ASSOCIATION FOR THE ADVANCEMENT OF ANIMAL BREEDING AND GENETICS

Proceedings of the Twentieth Conference

Translating Science into Action

Napier, New Zealand

20th October – 23rd October 2013

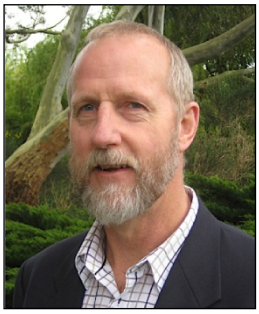
ISSN 1328-3227
ISBN 978-0-473-26056-9

Produced by:
Association for the Advancement of Animal Breeding and Genetics
C/- AAABG Secretary
Massey University
Palmerston North
New Zealand

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



A very warm welcome to the 20th AAABG meeting in Napier, New Zealand. This is the third time New Zealand has hosted the Biennial meeting, and we trust it will be as enjoyable as the previous occasions.

The theme for the conference is "Translating Science into Action". There have been similar themes for previous meetings. In 1997 we had "Breeding … responding to client needs", and in 1990 "Technology prospects and transfer". This should be unsurprising given that the third of our Association Objectives is "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". The primary point of the plenary session devoted to the topic in this meeting is to recognise that in this age of information overload, we might need to pause and think about how we provide new information to end-users. The electronic media are revolutionising our ability to transmit information to potential users and we may think that is the scientist's or extension officer's job done – the information transfer box is ticked. However, we have limited information to tell us how to best present information to cause behavioural change in our target farmer / rural professional community. There is likely a fertile research patch to address this gap in knowledge through joint research with social scientists and educationalists.

Further to the translating science into action theme, a Sheep Breeders' Day will take place on Thursday 24 October 2013 at the Hawke's Bay A&P Showgrounds. This is the 150th Show and it has been granted "Royal Event" status to recognise the importance of the occasion. The A&P Association is making a major effort to ensure that a suitably grand event transpires and it will be fantastic to have the AAABG represented at the Show.

Hawkes Bay has a vibrant primary sector with sheep, beef cattle, dairy and deer all featuring prominently in the economy, and the region is home to a good number of ram and bull-breeding operations. Hawkes Bay is also well known for its horticultural industries with several top wineries, many orchards and a variety of food crops such as squash.

As any event organiser knows, sponsorship helps to oil the financial wheels. We have been fortunate in getting strong support from several sponsors for which we are very grateful. The sponsors are recognised on a following page in these proceedings.

We trust that you will have an enjoyable time improving your knowledge in Animal Breeding and Genetics, and that you will enjoy the camaraderie of other attendees as you learn something about Napier and Hawkes Bay.

Hugh Blair
President, AAABG

ASSOCIATION FOR THE ADVANCEMENT OF ANIMAL BREEDING AND GENETICS

TWENTIETH CONFERENCE

COMMITTEE

President	Hugh Blair
President-Elect	Mike Goddard
Vice President	Geoff Nicoll
Secretary	Rebecca Hickson
Treasurer	Patrick Morel
Members	Rao Dukkipati
	Many New Zealand AAABG Members
Editor	Nicolas Lopez Villalobos
Proceedings Production	Bill Pattie
	Barrie Restall
	Sarah Siebert
Professional Conference Organiser	Sarah Siebert
	Massey University

CITATION OF PAPERS

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

For example:

Bowley F.E., Amer P.R. and Meier S. (2013) New approaches to genetic analysis of fertility traits in New Zealand dairy cattle. *Proc. Assoc. Advmt. Anim. Breed. Genet.* **20**: 37-40.

REVIEWERS

All papers, invited and contributed, were subject to peer review by two referees. The following people are acknowledged and thanked for their work in reviewing the papers.

