariate for days in milk at sampling was fitted for MSDIGF1. Cow age was fitted as a yearly class effect, with cows 10 years or older grouped together into the same class. Models for IGF1 traits were redefined to accommodate herd, sampling date and IGF1 assay batch/plate effects. Best linear unbiased prediction (BLUP) solutions for IGF1 in sires were also calculated, and correlated with sire Australian Breeding Values (ABVs) for fertility using Pearson's correlations. Sire ABVs were obtained from a recent genetic evaluation by the Australian Dairy Herd Improvement Scheme, which used data from all daughters and herds, rather than just herds recorded in this study, giving a relatively high

reliability for fertility (84%). The relationship between IGF1 and binary reproductive outcomes for oestrus and conception were analysed with ANOVA using residuals for IGF1 that had been adjusted for herd, season, age, sampling date and IGF1 assay batch/plate effects.

RESULTS AND DISCUSSION

Means, variation and heritability estimates for each trait are outlined in Table 1. Means and variation of plasma IGF1 were similar to that reported in other studies involving pasture-fed dairy cows (Obese 2003; Moyes 2004), lower than that recorded for beef cattle (Johnston *et al.* 2001), but higher than that reported for dairy cows fed total mixed rations (Beam and Butler 1998). Plasma IGF1 had a moderate heritability (0.23), which is within the range of 0.18 to 0.48 that has been reported for dairy and beef cattle in other studies (Davis and Simmen 2000; Grochowska *et al*. 2001; Johnston *et al*. 2001). Means for fertility traits suggest a substantial decline in fertility in Victorian dairy herds since a large field study by Morton (2000). The heritability of fertility traits was low ($\leq$ 0.10; Table 1). This was similar to most international studies in dairy cows (Grosshans *et al.* 1997; Pryce *et al.* 1998; Berry *et al.* 2003). Fertility is difficult to measure as the phenotypic outcome depends on many non-genetic factors, ie. environmental and management factors. This is why it may be more effective to improve the genetic potential of fertility by selecting for traits with higher heritabilities that are known to be associated with fertility.

Table 1. Means, coefficient of variation (CV), heritability (h^2) and phenotypic variation (σp) of plasma IGF1 measured at the dry period (Dry) and mating start date (MSD), the interval between MSD and first service (MSD1stserv) and MSD and conception (MSDConc), 3 and 6 week submission rates (3 wk Sub, 6 wk Sub), and 6 and 21 week conception rates (6 wk Con, 21 wk Con).

	Mean ± SD	CV	h^2 ± SE	σp
Dry IGF1 (ng/ml)	154.0 ± 60.2	39	0.23 ± 0.05	43.8
MSD IGF1 (ng/ml)	89.4 ± 39.0	44	0.23 ± 0.05	33.7
MSD1stserv (days)	19.0 ± 28.4	149	0.06 ± 0.04	24.3
MSDConc (days)	46.9 ± 57.3	122	0.10 ± 0.05	49.8
3 wk Sub (%)	0.70 ± 0.46	-	0.03 ± 0.02	0.40
6 wk Sub (%)	0.88 ± 0.33	-	0.09 ± 0.04	0.28
6 wk Con (%)	0.55 ± 0.50	-	0	0.46
21 wk Con (%)	0.78 ± 0.41	-	0.02 ± 0.03	0.37

SD = standard deviation, SE = standard error

Most fertility traits failed to converge in bivariate analyses because of very low heritabilities (when estimable) under univariate analyses. For MSD1stserv the genetic variance was forced to the zero boundary so that a genetic correlation was not estimable. Phenotypic correlations were only obtained between MSDIGF1 and MSD1stserv (-0.11 ± 0.02; correlation ± standard error) and MSDIGF1 and MSDConc (-0.06 ± 0.02). This suggested that higher IGF1 was associated with slightly reduced intervals to conception and a higher probability of observing oestrus soon after MSD (phenotypically). ANOVA between residuals for IGF1 and binary reproductive outcomes showed that cows that showed oestrus and were inseminated in the first 3 and 6 weeks after MSD had higher IGF1 (Dry and MSD) than cows that were not ($p<0.05$). Cows that conceived in the first 6 and 21 weeks after MSD had numerically greater IGF1 than cows that did not, but this was only statistically significant for the 6 week conception rate with MSDIGF1. Pearson's correlation coefficient between sire ABVs for fertility and sire BLUP solutions for IGF1 was 0.26 for Dry IGF1 and 0.16 for MSD IGF1. This gave further information on the positive genetic relationship

between IGF1 and fertility, given the higher reliability of fertility data using greater numbers of animals in the ADHIS database to calculate ABV for fertility.

CONCLUSIONS

In the herds studied, plasma IGF1 concentrations displayed moderate variation and moderate heritability. Fertility traits had high variation (phenotypic) but low heritability. Genetic correlations between IGF1 and reproductive performance could not be demonstrated with this dataset. A moderate genetic relationship was demonstrated between sire BLUP solutions for IGF1 and sire ABVs for fertility. The phenotypic relationship between IGF1 and reproductive performance was positive, but weak. These results suggest that in Victorian dairy systems, IGF1 could be gradually altered by genetic selection, but the affect on fertility would probably be small. Given the negative genetic correlation between milk yield and IGF1 previously shown, further investigations could examine the use of IGF1 in a multi-trait selection index, or the relationship between sire IGF1 and daughter fertility to improve the reliability of sire ABVs for fertility.

ACKNOWLEDGEMENTS

The authors would like to acknowledge the generosity and cooperation of the study herd owners and the University of Melbourne and Primary Industry Research Victoria employees who provided technical support. Statistical analyses were performed by Dr Kim Bunter from the Animal Breeding and Genetics Unit of the University of New England. This research was part of the Dairy IGF1 Project funded by Dairy Australia, Genetics Australia and QAF Ltd.

REFERENCES

Bauman, D. E. and Currie, W. B. (1980) *J. Dairy Sci.* **63**:1514.

Beam,S.W. and Butler, W.R. (1998). *J. Dairy. Sci* **81**:121.

Berry, D.P., Buckley, F., Dillon, P., Evans, R.D., Rath, M and Veerkamp, R.F. (2003). *J. Dairy. Sci.* **86**:2193.

Davis, M.E. and Simmen, R.C. (2000). *J. Anim. Sci*. **78**:2305.

Gilmour, A.R., Gogel, B.J., Cullis, B.R. and Thompson, R. (2006) "ASReml User Guide Release 2.0" VSN International Ltd, Hemel Hempstead, Hertfordshire, UK.

Grochowska, R., Sorensen, P., Zwierzchowski, L., Snochowski, M. and Lovendahl, P. (2001). *J. Anim. Sci*. **79**:450.

Grosshans, T., Xu, Z.Z., Burton, L.J., Johnson, D.L. and Macmillan, K.L. (1997). *Livest. Prod. Sci* **51**:41.

Johnston, D.J., Herd, R.M., Reveter, A. and Oddy, V.H. (2001). *Proc. Assoc. Advmt. Anim. Breed. Genet*. **14**:63.

Lucy, M.C., Staples, C.R., Thatcher, W.W., Erickson, P.S., Cleale, P.S., Firkins, J.L., Clark, J.H., Murphy, M.R. and Brodie, B.O. (1992). *Anim.. Prod*. **54**:323.

Morton. J. (2000). "The Incalf Project, Progress Report # 2". Dairy Research Development Corp.

Moyes, T.E. (2004). PhD Thesis, University of Melbourne.

Obese, F.Y.(2003). PhD Thesis. University of Melbourne.

Obese, F.Y., Humphrys, S., Macmillan, K.L. and Egan, A.R.. (2008). *J. Dairy Sci*. **91**:160.

Pryce, J.E., Esslemont, R.J., Thompson, R., Veerkamp, R.F, Kossaibati, M.A. and Simm, G. (1998). *Anim.. Sci.* **66**:577.

Sharma, B.K., Vandehaar, M.J. and Ames, N.K. (1994). *J. Dairy Sci.* **77**:2232.

Spicer, L. J., Alpizar, E. and Echternkamp, S. E. (1993). *J. Anim.. Sci*. **71**:1232.

Stirling, T.E., Stockdale, C.R., and Macmillan, K.L. (2008) *Proc. New Zeal. Soc. Anim. Prod.* **68**:98.