Mary Abdelsayed
Daniel Abernethy
Hawlader Al-Mamun
Peter Amer
Neville Amyes
Jason Archer
Paul Arthur
Kevin Atkins
Benoit Auvray
Wendy Bain
Alex Ball
Rob Banks
Bill Barendse
Steve Barwick
Rolf Beilharz
Donagh Berry
Samantha Bickell
Matt Bixley
Dominique Blache
Hugh Blair
Vinzent Boerner
Sunduimijid Bolormaa
Cynthia Bottema
Frances Bowley
Forbes Brien
Daniel Brown
Kim Bunter
Heather Burrow
Timothy Byrne
Yizhou Chen
Sam Clark
Bronwyn Clarke
Schalk Cloete
Rene Corner
David Cottle
Neil Cullen
Brian Dalrymple
Hans Daetwyler
Yvette de Haas
Dirk-Jan De Koning
Dion Detterer
Cathy Dodd
Kath Donoghue
Rao Dukkipati
Zoey Durmic
Sandra Eady
Sophia Edwards
A Engelbrecht
Majbritt Felleki
Mark Ferguson
Ester Flores
Neal Fogarty
Marina Fortes
Jose Garcia-Muñiz
Dorian Garrick
John Gibson
Sam Gill
Arthur Gilmour
Mike Goddard
Cedric Gondro
Oscar Gonzalez-Recio
Tom Granleese
Johan Greeff
Boyd Gudex
Boyd Gudex
Mekkonen Haile-Mariam
Bevin Harris
Sue Hatcher
Ben Hayes
Fiona Hely
John Henshall
Robert Herd
Susan Hermesch
Cord Heuer
Rebecca Hickson
Janelle Hocking Edwards
Nicholas Hudson
Janusz Jamrozik
Gilbert Jeyaruban
David Johnson
Patricia Johnson
David Johnston
Elisabeth Jonas
Neville Jopson
Haja Kadarmideen
Mohammad Kalaldeh
Michael Keehan
Kathryn Kemper
Paul Kenyon
Mehar Khatkar
James Kijas
Brian Kinghorn
Konstantin Konstantinov
Michael Lee
Li Li
Nicolas Lopez Villalobos
Nils Lundeheim
Iona MacLeod
Daniel Martin-Collado
John McEwan
David McGill
Brian McGuirk
Bill McKiernan
Karin Meyer
Filippo Miglior
Nasir Moghaddar
Patrick Morel
Steve Morris
Sue Mortimer
Carel Muller
Gert Nieuwhof
Willem Olivier
Peter Parnell
Natalie Pickering
Laurie Piper
Wayne Pitchford
Laercio Porto Neto
James Preston
Jennie Pryce
Ian Purvis
Herman Raadsma
Imtiaz Randhawa
Antonio Reverter
James Rowe
Lise Sandenbergh
Bruno Santos
Anthony Schlink
Chris Schrooten
Ross Shepherd
Richard Sherlock
Jude Sise
Jen Smith

Richard Spelman
Katarzyna Stachowicz
Andrew Swan
Imke Tammen
Ross Tellam
Carel Teseling
Peter Thomson
Bruce Tier
Julius van der Werf
Alison Van Eenennaam
Ilse van Grevenhof
Japie Van Wyk
Carlos Vazquez
Klara Verbyla
Phil Vercoe
Steve Walkden-Brown
Eileen Wall
Bradley Walmsley
Shalanee Weerasinghe
Brian Wickham
Anne Winkelman
Matthew Wolcott
Rob Woolaston
Mark Young
Emily Young
Yuandan Zhang

SPONSORS OF THE 20TH AAABG CONFERENCE

The financial assistance of the following organisations is gratefully acknowledged.

PLATINUM SPONSOR

SILVER SPONSORS

Coffee Lounge Sponsor

Internet Café Sponsor

BRONZE SPONSORS

Writing Materials Sponsor

Supporting Student Sponsors

Concurrent Session 2.1 Sponsor

Conference Satchel Sponsor

Opening Session Sponsor
Plenary Session 1.1

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

Elected September 2011

B.P. Kinghorn
A. McDonald

Elected October 2013

H. Burrow
P. Fennessy
G. Nicoll
P. Parnell

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

HEATHER BURROW

Heather was raised in the bush and began her science career by undertaking a BA at UNSW (1976) in conjunction with Prince Henry and Prince of Wales Hospitals as part of a pilot graduate nursing degree. Through a convoluted series of events she ended up working for CSIRO at Rockhampton. In 1991 she completed a Postgraduate Certificate in Animal Breeding and Genetics at UNE and in 1998 she was awarded her PhD from UNE. During the past 10 years she has completed multiple courses with the Australian Institute of Company Directors and an MBA (UNE) where she was commended for her high level of achievement.

Heather is an incredibly committed and tireless worker for her beloved Northern Australian Beef Industry and has been so for over 30 years. She initiated, conducted and led ground-breaking research on temperament of beef cattle and its relationships with economically important attributes (1986-2003), an issue that has become a major focus for research groups worldwide over the past decade because of its impact on animal welfare, on-farm productivity and beef quality. During the first Beef Cooperative Research Centre (1993-1999) she led the major Northern Breeding component. During the second Beef CRC (1999-2005), Heather continued leadership of the Northern Program and was Deputy Chief Executive Officer. Also during this time (1999-2008) she led an ACIAR project with some of the world's most disadvantaged farmers in South Africa, enabling them to become commercially-oriented beef producers trading through commercial beef value chains. In 2004 she was appointed CEO and led the bid for the third CRC and left CSIRO as a Senior Principal Research Scientist and. This was successful (2005-2012) and during the early stages Heather led the process of transitioning from an unincorporated joint venture to an incorporated company with a transformed board and management committee structure.

Heather has an exceptional record of communication to scientific and beef industry audiences through oral and written presentations as a regular invited presenter at scientific, government and industry forums in Australia and internationally. She is author or co-author of ~190 refereed journal and conference papers, five book chapters and a producer-oriented technical manual. She also provided leadership for delivery of six Special Editions in two scientific journals, including initiation and advocacy of the concept with the publishers. Heather has been on the committee of at least two AAABG conferences including President in 2005 and Chairman of a very successful Genomics Conference in 2011. Heather's passion, ability to communicate to a wide range of people, embracing of management of science and work ethic is inspirational. Thank you for your contribution to our Beef Cattle Industry and our Association. You are a very worthy Fellow.

PETER FENNESSY

Peter Fennessy graduated from the then Lincoln College with a B Agr Sc (1969) and an M Agr Sc (Hons) in 1971. He worked as a research scientist at the Ministry of Agriculture and Fisheries Invermay Research Centre for a year before commencing a PhD programme at the Waite Research Institute which he completed in 1976. Peter then returned to Invermay Research Centre where he worked on a wide range of programmes as a research scientist. In 1992, the government established AgResearch as a Crown-Owned Research Institute, and Peter took the role of General Manager of its Sheep, Deer and Equine Division for five years. After leaving AgResearch in 1997, Peter established his own business consulting to the agricultural and biotechnology industries, and went on to establish Abacus Biotech Ltd (now AbacusBio Ltd) in 2001, mainly with a group of his Invermay colleagues. Peter was Managing Director of AbacusBio through its start-up period until 2010, after which he stepped aside to return to the consulting work he enjoys so much.

Peter has an incredible passion for science, an encyclopaedic knowledge of a wide range of fields and the energy and ability to make science and technology work in businesses. His many interests have included nutrition and metabolism, intake and growth, biological efficiency, genetics of growth and carcass traits, selection and breeding, deer antlers, pubertal and seasonality traits, photoperiodic manipulation of growth and reproduction, artificial reproductive technologies, evolution and its implications in farmed species, interspecies hybrids, gene mapping and quantitative traits. This work has been undertaken mainly in sheep and deer and perhaps not surprisingly given his interests in thoroughbreds, also in horses. These efforts have contributed enormously to, and significantly underpinned, the growth of the deer and sheep farming industries in New Zealand. Peter's contribution to the New Zealand sheep and deer industries was recognised in 1990 with the awarding of the New Zealand Society of Animal Production's McMeekan Memorial Award for an outstanding individual contribution to animal science in New Zealand.

Amongst his many achievements, Peter established the Invermay Lean and Fat Coopworth selection lines in 1979 to address the "overfat" lamb problem. The selection lines were used to understand the genetics of carcass composition, in particular to investigate responses to selection for leanness in the sheep industry. This contributed to the rapid uptake of ultrasound scanning to improve carcass composition in the sheep industry. These lines were subsequently used in a large scale programme to detect quantitative trait loci for growth and meat quality traits.

Peter was a driving force in the creation of the AgResearch Molecular Biology Unit (now the AgResearch Genomics Group), which was a joint venture between AgResearch and Otago University. The Molecular Biology Unit developed the first map of the sheep genome using microsatellite markers, which was used in a number of large scale QTL experiments. Peter was also a key person in the establishment of Ovita, the Sheep biotechnology research consortium that commenced in 2001. Peter was a director and a key person behind the scenes in setting up the consortium, which has resulted in eight commercial DNA tests for production traits and genomic predictions for 28 traits including facial eczema in a variety of New Zealand sheep breeds.