BIOETHICS AND DNA DIAGNOSTICS IN ANIMALS – ARE THERE LESSONS TO BE LEARNED FROM GENETIC TESTING IN HUMANS?

I. Tammen and H.W. Raadsma

Reprogen. Faculty of Veterinary Science, University of Sydney, Camden, NSW 2570

SUMMARY

Publications in the field of human genetics often include or focus on ethical considerations of genetic research and DNA diagnostics. Although similar ethical dilemmas exist in relation to DNA testing of animals, these appear to be discussed less frequently in the scholarly literature. This paper is an attempt to identify similarities and differences in ethical concerns about genetic testing when applied to either the human or animal context. Considering the greater debate in humans, there are likely lessons to be learned on how to approach specific ethical dilemmas. Within animal applications the issue of commercialization and public access to knowledge is likely to attract the great ethical debate. This paper is not aiming to prescribe what is 'right' or 'wrong' about genetic testing of animals but hopes to instigate awareness of ethical dilemmas and prompt further discussion within our profession and with breeding organisations, animal breeders, animal owners as well as the public.

INTRODUCTION

A recent publication in Nature Biotechnology titled 'Most gene test sales are misleading' (Wallace 2008) has provided some motivations for this paper. Our experiences as researchers and educators in animal genetics have been an even greater inspiration, particularly as our research has been focused on the molecular characterisation of inherited disorders and productive traits in livestock.

Researchers, service providers, animal owners and breeders as well as the public are well aware of the many ethical dilemmas in animal DNA testing (see examples in Table 1); however, a systematic approach to deal with these complex issues seems to be lacking. We propose that 'comparative' bioethics can be helpful to identify, clarify and approach ethical dilemmas in animal genetic research and more specifically in animal DNA testing. In the following, We will briefly define bioethics, summarize key information about DNA diagnostics, draw on human literature to summarize key ethical concerns in DNA diagnostics, briefly discuss how these can be seen to relate to animal genetics and conclude with some suggestions on what lessons can be learnt from human approaches.

WHAT IS BIOETHICS?

Bioethics represents a recently developed category of practical ethics, which deals with the application of ethics in the context of biological sciences. Ethics in its simplest definition can be understood as the study and philosophical framework of what should be done, or the study of how we should live in relation to others. Different systems or theories of ethics co-exist, and they generally agree that ethics requires systematic approach, is prescriptive, universalisable and of overriding importance (Kerridge *et al*. 20.05). Ethics is largely concerned with human flourishing and well-being; however, depending on the theory of ethics animals are considered to be of similar or lesser value. Kerridge *et al.* (2005) highlight the importance for professionals to actively engage in ethical discourse, they caution that no single approach can ever describe or resolve complex ethical dilemmas and thus conclude that ethics is about ongoing, respectful, transparent interdisciplinary discourse. Although no single ethical theory has emerged to dominate this discourse, principle-based approaches have been seen as very useful in providing a flexible

framework in such discussions (Kerridge *et al.* 2005). Balancing of four fundamental principles (identified as justice, autonomy, beneficence and non-malfeasance) is the basis for shared moral reflection and provides a framework for context specific action guides (Beauchamp and Childress 1994). Mepham (1996) has applied and modified this principle-based approach by proposing the use of an ethical matrix to engage with the complex and controversial issues in animal biotechnologies and to assist with the development of regulatory frameworks.

DNA DIAGNOSTICS

Genetic testing is also a relatively new field in the discipline of genetics, and recent improvements in technology and the broadening of applications have resulted in DNA testing being widely used in humans and animals. Traditional applications of DNA testing include direct and indirect DNA tests for inherited diseases/traits with Mendelian inheritance and 'genetic fingerprinting' for parentage testing and forensic studies. Genetic testing has been invaluable in many research fields such as anthropology and conservation genetics. More recent developments have lead to commercialisation of DNA tests for complex traits and multifactorial diseases as well as tests for the identification of breed or race. In addition, DNA tests are used for the detection of pathogens and species identification in animal derived food products.

Consequently 'DNA testing' can not be understood as a single entity – there are differences in regards to technologies used for genotyping, differences in the types of genetic variation that are detected, differences in nature of application, and most importantly differences in relation to the conclusions that are drawn from the genotyping results. Most DNA tests are recognised as accurate with a low probability of failure or misclassification if performed in accordance within technical standards (e.g. SCAHLS 2008, Standards Australia 2005). However, as methodologies to derive conclusions from genotyping are becoming more complex, the validity and/or accuracy of some of the interpretations of DNA tests have been questioned (e.g. Gollust *et al.* 2007; Wallace 2008).

ETHICS AND DNA DIAGNOSTICS

DNA diagnostics in humans. Genetic testing has been differentiated from other diagnostic testing as genetic information is considered to be ubiquitous, familial and often predictive (ALRC96 2003). These aspects can be seen as key strength of DNA tests, but are also the reason for concern: e.g. DNA can be obtained without knowledge or consent of those tested; DNA tests for an individual can intentionally or inadvertently reveal information about relatives, and predictive information about disease risk can be misunderstood or misused (ALRC96 2003). Detailed analyses have revealed a multitude of ethical issues (NHMRC 2000; Kerridge *et al*. 2005). It is beyond the scope of this paper to discuss these in detail but several topics are listed in Table 1. Debate of these ethical issues has lead to the development of ethical and legal guidelines on how we should deal with some of the emerging issues in this field (e.g. ALRC96 2003; ALRC99 2004)

DNA diagnostics in animals. Are any of the ethical issues in humans genetic testing (Table 1) of relevance in the animal context? Are there any additional issues that need to be considered? In an attempt to answer these questions two issues become evident: firstly, there is no consensus on what the moral status of animals is (e.g. Li 2002), and secondly, issues related to DNA diagnostics in animals are largely about human interactions and relationships – those between researchers, test providers, animal owners, breed organisations and the general public.

We will focus here on ethical questions in relation to humans and not engage in a debate about moral status of animals – not because we consider this of lesser importance, but because such a debate would be beyond the scope of this paper. In Table 1, ethical issues identified in human

genetics are used as a starting point for the identification of ethical dilemmas in animal DNA testing. The table provides examples only and is not an attempt to cover all possible or even common ethical dilemmas. As pointed out in the introduction it is not the aim of this paper to propose solutions to these questions but to encourage awareness and constructive debate, as well as to suggest a framework for improvement of DNA technologies in animal applications.

Table 1. Ethical issues in humans and comparative ethical questions in animal DNA testing

Ethical issues in human DNA testing relate to:	Examples of comparative ethical questions in animal DNA diagnostics
Equity of access	Do all animal owners (in all countries) have access to specific DNA tests? Is it just that some DNA tests for species seem to be more expensive than similar tests in other species? Does dependence on industry funding to develop/commercialise DNA tests create inequity – i.e. are only large breed organisations able to support this?
Allocation of resources	Is spending money on genetic research and testing of animals the best way to improve animal productivity and welfare? Or indeed human well-being?
Consent	Do ethics committees need to consent to non-invasive sampling for research? Who should consent to DNA testing of animals - the owner, the breeder, the purchaser of semen/embryo's?
Privacy / confidentiality / 'The right not to know'	Should breed organisations publish results of DNA testing? Should researchers or breed organisations identity founder animals of inherited diseases? Do animal breeders or breed organisations have 'the right not to know' When selling a DNA tested animal (or semen or offspring of a tested animal) should we notify the buyer of the test results?
Discrimination	What should we do with animals heterozygous for disease alleles – as animal owners, breeders or breed organisations? Does it matter if the disease is lethal, late onset or linked to a favourable allele? Is it fair to 'discriminate' against breeds that have been reported to have an inherited disease / or breeders that have advertently or inadvertently breed an animal with an inherited disease? Should differential rates of insurance apply? Should carriers be banned from shows/ registration in the herd book?
Predictive tests	Do animal owners understand the results of DNA tests – especially for those tests where complex interpretations of genotypes are required and reported? How reliable should these tests be before they can be commercialised? Should tests be periodically updated for predictive capacity as gene frequencies change or additional information becomes available? What role should researchers have in presenting an unbiased objective assessment of their discoveries? Is peer review sufficient?
Gene patenting	Should we patent DNA tests for animals?
Storage of material and information	Should we use samples submitted for diagnostics for future research? If we do – should we inform owners of the results?
Commercialisation & direct–to-consumer marketing	Should we require independent accreditation and/or validation of research before commercialisation? Considering trade secrets and increasingly complex data sets for multifactorial traits/disease is a requirement for independent validations financially viable? Should we regulate / self-regulate marketing? Should we include 'genetic counselling' as a requirement for reporting? Considering that technologies and methods are rapidly evolving what should service providers do if more accurate tests become available or existing tests are identified as misleading?