In addition to his research accomplishments, Peter has also had a huge number of interactions with commercial farmers, breeders and farm businesses with a strong focus on how to make science work on farm.

For his outstanding contributions to the science of genetics and animal improvement The Association for the Advancement of Animal Breeding and Genetics is pleased to enrol him as a fellow of the Society.

GEOFF NICOLL

Geoff Nicoll graduated with a BAgrSc degree from Massey University in 1972. He followed this with an MAgrSc(Hons) degree (1975) under the tutelage of Professor Al Rae while at the same time working as a Junior Lecturer in the Sheep Husbandry Department. Upon leaving Massey, Geoff took up a farm advisory position and then became a research scientist before travelling to Grange, Ireland, to undertake his PhD with the National University of Ireland. Geoff returned to New Zealand in 1981 and was seconded from the Lands and Survey Department to work as a scientist at Whatawhata Hill Country Research Station, and subsequently at MAF's Ruakura Agricultural Centre. In 1987, Geoff became Head of the Genetics and Nutrition Unit for Landcorp Farming Ltd. In 2011, the Landcorp Genetics operation merged with Rissington Breedline to form a joint venture company called Focus Genetics. Geoff has recently announced his retirement.

As Head of Landcorp's Genetics and Nutrition Unit, Geoff was responsible for (amongst other roles) the scientific and technical integrity of Landcorp's sire-breeding programmes. This operation provided seedstock for Landcorp's considerable animal resources, which included about 600,000 ewes, 70,000 beef cows and 65,000 hinds. This necessitated the management of 11 sire-breeding programmes involving some 25,700 fully performance-recorded animals in 17 individual flocks and herds.

Geoff made major contributions in overseeing the introduction of new technologies into the Landcorp breeding schemes including use of DNA markers, CT scanning and the formation of new composites such as Landcorp Lamb Supreme and Landcorp Landmark. Not all explorations went as successfully as these prior examples, but for those pushing the boundaries of technology, the occasional misfire is bound to happen. With Geoff's animal breeding and genetics expertise the application of CT scanning achieved 30% greater genetic gain in weight of meat in the carcass, while at the same time halving the weight of fat and doubling the improvement in eye muscle area, compared with using ultrasound-based selections. Another of Geoff's successes was the importation of semen from rams carrying a major gene for muscling (Carwell). Research by staff at Landcorp and AgResearch demonstrated that animals expressing the gene had approximately 8% more muscle weight in the loin, with no significant effect elsewhere in the carcass, when compared to contemporaries that did not express the gene. Follow-up research located the gene responsible. Geoff was a major contributor at all stages in this work including the management of the genetic resource animals, the design and implementation of the experimental programme and as Landcorp's representative on the management committee for the programme. It should be clear from these examples that Geoff added significant value to an already well-oiled genetic improvement programme. While the initial intent was to only provide seed-stock for Landcorp farms, surplus animals have been made available to the wider industry, enabling others to benefit from Geoff's successes.

Geoff's advice on genetic improvement has been regularly sought by industry breeders both within New Zealand and offshore, and he was regularly invited to speak at international meetings on the implementation of genetics and breeding theory to on-farm genetic improvement. Geoff has a strong publication record, with the majority of these being technical reports and conference papers which spoke to his primary audiences. Geoff published his first AAABG paper in 1987 and has been a regular contributor ever since. Geoff was a member of the Eighth AAABG Committee in 1990 and is currently AAABG Vice President. He has also contributed significantly to other professional societies including the New Zealand Society for Animal Production

(Secretary 1983 to 1985; President, 1994) and the Asian-Australasian Association of Animal Production (Secretary-General 1985 to 1987).

Dr Geoff Nicoll has made an outstanding contribution to the New Zealand livestock industry, not only directly through his leadership in the application of genetic principles to benefit Landcorp's livestock, but also through his support for good science and his production of scientific papers. His ability to develop and translate science in a commercial and practical manner is exceptional.

For Geoff's major contribution to the application of animal breeding and genetics to livestock industries in New Zealand, the Association for the Advancement of Animal Breeding and Genetics is pleased to elect him as a Fellow of the Association.

PETER PARNELL

Dr Peter Parnell has made an outstanding contribution to research and application of genetic technologies in Australia.

Peter graduated from the University of New England in 1981 with a Bachelor of Rural Science with 2nd Class Honours. He went on to do a PhD at the Animal Genetics and Breeding Unit under the supervision of Professor Stewart Barker and Dr Keith Hammond, completing his thesis in 1988.

From March 1985 to April 1995 Peter worked as a Senior Research Officer for the NSW DPI leading the research project into Feed Efficiency of high and low feed efficiency lines of Angus cattle. He established Net Feed Efficiency as the measure of Feed Efficiency in these selection lines.

From April 1995 to April 1996 Peter was employed as the breed development officer of the Murray Grey Breed Society and introduced a scientific approach to this breed society.

From April 1996 to July 2001 Peter was the Breed Development Manager for Angus Australia during which he developed a very strong performance based ethos for the society.

From August 2001 until April 2009 Peter was the Centre Director of the NSW DPI Beef Industry Centre of Excellence based in Armidale. From July 2004 to April 2009 he was NSW DPI's Research Leader of Beef Genetics and Improvement.

During this time Peter played a key role in the design of the Maternal Productivity Project, which was a major program in the Co-operative Research Centre for Genetic Technologies in Beef Cattle.

He was also the leader of the very successful CRC Accelerated Adoption Project which utilised "Beef Profit Partnerships" (BPPs) across Australia and New Zealand. BPPs were a system of partnerships between beef businesses, value chains and the broader beef industry designed to accelerate improvements, innovations and adoption for sustainable and quantifiable impact on business profit.

From August 2003 until April 2009, Peter was an Adjunct Associate Professor of the University of New England.

Since May 2009 Peter has been the Chief Executive Officer of Angus Australia with a strong input into the applications of genetic technologies to the Australian cattle breeding industry.

Peter is currently a consultant to the Southern Beef Technologies Service (SBTS), which is a joint venture between Meat and Livestock Australia, ABRI and 15 temperate cattle breed societies aimed at increasing the understanding and use of genetic technologies by beef cattle seedstock and commercial breeders.

Peter has authored 25 refereed research publications including a number of papers presented to AAABG Conferences and has presented over 70 conference papers.

Peter has a passion for the application of innovation into the beef industry. He has excellent communication abilities, which have allowed him to translate research outcomes into readily useable on-farm applications.

For his outstanding contributions to the genetic improvement of the Australian beef herd, the Association for the Advancement of Animal Genetics and Breeding is pleased to enrol him as a Fellow of the Association.

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 2005 Medal recipient, Dr. Keith Hammond, is reproduced in these proceedings.

Trustees of the Helen Newton Turner Trust are:

- Dr. Richard Sheldrake AM (Chairman), representing NSW Department of Primary Industries
- Professor Brian Kinghorn, representing the University of New England
- Mr. Scott Dolling, representing the Association for the Advancement of Animal Breeding and Genetics
- Dr Roly Neipier Representative of the National Farmers Federation
- Dr. Robert Banks, Director, Animal Genetics and Breeding Unit

MEDALLISTS

1994	J.W. James	*2003*	F.W. Nicholas
1995	L.R. Piper	*2007*	Lucinda Corrigan
1997	J. Litchfield	*2009*	J. Ryves Hawker
1998	J.S.F. Barker	*2011*	Robert Banks
1999	C.W. Sandilands	*2013*	MikeGoddard
2001	G.A. Carnaby		

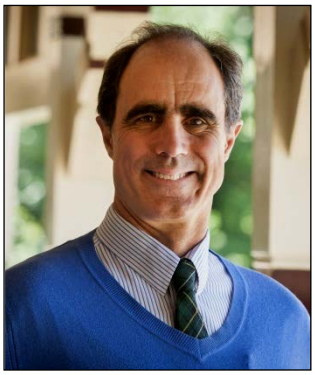
HELEN NEWTON TURNER AO

HELEN NEWTON TURNER MEDALLIST ORATION 2011

HOW CAN LIVESTOCK GENETIC IMPROVEMENT EVOLVE FASTER?

R.G. Banks

AGBU, University of New England, Armidale, NSW 2350



It is a great honour to be the recipient of this award, given in memory of Helen Newton Turner. I only the privilege of meeting Helen Newton Turner once, but her work informed my initial steps in the role of National LAMBPLAN Coordinator. Soon after commencing in that role, I discovered in the UNE library a report to Standing Committee of Agriculture in which she outlined the necessary key elements of a genetic improvement scheme for meat sheep in Australia. Her report was written in I think 1956, a long time before I started in LAMBPLAN, but all the fundamentals were there, and to some extent the development of LAMBPLAN followed her recommendations.