CONCLUDING REMARKS

We have proposed here to use a 'comparative ethics' approach, where the 'rich' ethical debate in human DNA testing can be used to identify, clarify and approach ethical dilemmas in animal genetic testing in a more systematic approach. Table 1 suggests that this could be a useful approach, as very similar ethical questions exist. The severity of the ethical dilemmas appear to be greater in humans but the issues are possibly even more complex in animal testing, particularly if we consider the debate about moral status of animals.

The recent developments in DNA testing for multifactorial diseases and traits as well as testing for race or breed affiliation for both animals and humans have highlighted concerns about direct-to-consumer marketing, overselling as well as premature commercialisation. In addition to ethical concerns highlighted above, we need to be aware of the great risk that consumer confidence in DNA technologies in general can be lost easily if these new predictive tests don't deliver.

Genetic counselling is not a requirement for animal DNA testing but considering its importance in human diagnostics building up of further capabilities in this area should be encouraged, especially in the context of companion animal testing.

The aim of this paper was not to propose solutions to specific ethical questions but to explore what approaches could be used to do so. Principle-based approaches have proven useful in medical ethics (Kerridge *et al.* 2005) and have already been adapted to the animal context. The ethical matrix has been devised by Mepham (1996) as a framework for rational ethical analysis in animal biotechnologies and could be particular useful in exploring the ethical dilemmas exposed here. This will not lead to simple answers or precise action guides but could frame the debate between researchers, service providers, breed organisations, animal breeders and the public. Such a debate should include the constant review of relevant issues and where appropriate identify where and when additional guidelines, (self-)regulation or even legislation might be needed.

ACKNOWLEDGEMENTS

We are thankful to Peter Thomson and Julie Cavanagh for comments

REFERENCES

ALRC 96 (2003) "Essentially Yours: The Protection of Human Genetic Information in Australia". http://www.austlii.edu.au/au/other/alrc/publications/reports/96/

ALRC 99 (2004) "Genes and Ingenuity: Gene Patenting and Human Health". http://www.austlii.edu.au/au/other/alrc/publications/reports/99/

Beauchamp, T.L. and Childress, J. (1994) "Principles of biomedical ethics". Oxford University Press, Oxford.

Gollust, S.E., Hull, S.C. and Wilfond, B.S. (2002) *J. Am. Med. Assoc.***288**:1762.

Kerridge, I., Lowe, M. and McPhee, J. (2005) "Ethics and Law for the Health Professions", 2nd ed. Federation Press, Sydney.

Li, H.-L. (2002) *J. Med. Philos.* **27**:589.

Mepham, T.B. (1996) In "Progress in Dairy Science", editor C.J.C. Phillips (ed), CAB International, Wallingford, pp 375-395.

NHMRC (2000) "Ethical aspects of human genetic testing: an information paper." http://www.nhmrc.gov.au/your_health/issues/genetics/index.htm

SCAHLS (2008) "Veterinary Laboratory Guidelines for Nucleic Acid Detection Techniques March 2008" http://www.scahls.org.au/

Standards Australia (2005) "AS ISO/IEC 17025-2005: General requirements for the competence of testing and calibration laboratories".

Wallace, H. (2008) *Nat. Biotechnol.* **26**:1221.

QTL ANALYSIS OF BEEF FAT COLOUR AND THE EFFECT OF BCDO2

R. Tian[1], W.S. Pitchford[1] and Cynthia D.K.Bottema[1].

[1]School of Agriculture, Food and Wine, University of Adelaide, Roseworthy SA 5371 Australia

SUMMARY

Most agricultural economic traits are controlled by both the environment and genetics. Locating quantitative trait loci (QTL) using molecular markers remains a key approach to identify the genes controlling those traits. In this study, the QTL for fat colour in beef cattle were mapped. The mapping herd was a double-backcross design using two extreme *Bos taurus* breeds (the Jersey dairy breed and Limousin beef breed). Three fat colour traits were measured (beta-carotene concentration in subcutaneous fat and fat colour scores of biopsy and carcass fat), and a total of 16 QTL on 13 chromosomes were detected at the 5% genome-wise significance level.

A potential candidate gene in QTL on BTA 15, *β, β-carotene-9', 10'-dioxygenase* (*BCDO2*) was investigated. A SNP in exon 3 which creates a stop codon was previously identified, and inclusion of this *BCDO2* SNP as a fixed effect in the linkage analysis of the fat colour traits changed the F-values for a number of QTL. As expected, the QTL on BTA 15 was no longer significant. As the presence of the *BCDO2* genotype in the model reduced the residual variance, three additional QTL were detected for biopsy fat colour score and one additional QTL for carcass fat colour score.

INTRODUCTION

Beef with yellow fat is considered undesirable by the consumers in most European and Asian markets. Presumably, this is because beef with yellow fat is perceived as being from old or diseased animals. More than AUS$18 million is lost to Australian beef producers annually due to the rejection of carcasses with high yellow fat colour scores by the Asian markets, such as Japan and Korea (Browne 1992). The yellowness is caused by the deposition of beta-carotene from green feed into the adipose tissue (Kruk 2001). Over the past decade, there has been an increasing emphasis on the development of molecular genetic tools, such as DNA markers, to improve beef production and quality through marker-assisted selection. In this study, a whole genome scan for QTL affecting fat colour related traits in half-sib families generated from Limousin and Jersey breed was conducted.

MATERIALS AND METHODS

Cattle resources. Purebred Jersey (J) and purebred Limousin (L) cattle were crossed to produce three pairs of half brothers F_1 progeny (X). One of the half brothers from each pair was mated to the purebred Jersey or Limousin cows in Australia and the other half brother was mated in New Zealand. A total of 366 backcross progeny (205 XJ and 161 XL) were born in the autumn over 3 years from 1996 to 1998 in Australia. β-carotene concentration was measured in subcutaneous fat sampled at slaughter and in fat biopsy samples taken from around the tail at 12 months of age. Fat colour score of the adipose biopsy samples was estimated on a 5-point scale (l = white to 5 = very yellow) immediately after removing the fat from biopsy site and rinsing with water. Yellowness in the subcutaneous fat was also scored visually after slaughter, assessed on a 10-point (l=white to 10=very yellow) scale by trained assessors using standard colour chips (AUS-MEAT 1990) (Kruk 2001). Fat colour and beta-carotene data were examined for distribution before the analysis. Only the beta-carotene distribution was positively skewed. Therefore, the beta-carotene values were logarithmic transformed to reduce the skewness (Esmailizadeh 2006). DNA was extracted from all

cattle, and approximately 190 microsatellite markers were genotyped per animal, providing a whole-genome scan at roughly even 20 cM intervals.

Model of analysis. Linkage of each trait to markers on the chromosomes was tested using the QTL Express software package (Seaton *et al.* 2002; http://qtl.cap.edu.ac.uk, accessed 3rd November 2006). The analysis uses the multimarker approach for interval mapping in the half-sib families, as described by Knott *et al.* (1996). Within every half-sib family, a QTL was fitted at 1-cM intervals along the chromosome. The least squares regression model was used for the QTL mapping and included the fixed effects of breed of dam (J or L) and cohort (96H, 96S, 97H, 97S, 98H, 98S). 5% genome-wise significance thresholds were determined using permutation methods to account for multiple testing (Churchill and Doerge 1994). In general, an *F* statistic of greater than 7 was set as the threshold (Esmailizadeh 2006). A t-test was calculated for the most likely position of a QTL. The *BCDO2* gene effect (θ_k) was excluded in the first model (model 1) and included in the second model (model 2).

$Y_{ijkl} = \mu + \alpha_i + \beta_j + \delta_l + bx + e_{ijkl}$ (model 1)

$Y_{ijkl} = \mu + \alpha_i + \beta_j + \theta_k + \delta_l + bx + e_{ijkl}$ (model 2)

Where

μ overall mean,

α_i effect of the i^{th} breed (Jersey or Limousin),

β_j effect of the j^{th} cohort (six levels),

θ_k effect of the k^{th} genotype (AA, GA, GG),

δ_l effect of l^{th} sire, b_l is the allele substitution effect of the QTL within family l, x_l is the probability that animal *m* inherited the arbitrarily assigned first haplotype of sire l,

bx b is the allele substitution effect of the putative QTL, x is the probability that animal l inherited the arbitrarily assigned first haplotype of the sire, and

e_{ijkl} is the residual effect.

RESULTS AND DISCUSSION

A nonsense mutation SNP (W240X) was found in the *BCDO2* gene (Tian, *et al* accepted). This SNP was located at base 240 from the start of coding region, and causes an amino acid change at position 80 from a tryptophan (encoded by the G allele) to a stop codon (encoded by the A allele). Association analysis of this nonsense mutation showed significant differences in subcutaneous fat colour and beta-carotene concentration amongst cattle with different *BCDO2* genotypes. Therefore, the *BCDO2* SNP W240X was fitted in the model to detect new QTL for fat colour related traits.

Allele substitution or sire QTL effects, estimated QTL locations and F-statistic value were calculated (Table 1). Without the *BCDO2* SNP W240X in the model, a total of 16 QTL segregating in all the individual sire families on 13 chromosomes were detected at the 5% genome-wise significance level. The largest QTL was for carcass fat colour near 74 cM on chromosome 12.

When the SNP W240X was fitted as a fixed effect in the model, there are some changes the F values and positions of the QTL. A large decrease occurred in the BTA 15 QTL in family 398 for biopsy fat colour. The test statistic diminished at 45cM for BTA15 as expected, since this is the location of the *BCDO2* gene. However, another 4 new QTL were observed, for which the test statistic increased from non-significant to significant on BTA1, BTA7, BTA8 and BTA14.

Vitamin A is essential for mammals. One of the functions of carotenoids is to serve as the precursor of vitamin A. The enzyme *BCDO2*, which specifically acts at the 9' 10'-double bond of β-carotene, results in the formation of β-ionone and two molecules of β-apocarotenal with different chain lengths (Von Lintig and Vogt 2004). Given that the *BCDO2* protein is 530 amino acids in length, this change will result in the truncated polypeptide and presumably the loss of *BCDO2* protein function. So the excentric cleavage pathway of β-carotene will be affected and thus β-

carotene accumulation in the adipose tissue and changes the colour in beef fat. As a potential candidate gene on chromosome 15, this SNP was fitted as a fixed effect in the model. The test statistic for QTL on BTA15 diminished as expected. No other QTL diminished, so there was no evidence of other epistatic QTL in this analysis.

Table 1. QTL segregating in single sire families

BTA	*Trait*	*family*	*with/without SNP W240X*	*QTL position*	*F value*	*sire QTL effects*
1	Fcam[1]	sire 361	model without SNP W240X	5cM	7.68	-0.49 (±0.18)
1	Fcam	sire 361	model with SNP W240X	5cM	7.14	-0.48 (±0.18)
1	Fcbiop[2]	sire 398	model without SNP W240X	23cM	2.47	-0.25 (±0.16)
1	Fcbiop	sire 398	model with SNP W240X	0cM	6.72	-0.34 (±0.13)
2	Bcbiop[3]	sire 361	model without SNP W240X	64cM	7.61	0.23 (±0.09)
2	Bcbiop	sire 361	model with SNP W240X	64cM	7.11	0.23 (±0.09)
6	Fcam	sire 361	model without SNP W240X	54cM	7.02	0.47 (±0.18)
6	Fcam	sire 361	model with SNP W240X	54cM	6.69	0.46 (±0.18)
6	Fcbiop	sire 361	model without SNP W240X	46cM	9.03	-0.26 (±0.09)
6	Fcbiop	sire 361	model with SNP W240X	46cM	8.96	-0.26 (±0.09)
6	Fcbiop	sire 368	model without SNP W240X	68cM	7.96	-0.25 (±0.09)
6	Fcbiop	sire 368	model with SNP W240X	67cM	7.62	-0.25 (±0.09)
7	Bcbiop	sire 368	model without SNP W240X	46cM	5.58	0.16 (±0.07)
7	Bcbiop	sire 368	model with SNP W240X	43cM	7.21	0.18 (±0.07)
8	Bcbiop	sire 398	model without SNP W240X	0cM	3.65	-0.26 (±0.13)
8	Bcbiop	sire 398	model with SNP W240X	0cM	7.33	-0.33 (±0.12)
8	Fcam	sire 361	model without SNP W240X	53cM	11.8	0.74 (±0.22)
8	Fcam	sire 361	model with SNP W240X	52cM	12.29	0.77 (±0.22)
11	Fcbiop	sire 368	model without SNP W240X	98cM	15.29	0.3 (±0.08)
11	Fcbiop	sire 368	model with SNP W240X	104cM	16.89	0.35 (±0.09)
12	Bcbiop	sire 361	model without SNP W240X	98cM	8.46	-0.35 (±0.12)
12	Bcbiop	sire 361	model with SNP W240X	98cM	8.38	-0.35 (±0.12)
12	Fcam	sire 368	model without SNP W240X	74cM	17.87	-0.73 (±0.17)
12	Fcam	sire 368	model with SNP W240X	74cM	18.23	-0.73 (±0.17)
14	Fcam	sire 398	model without SNP W240X	0cM	2.34	0.39 (±0.26)
14	Fcam	sire 398	model with SNP W240X	21cM	10.39	0.7 (±0.22)
14	Fcbiop	sire 361	model without SNP W240X	0cM	15.45	-0.34 (±0.09)
14	Fcbiop	sire 361	model with SNP W240X	0cM	15.75	-0.34 (±0.09)
15	Fcbiop	sire 398	model without SNP W240X	45cM	7.23	0.54 (±0.2)
15	Fcbiop	sire 398	model with SNP W240X	89cM	4.25	0.27 (±0.13)
16	Bcbiop	sire 361	model without SNP W240X	51cM	7.64	-0.23 (±0.08)
16	Bcbiop	sire 361	model with SNP W240X	51cM	7.66	-0.23 (±0.08)
16	Bcbiop	sire 368	model without SNP W240X	85cM	6.94	-0.18 (±0.07)
16	Bcbiop	sire 368	model with SNP W240X	85cM	6.91	-0.18 (±0.07)
17	Bcbiop	sire 368	model without SNP W240X	70cM	7.53	-0.19 (±0.07)
17	Bcbiop	sire 368	model with SNP W240X	70cM	6.83	-0.18 (±0.07)
20	Fcam	sire 368	model without SNP W240X	17cM	7.82	0.48 (±0.17)
20	Fcam	sire 368	model with SNP W240X	17cM	9.04	0.51 (±0.17)
24	Fcbiop	sire 361	model without SNP W240X	46cM	6.9	-0.31 (±0.12)
24	Fcbiop	sire 361	model with SNP W240X	46cM	6.84	-0.31 (±0.12)

[1]Fcam: fat colour of carcass on a 10 point scale. [2]Fcbiop: fat colour of biopsy samples on a 5 point scale. [3]Bcbiop: β-carotene concentration of fat biopsy samples (μg/g fat)

CONCLUSION

In summary, inclusion of the *BCDO2* SNP as a fixed effect in the linkage analysis of the fat colour traits changed the F-values for a number of QTL and four new QTL for fat colour traits were identified. The new QTL were presumably observed because the *BCDO2* SNP accounted for some of the residual variation in these fat colour traits. Based on these QTL, a number of additional candidate genes have been selected and are currently being investigated for their potential association with beef fat colour.

REFERENCES

Browne, G.M. (1992) *Proceedings Aust. Soc. Anim. Prod.* **19:**91
Churchill G.A. & Doerge R.W. (1994) *Genetics* **138:**963
Esmailizadeh, A.K. (2006) PhD Thesis, University of Adelaide
Kruk, Z.A. (2001) PhD Thesis, University of Adelaide
Knott, S.A., Elsen, J.M. & Haley, C.S. (1996) *Theor. Appl. Genet.* **93:**71
Tian R, Pitchford W.S, Morris C.A., Cullen N.G., Bottema C.D.K.. *Animal Genetics* (accepted).
Von Lintig, J.V., and Vogt K. (2004) *J. Nutr*. **134:**251

GENOTYPE x ENVIRONMENT INTERACTIONS IDENTIFIED IN SOUTHERN AUSTRALIAN BEEF PRODUCTION

S.L. Truran[1], M.P.B. Deland[2], M.L. Hebart[3], A.P. Verbyla[1,4] and W.S. Pitchford[1]

Cooperative Research Centre for Beef Genetic Technologies
[1]School of Agriculture, Food and Wine, University of Adelaide, Roseworthy SA 5371
[2] Struan Research Centre, SARDI, Naracoorte SA 5271
[3] SARDI, Roseworthy SA 5371
[4] CSIRO, Waite Campus, Glen Osmond SA 5064

SUMMARY

Data from the "Southern Crossbreeding Project" were utilised for the analysis of genotype by environment (G x E) interactions in a southern Australian production system. "Domestic heifers" were slaughtered for the domestic market at an average of 16 months (218kg Hot Standard Carcass Weight and 9.7mm rump fat at the P8 site) whilst "Export steers" were slaughtered for the export market at an average of 23 months (323kg HSCW and 14.7mm P8 fat). Performance of each sex was treated as a separate trait to assess the interaction between carcass traits at different market end points. Data were analysed using a univariate animal model containing the fixed effects of sex, sex x breed and sex x management group. Traits other HSCW were analysed on a weight constant basis by fitting HSCW × sex as a fixed effect. The random effects were animal (pedigree) for each sex resulting in a genetic variance for each sex and the correlation between heifers and steers (3 genetic and 2 residual variance components). For all carcass traits, the genetic correlation between market end points was less than one, although only significantly so for loin eye muscle area. It is not possible to determine the extent to which the significant interaction was the result of different weight endpoints, ages or differences between heifers and steers.

INTRODUCTION

Genotype by environment (G × E) interactions can arise due to both natural variations and management differences. If they are significant, this creates a potential industry issue if an individual's breeding performance is altered across environments (Falconer 1952). Data used in this study were obtained from The Southern Crossbreeding project, conducted in South Australia with calves born from 1994–1997. For this project, heifers and steers were slaughtered at different time points to meet both the domestic and export markets, respectively. The objective of this study was to analyse G × E interactions, defining the environment as the different market end points. In concurrence with previous studies (Johnston *et al.* 2003; Reverter *et al.* 2003), it was hypothesised that potential variation between market end points may be identified for certain traits, but no significant G × E interaction would be identified. Researching G × E interactions between market end points is important for producers because if G × E does exist between market endpoints, producers will have to select genotypes to match the market specifications to ensure maximum profit (Reverter *et al*. 2003).

MATERIALS AND METHODS

Animals and management. The project was conducted both at Struan Research Centre and Wandilo (a nearby property), in the south east of South Australia. The region is characterised by a Mediterranean climate. Jersey, Wagyu, Angus, Hereford, South Devon, Limousin and Belgian Blue sires were mated to Hereford cows in a top cross design (Pitchford *et al*. 2002). Calves were born at either Struan or Wandilo in three management groups. Calves were weaned at 9 months of

age in early January each year. Following weaning, all calves were managed at Struan where they were randomly allocated to between one and three post-weaning management groups (Pitchford *et al*. 2002). Calves were grown until 12 to 18 months of age and then transported to a commercial feedlot. The feedlot ration included a minimum of 60% grain with approximately 12 MJ ME/kg DM and 13% protein (Pitchford *et al*. 2002). The exception to this was the 1997 Export Steers that reached marketable weights without requiring grain finishing due to good available pasture.

Carcass traits. Heifers (n=636) were slaughtered for the domestic market at an average of 16 months (218kg hot standard carcass weight (HSCW) and 9.7mm P8 fat (rump fat depth at the P8 site)) whilst Export Steers (n=691) were slaughtered for the export market at an average of 23 months (323kg HSCW and 14.7mm P8 fat) at various commercial abattoirs throughout south eastern Australia. Carcass traits recorded included HSCW based on a standard trim (AUSMEAT, 1995), rump P8 fat, eye muscle area (EMA) and intra-muscular fat content (IMF). EMA was adjusted in relation to the site of quartering as per adjustments cited by Pitchford *et al*. (2006).

Statistical analyses. Analyses were conducted using ASREML 2.0 (Gilmour *et al*. 2000). Data were analysed using a univariate animal model containing the fixed effects of sex, sex × breed, sex × management group. HSCW × sex was fitted as a fixed effect for all traits excluding HSCW. The random effects were animal (pedigree) for each sex, resulting in a genetic variance for each sex and the correlation between heifers and steers (3 genetic and 2 residual variance components). An additional parameter estimated using CORH provided the genetic correlation between Domestic Heifers and Export Steers. Fat traits were log transformed because of scale effects on the variance. The likelihood ratio test statistic ($-2\Delta l$) (Kendall and Stuart, 1973), which is distributed as a χ^2 distribution, was used to test significance for values with 3.84 used as the threshold.

RESULTS

The raw means highlight that for the four traits analysed, males recorded greater values than females (Table 1). The differences reflect the market specifications each sex was slaughtered at. Heritabilities were calculated for the 4 carcass traits for both Export Steers and Domestic Heifers (Table 1). Heritability estimates for HSCW were high, especially for Domestic Heifers. Estimates were lower for IMF but moderate for the other traits analysed.

Table 1. Mean, phenotypic variance (V_p), heritabilities (h^2) and genetic correlations for traits analysed for Domestic heifers and Export steers. Significance was tested using the likelihood ratio test statistic which is distributed as a χ^2.

Trait	Domestic Heifers			Export Steers			Genetic Correlation	χ^2 prob.
	Mean	V_p	h^2	Mean	V_p	h^2		
HSCW	218	425.4	0.82	323	988.9	0.54	0.725	0.07
P8 Fat*	9.7	0.118	0.31	14.7	0.109	0.38	0.828	0.49
EMA	67.5	50.40	0.30	75.1	69.84	0.37	0.337	0.04
IMF*	3.72	0.088	0.21	5.34	0.111	0.15	0.999	0.96

* Data were transformed for analysis

Genetic Correlations. Genetic correlations for all traits measured are presented in Table 1. Robertson (1959) considered a genetic correlation above 0.8 to exclude G × E interactions. However, within this study we consider any genetic correlations less than unity to indicate some level of G × E interaction, with a significant interaction tested using the likelihood ratio test statistic. For the fat traits, the genetic correlation between Domestic Heifers and Export Steers was either close to (P8 fat) or unity (IMF). However, HSCW (P=0.07) and EMA (P<0.05) were less than unity.

DISCUSSION

In the absence of G × E interactions, breeders can be confident in their selection choices across a range of production systems. Alternatively, detecting G × E interactions allows for producers to be aware of the role of specific genotypes in their production environment.

Heritabilities estimated differed between Export Steers and Domestic heifers. However, generally heritabilities were in a similar range, with the exception of HSCW. For HSCW and IMF, domestic heifers had a higher heritability than export steers. Similarly, a trend for slightly higher heritabilities for domestic market carcasses (220 kg carcass weight) in comparison to export carcasses (>280kg carcass weight) was identified by Johnston *et al*. (2003). Actual heritabilities for the traits EMA and P8 fat differ between studies.

The genetic correlations identified for the traits in this project indicate that there are G × E interactions present between market end points in southern production systems for HSCW and EMA. The analysis, excluding HSCW as a covariate, identified that HSCW has a χ^2 probability of 0.07. Indicating a strong interaction between G × E for this trait. Although HSCW was not significant (P<0.05), the interaction for EMA was significant (χ^2 = 0.04), proving a interaction between market endpoints for this trait. The G × E interaction identified between these two traits indicates differences in growth between the two cohorts. Arguably, males and females may have different muscular and skeletal patterns. However, the high level of G × E detected for EMA and HSCW was not detected in the other carcass traits (P8 fat and IMF).

In previous studies, there has been limited evidence of G × E interactions at different market end points. Johnston *et al*. (2003) and Reverter *et al*. (2003) found there was no significant genotype by environment interaction between three different market end points (220, 280 and 340 kg carcass weight) for both tropical and temperate breeds, concluding that the effect of genotype by market weight interactions was small. Johnston *et al*. (2003) attributed this to an increasing genetic expression of traits with increasing market weight, possibly reflecting increases in scale. In contrast with this current study, Johnston *et al*. (2003) analysed various carcass trait measurements at 3 different stages prior to slaughter, rather than at slaughter. Reverter *et al*. (2003) found some level of G × E interactions for fat colour (r_G=0.73) in temperate breeds and retail beef yield (r_G=0.53) in tropical breeds. Although this suggests G × E interactions at different market end points, Reverter *et al*. (2003) suggest these results may be due to other factors, for example differences in trimming between carcasses for RBY.

This study suggests that a significant G × E interaction exist in HSCW and EMA for different market end points. As previous studies have shown no significant G × E interactions between the domestic and export market endpoints, it can be assumed the differences observed in this study are due to differences in sex.

ACKNOWLEDGMENTS

The Southern Crossbreeding Project was supported by the South Australian Cattle Compensation Trust Fund, the J.S. Davies Bequest to the University of Adelaide and AW and PR Davis Pty. Ltd.

REFERENCES

AUSMEAT (1995) Handbook of Australian Meat, 5th ed., Australian Meat and Livestock Corporation, Brisbane, Australia

Choat, W.T., Paterson, J.A., Rainey, B.M., King, M.C., Smith, G.C., Belk, K.E., and Lipsey, R.J., (2006) *J. Anim. Sci*. **84**:1820.

Falconer, D.S., (1952) The Problem of Environment and Selection, The American Naturalist, Chicago Press, USA

Gilmour, A. R., Thompson, R., Cullis, B.R., and Welham, S.J., (2000) ASREML, User Notes., New South Wales Agriculture, Orange, NSW, Australia

Johnston, D.J., Reverter, A., Burrow, H.M., Oddy, V.H., and Robinson, D.L., (2003) *Aust. J. Agric. Res*. **54**:107.

Kendall, M.G., and Stuart, A., (1973) The Advanced Theory of Statistics, Vol. 2: Inference and Relationship, Charles Griffin and Co., London, United Kingdom

Pitchford, W.S., Deland, M.P.B., Siebert, B.D., Malau-Aduliand, A.E.O., and Bottema, C.D.K., (2002) *J. Anim. Sci*. **80**:2825.

Pitchford, W.S., Mirzaei, H.M., Deland, M.P.B., Afolayan, R.A., Rutley, D.L., and Verbyla, A.P., (2006) *Aust. J. Exp. Agric*. **46**:225.

Reverter, A., Johnston, D.J., Perry, D., Goddard, M.E., and Burrow, H.M. (2003) *Aust. J. Agric. Res*. **54**: 119.

Robertson, A., (1959) *Biometrics*, **15**:469.

QTL MAPPING IN MULTIPLE FAMILIES USING LOGISTIC REGRESSION

Y. D. Zhang and B. Tier

Animal Genetics and Breeding Unit*, University of New England, Armidale NSW 2351

SUMMARY

This study compares logistic regression (LR) with maximum likelihood (ML) methods for mapping quantitative trait loci (QTL) in multiple half-sib families under selective or full genotyping strategies, with various levels of marker informativeness and marker interval. In ideal conditions involving evenly located polymorphic markers and all individuals genotyped, both LR and ML methods showed a high power of detecting QTL and produced accurate estimates of QTL locations and effects. Under selective genotyping strategy, the power of ML is limited by regions with low information content. The LR method performed better than ML and is a straight-forward and robust method for this case.

INTRODUCTION

There are many methods available for detecting QTL in half-sib populations when all individuals are genotyped (eg Zeng 1993, 1994; Kao *et al.* 1999; Kerr *et al.* 2005). These methods generally rely on good genetic information, that is, polymorphic markers spaced evenly at close intervals along the chromosome. This ideal case is uncommon in real genome scans, because of a lack of informative markers in some chromosomal regions. Selective genotyping is used to reduce the cost of genotyping and achieve a reasonable power of detecting QTL (Lebowitz *et al.* 1987). Low amounts of genetic information, selective genotyping or a spike in the phenotypic distribution can generate spurious QTL peaks (eg Broman 2003; Feenstra and Skovgaard 2004). In a previous study, the logistic regression (LR) method effectively detected QTL in single half-sib families and avoided the spurious QTL peaks (Zhang and Tier 2005). Using simulated data, this study compares LR with the maximum likelihood (ML) methods for detecting QTL in multiple families with various levels of marker informativeness, marker interval and under two types of genotyping schemes (full or selective).

MATERIALS AND METHODS

Data. Ten datasets were simulated and analysed to examine the performance of LR and ML methods (five for each genotyping scheme). For each dataset one chromosome and 4 half-sib families (two families without QTL, the other two with QTL segregating but their effects in opposite directions) were simulated. Each progeny had a different dam. Genetic information for full genotyping datasets (1 to 5) and selective genotyping datasets (S1 to S5) is shown in Table 1. Residuals were drawn from a normal distribution with a mean of zero and variance of 10. The QTL effect (difference between two homozygous genotypes) was assumed to be half of the residual standard deviation (1.6). For selective genotyping, 100 individuals in each single family were simulated and the 25 individuals with highest and 25 with lowest phenotypic values were genotyped. Fifty individuals per family were simulated and genotyped in the full genotyping cases. As a result, 200 genotyped individuals in each dataset were analyzed. For each case, 100 replicates were generated and analyzed using both LR and ML methods. The chromosome was searched in steps of 5 cM. A chromosome wide significant threshold value was obtained using 300 or 500 permutations across each test point using LR or ML, respectively. Genotypes in each family were shuffled. The power was defined as a percentage of replicates in which a chromosome wide

* AGBU is a joint venture of NSW Department of Primary Industries and University of New England.

significance level of 5% was achieved. The QTL position was identified by the test location with the highest test statistic. The average QTL position and effect were calculated using all replicates where a 5% chromosome wide significance was achieved.

Table 1. Genetic information used in each multiple family dataset, number (N), position (cM) and information content range (MIC) of markers, minimum MIC with position in parenthesis. Simulated QTL position is also shown for full and selective genotyping cases

Dataset	N	Marker position (cM)	QTL position	MIC range	Min MIC (cM)
1, S1	6	0, 26, 51, 77, 102, 128	64	0.90-1.0	0.75
2, S2	6	0, 18, 30, 76, 102, 127	64	0.90-1.0	0.65(54)
3, S3	6	0, 26, 51, 97, 109, 127	64	0.90-1.0	0.65(74)
4, S4	5	0, 81, 106, 132, 157	93	0.90-1.0	0.33(40)
5, S5	5	0, 81, 106, 132, 157	93	0.75-0.85	0.29(38)

Marker information content. Paternal haplotypes were estimated using the Lander-Green algorithm (Lander and Green 1987). Prior QTL transmission probabilities (TP) for each individual at test points were estimated using genotypes and the recombinant fragment of flanking markers, as described by Kerr *et al*. (2005). The variation of TP is used as an indicator of the marker information content (MIC), which is 4 times the TP variance, ranging from 0.0 to 1.0.

Statistical model. The multiple-family approach integrates the single half-sib family analyses. The model for the single family analysis is the same as described in the previous study (Zhang and Tier 2005). Assuming that a bi-allelic QTL with alleles Q and q contributing to the variance of a quantitative trait. The quantitative trait value y_i for individual *i*, being pre-adjusted for fixed effects and polygenic effects, can be related to the QTL by the model $y_i = \mu + \alpha x_i + e$, where α is the effect of the putative QTL, x_i is a probability that a progeny inherits allele Q from the sire, and e is the residual. The logistic regression model is fitted and implemented as described by Dobson (2002). For the multiple-family analysis, different QTL phases were considered for different families. For a family the phase was determined by the largest likelihood. The log likelihood of the multiple-family estimation was the sum of log likelihoods from each family with respect to its most likely phase. The estimated LR coefficient and the total variance were used to calculate the QTL effect α, as described by Henshall and Goddard (1999). The ML method was implemented as described by Kerr *et al.* (2005), the α value is estimated as described by Zeng (1994). In the case of selective genotyping, the α values from ML were adjusted for the consequences of selective genotyping, using the method of Darvasi and Soller (1992). Empirical threshold values were determined using permutation tests (Churchill and Doerge 1994).

RESULTS AND DISCUSSION

Results are presented for the two genotyping strategies separately to compare performance of LR and ML methods. Differences between the two strategies are not compared here, because the underlying simulated datasets, for example, 1 and S1, are not the same.

Full Genotyping. As shown in Table 2, under the ideal conditions involving evenly located polymorphic markers and all individuals genotyped (dataset 1), both LR and ML showed a high power of detecting QTL and produced accurate estimates of QTL locations and effects, although QTL effects estimated using LR were slightly lower than the simulated value. The results for ML

methods are consistent with findings by Kerr *et al.* (2005), who also reported that a 100% detection power was achieved when sib-family size was 100. Although all animals were genotyped, when the marker genetic information was low (at various levels in datasets 2 to 5), power decreased dramatically, particularly for ML. Though the estimates of QTL effects were very conservative for both LR and ML methods the estimated QTL positions were close to that simulated.

Table 2. Results of QTL analyses using Logistic Regression and Maximum Likelihood methods in the full and selective genotyping schemes. The power of each analysis, QTL position (cM) and effect (simulated at 1.6) are presented (standard error in parenthesis)

Dataset	Logistic regression			Maximum likelihood		
	Power[1]	Position	Effect	Power	Position	Effect
			Full genotyping			
1	84	62	1.39(0.04)	87	63	1.64(0.06)
2	31	71	1.21(0.10)	27	70	1.36(0.11)
3	27	57	1.17(0.10)	18	56	1.21(0.13)
4	28	106	1.10(0.08)	24	95	1.47(0.13)
5	31	101	1.17(0.09)	21	94	1.42(0.13)
			Selective genotyping			
S1	44	59	1.64(0.08)	42	59	1.49(0.11)
S2	36	68	1.60(0.10)	39	65	1.63(0.10)
S3	38	55	1.54(0.09)	37	63	1.68(0.12)
S4	50	96	1.66(0.08)	37	88	1.85(0.14)
S5	51	97	1.48(0.08)	34	53	2.73(0.14)

[1] percentage of 100 replicates with a 5% chromosome wide significance level.

Selective Genotyping. Both LR and ML methods showed a similar power, QTL position and effects when markers were highly informative and evenly spaced (S1). When an interval containing the QTL was extended to 57 cM, QTL positions assessed by both methods also did not deviate very much from the true position when marker genetic information was not very low (for example, minimum MIC was 0.65 in datasets S2, S3). When a large marker interval (up to 80 cM) occurred next to the interval containing the QTL, LR performed better than ML in the positioning QTL and the power to detect QTL (S4). When marker information content was ever lower (0.6 – 0.8, 0.29 at the lowest point) and with a large marker interval neighboring the QTL region, LR also produced better estimates of QTL location and effect than those from ML (S5) and displayed high power. In this adverse case, ML is likely to produce spurious QTL peaks, far away from the true position and towards the region with low genetic information, and with an inflated estimate of the QTL effect (2.73 for dataset S5).

Chromosomal regions of low MIC frequently occur in QTL projects. Those regions are attributable to a lack of informative markers. ML can produce spurious peaks of test statistic in low information regions under selective genotyping. Feenstra and Skovgaard (2004) developed a

2-component mixture model to avoid such spurious peaks and applied to a back-cross design. However, to apply this method to F2 or half-sib design, a 3-component model is required and is yet to be developed. The LR method can overcome such problems and facilitate the efficient application of selective genotyping. In the previous study on a single family (Zhang and Tier 2005), both LR and ML methods displayed higher power than in this multiple family approach, suggesting this approach could effectively reduce the positive rate.

CONCLUSION

In the ideal conditions involving evenly located polymorphic markers and all individuals genotyped, both LR and ML methods showed a high power of detecting QTL and produced accurate estimates of QTL locations and effects. Under selective genotyping, the power of the ML method is limited in low information regions, being very likely to produce spurious QTL peaks. In this case, LR performed better than ML in detecting QTL. The LR method provides a straightforward and robust solution for this situation.

REFERENCES

Broman, K. W. (2003) *Genetics* **163**:1169
Churchill, G. A. and Doerge, R. W. (1994) *Genetics* **138**:963.
Darvasi, A. and Soller, M. (1992) *Theor. Appl. Genet.* **85**:353.
Dobson, A. J. (2002) "An Introduction to Generalized Linear Models" 2nd ed. Chapman and Hall, London.
Feenstra, B. and Skovgaard, I. M. (2004) *Genetics* **167**:959.
Henshall, J. M. and Goddard, M. E. (1999) *Genetics* **151**:885.
Kao, C. H., Zeng, Z. B. and Teasdale, R. D. (1999) *Genetics* **152**:1203.
Kerr, R. J., McLachlan, G.M. and Henshall J.M. (2005) *Genet. Sel. Evol*. **37:**83.
Lander, E. S. and Green, P. (1987) *Proc. Natl. Acad. Sci. USA* **84**:2363.
Lebowitz, R. J., Soller, M. and Beckmann, J. S. (1987) *Theor. Appl. Genet.* **73**:556.
Zeng, Z. B. (1993) *Proc. Natl. Acad. Sci. USA* **90**:10972.
Zeng, Z. B. (1994) *Genetics* **136**:1457.
Zhang, Y. D. and Tier, B. (2005) *Proc. Assoc. Advmt. Anim. Breed. Genet.* **16**:354.

CATTLE RESIDUAL FEED INTAKE CANDIDATE GENES

N.A. Zulkifli, M. Naik, W.S. Pitchford and C.D.K Bottema

CRC for Beef Genetic Technologies
School of Agriculture, Food and Wine, The University of Adelaide,
Roseworthy SA 5371, Australia

SUMMARY

One option that can be used to select cattle for improving feed efficiency is to select for genes affecting residual feed intake. Previous studies using two groups of cattle, the Trangie Angus residual feed intake (RFI) selection line cattle and the Davies Jersey x Limousin gene mapping cattle have identified five chromosomal regions or QTL that have significant effects on net feed efficiency. These regions are on cattle chromosomes 1, 6, 8, 11 and 20. Recent work has also implicated mitochondrial function as being important in net feed efficiency in livestock. Using information from these chromosome regions and literature on mitochondrial function, candidate genes were selected for net feed efficiency. By sequencing the candidate genes, DNA variants including single nucleotide polymorphisms (SNPs) and insertion and deletions (in/dels) were detected. The DNA variants will be genotyped in the progeny for association studies.

INTRODUCTION

Residual feed intake is a measure of feed efficiency which is an economically important trait in livestock. Residual feed intake is affected by many factors including both diet and genetics. Residual feed intake of an animal depends on the ability of the animal to consume less feed than expected based on their weight gain and weight maintained during the feed testing. This occurs by improving the utilisation of nutrients and energy from the feed for maintenance and growth. Recent work has implicated mitochondrial function as being important in net feed efficiency in livestock (Naik 2007). It is well known that genetics and diet have profound influence on mitochondrial function (Bottje *et al.* 2002).

Mitochondria are the site of energy production in the cell and produce the majority of the cellular ATP (Kolath *et al.* 2006). The electron transport chain in the mitochondria is a sequence of electron carrier molecules that shuttle electrons during the redox reactions that release energy used to make ATP. Most components of the chains are proteins, which exist in multi-protein complexes numbered I through IV. During electron transport along the chain, electron carriers alternate between reduced and oxidized states as they accept and donate electrons. However, mitochondrial inefficiency may occur as a result of electron leakage from the chain. As a consequence, 2 to 4% of oxygen consumed by mitochondria may be incompletely reduced to reactive oxygen species (ROS) rather than being completely reduced to water, due to the univalent reduction of oxygen by the electrons (Boveris *et al.* 1973).

These reactive oxygen species produced in the mitochondria are very destructive. Reactive oxygen species can cause oxidative damages to nucleic acids, lipids and proteins, as well as damaging organelles such as the mitochondria itself (Nelson *et al*. 2008). Thus, the reactive oxygen species themselves can cause the mitochondria to function less efficiently and produce even more ROS.

Studies in chickens have shown that mitochondria obtained from chickens of low feed efficiency exhibit greater uncoupling of the electron transport chain (Bottje *et al*. 2006). It was also observed that there is a higher level of reactive oxygen species production in the mitochondria of the low feed efficiency chickens. Similar studies in chickens revealed that there is a greater electron leakage in mitochondria from low feed efficiency chickens in comparison to

mitochondria from high feed efficiency chickens (Bottje *et al*. 2002).

Hence, the objectives of this study were to identify genes involved in mitochondrial function and the regulation of ROS which may affect feed efficiency in cattle. Several QTL affecting net feed efficiency were mapped in Jersey x Limousin backcross progeny (Naik 2007). Candidate genes within these QTL that are involved in mitochondrial function or the reactive oxygen species regulation were selected and screened for DNA variants that might be used as DNA markers for selecting animals of high feed efficiency.

MATERIALS AND METHODS

Materials. Genotype and phenotype data from Davies cattle gene mapping project were used in this study (Sellick 2007). The breeds used were Limousin, a beef breed with a moderately large frame, and Jersey, a dairy breed with a small frame. Since the breeds are phenotypic extremes, the trait variation in the progeny of their crosses will be maximized. In 1993, the first phase on this study was conducted by mating purebred Jersey and Limousin to produce their first cross progeny. The first cross progeny were born in 1994 and 1995. In the second phase, three Limousin x Jersey F_1 sires were mated to pure Jersey and Limousin dams in Australia and New Zealand to produce double backcross animals, namely Limousin cross progeny and Jersey cross progeny. There were 161 Limousin cross progeny and 205 Jersey cross progeny in Australia.

Methods. The sequence of the candidate genes were obtained from the Bovine Genome project database: Ensemble (www.ensemble.org) and Biomanager (www.angis.org). Primers were designed using Primer3 software and tested against OLIGO 4.04 software to avoid hairpin structures and primer dimer formation and to minimize the GC content. The primers designed were then optimised for PCR. Once the appropriate conditions were obtained for each pair of primers, genomic DNA from the three Davies F_1 sires were amplified (Palm Cycler, Corbett Research). The genomic DNA was extracted from whole blood using the Jetflex Genomed genomic DNA purification kit (Astral Scientific) following the manufacturer's protocol (Sellick 2007). The amplification was followed by gel electrophoresis for product size confirmation. The amplified PCR products were then purified using PCR purification kit (DNA Purification Kit, Mo Bio Laboratories) to remove excessive primers and dNTPs. After obtaining the purified PCR products, the DNA was sequenced using BigDye terminator cycle sequencing (Applied Biosystems). The sequencing reactions were conducted using BigDye terminator ready reaction mix, 25 pmol of primer and 30-100 ng purified PCR products. This was followed by analysing the sequencing products using an Applied Biosystems 3730 DNA Analyser at Institute of Molecular and Veterinary Science (IMVS), Adelaide.

RESULTS AND DISCUSSION

The selection of candidate genes (Table 1) was based on the results of QTL mapping. The selection was also supported by previous mitochondrial proteomics studies (Naik 2007). Most of the genes selected are involved in the mitochondrial electron transport chain. It is crucial that the electron transport chain works effectively, as excessive (ROS) production may result in oxidative stress in the cell. Catalase has a major role in controlling the level of ROS by converting the superoxides to water and oxygen. So although *catalase* (CAT) gene was not located in a highly significant QTL, it was selected based on function and the mitochondrial proteomics studies.

The coding regions, including all exons and untranslated regions (5' and 3'), of the genes were sequenced in 3 F1 sires whose progeny were linkage mapped for RFI. DNA variants are not commonly found in the coding regions of genes. In order to ensure that DNA markers were

found for each gene, the introns flanking the exons and the 5' and 3' regions flanking the genes were also partially sequenced.

Table 1. Selected candidate genes and their function.

Candidate Genes	Chromosome	Function
Complex I-SGDH, 75kDa (NDUFB5)	BTA 1	Involved in electron transport chain
Superoxide dismutase 1, soluble (SOD1)	BTA 1	Involved in binding copper and zinc ions and destroying free superoxide radicals
Aldolase B (ALDOB)	BTA 8	Involved in fructose metabolism
Adenylate kinase (AK1)	BTA 11	Involved in maintaining cellular energetic economy
Complex I-19 kDa (NDUFA8)	BTA 11	Involved in electron transport chain
Hydroxyacyl co-enzyme A dehydrogenase (HADHB)	BTA 11	Involved in synthesising mitochondrial trifunctional protein
Succinyl Co-A synthetase (SUCLG1)	BTA 11	Involved in generating high energy phosphate
Catalase (CAT)	BTA 15	Involved in reactive oxygen species (ROS) metabolic pathway

Of the 8 candidate genes, the coding regions of five genes have been completely sequenced in the three F1 sires. The majority of the DNA variants identified were SNPs but there were some in/dels (Table 2). Most of the DNA variants either occurred in the introns or were silent mutations in exons. These are, therefore, not likely to affect the function of the protein.

Table 2. DNA variants found in candidate genes.

Gene	# Exons	# Exons sequenced	DNA variants	Potential functional DNA variants
CAT	13	13	3 SNPs (2 exonic, 1 intronic)	
HADHB	16	16	4 SNPs (2 intronic, 2 exonic)	stop codon (exon 4)
ALDOB	9	9	3 SNPs (2 exonic, 1 intronic)	alanine ➔ threonine (exon 8)
AK1	6	5	1 SNP (intronic), 1 in/del (intronic)	
SUCLG1	9	9	3 SNPs (1 exonic, 2 intronic), 1 in/del (intronic)	
NDUFB5	8	2	6 SNPs (intronic)	
NDUFA8	4	1	1 SNP (exonic)	serine ➔ tyrosine (exon 3)

However, several potentially functional SNPs were also discovered. In the *aldolase B* gene (ALDOB), three DNA variants were detected, including one variant that changes an alanine amino acid residue to a threonine in exon 8. In the *complex I - 19 kDa* gene (NDUFA8), a non-conservative amino acid substitution of serine for tyrosine was found in exon 3. In the *hydroxyacyl co-enzyme A dehydrogenase* gene (HADHB), four DNA variants were discovered including a stop codon in exon 4.

Thus, SNPs that cause either a missense or a nonsense mutation have been found in 3 candidate genes thus far. The genotyping of the DNA variants is currently underway. In addition, the exons in the remaining candidate genes are being sequenced. Apart from the genotyping, mitochondrial biochemical assays will be used to estimate oxidative phosphorylation activity in animals of high and low efficiency. By combining the results from these studies, we hope to increase our understanding of net feed efficiency and its relationship with mitochondrial function and explore the potential for dietary manipulation to improve feed efficiency in cattle..

ACKNOWLEDGMENTS

We would like to thank the J.S. Davies Bequest for funding the residual feed intake QTL mapping project at the University of Adelaide. We would also like to thank the Ministry of Higher Education (MOHE), Malaysia and the National University of Malaysia (UKM), Malaysia for providing scholarships for N. Zulkifli.

REFERENCES

Bottje, W., Pumford, N.R., Dirain, C.O., Iqbal, M., and Lassiter, K. (2006) *Poult. Sci.* **85**:8.

Bottje, W., Tang, Z.X., Iqbal, M., Cawthon, D., Okimoto, R., Wing, T., and Cooper, M. (2002) *Poult. Sci.* **81**:546

Boveris, A., and Chance, B. (1973) *Biochem. J.* **134**:707

Kolath, W.H., Kerley, M.S., Golden, J.W., and Keisler, D.H. (2006) *J. Anim. Sci.* **84**:861

Naik, M. (2007) PhD Thesis, University of Adelaide University, Australia

Nelson, D. L., and Cox, M.M. (2008) "Lehninger Principles of Biochemistry" 5th ed. W.H. Freeman and Company, New York

Sellick, G.S., Pitchford, W.S., Morris, C.A., Cullen, N.G., Crawford, A.M., Raadsma, H.W., and Bottema, C.D.K. (2007) *Animal Genetics*. **38**:440

AUTHOR INDEX