The livestock industries remain an important part of our economy, in particular our rural economy. The combined GVP is of the main livestock industries is at least $15bn pa on-farm, and the multiplier value through to domestic and overseas consumers many times that.

Breeding animals that suit Australian production systems and our markets has been a challenge since the first days of European settlement, but Australia has been blessed with active and talented animal breeders, and more recently researchers and extension personnel, and to have worked with many of them during the last 25 years has been a privilege.

In the period since the first awarding of the Helen Newton-Turner medal in 1994, there has been consolidation of national genetic evaluation systems for all major livestock species in this country, and in parallel, development and use of methods for direct genotyping. At the same time, we have with one or two exceptions achieved genetic improvement at similar rates to other countries and there is still considerable room for improvement in the effectiveness of our efforts in research, development and implementation, and in rates of improvement across all species.

What contributes to that performance gap, and what might be done about reducing it? I suggest that we need to examine a number of aspects of animal breeding research and its implementation in this country with the aim of re-balancing our efforts.

The following points are all aspects of how we think about and invest in animal genetic improvement. What we conclude on each, and how we respond, are important questions:

- Given that the principles of variation, selection and multiplication – the evolutionary algorithm – are completely confirmed, what is research?

Research by definition implies that answers are uncertain. However, this is not generally the case in the broad sense in animal breeding, since we have wide knowledge of typical parameters, and very reliable knowledge of how to apply selection. One area where uncertainty will not disappear lies in estimating the merit of new animals and in the relationships between genotypes and animal performance. In this context, new crops of animals can arguably be viewed as a research project – in the sense that their data informs genomic prediction, quite literally so.

- Genomics has been promoted as a disruptive technology, potentially ushering in a new era in which progress would be faster, directed at a much more interesting set of traits, and more glamorous – no more reliance on tedious recording of old-fashioned traits! It may well be disruptive, but very likely because its heavy dependence on data will radically transform industry structures, in ways explored and analysed by Bichard, James and others in the 1970s, and in particular, focus attention on the commercial relationships around data.

Genomic selection is making very clear the paramount importance of good data on the relevant traits on informative animals. To date in Australia in all species optimising investment in data has been left to the individual breeder, and it can be argued that much of our focus in national genetic evaluation schemes has been on making the best use of data of very variable quality. Genomic technologies allow us to uncouple recording and estimation of breeding values, and this will force attention of data quality, quantity and on efficient structures for recording and for exploiting the resulting information.

- Principles and recommendations are emerging for the technical requirements underpinning genomic selection, and are clear for parameter estimation in general. Given that these are essentially minimum data requirements, should collective funds be applied to breeds that either don't, or barely meet these minima?

As a rule of thumb, a breeding population (more typically referred to as a breed) requires approximately 5-6,000 recorded and genotyped animals per trait to achieve useful accuracy of genomic prediction. Further, as selection proceeds, this will need to be "topped up" or "refreshed" with approximately 1-2,000 plus new animals recorded and genotyped per trait per year. For many breeds of extensive livestock in this country, these numbers are close to or exceed their total current recording populations. This raises the interesting question of whether anyone should care about this, and if so, what we might do about it.

- Can we afford to conduct research on animals other than those at the apex of the breeding pyramid? Do we work closely enough with leading breeders?

Over the last 30 years in Australia there has been a steady decline in public sector support for agricultural R&D, including research stations and large-scale genetics research projects such as selection experiments. To some extent the investments in the Information Nucleus projects in beef and sheep have slowed this trend, but importantly these have involved heavy sampling of elite industry genetics.

Research scientists, together with industry and government funding bodies, will increasingly need to maximise the cost-effectiveness of research projects, and in doing so take account of time to adoption. This focus invariably means that earlier adoption is better than later, and if research is conducted on the active population, there are benefits in not having to re-estimate between the experimental and active populations.

It is hard to see that this will not lead to at least a section of the commercially and genetically active population becoming simultaneously a section of the research population, and that where possible research activities will be integrated into normal commercial breeding operations. This may require some funding flowing directly to the herds or flocks involved, but overall savings will come from speed and genetic relevance.

- Is it possible that industries simply do not have the capacity to benefit from quite rapid change in some traits, for instance in product quality? Is gradual improvement all that can be harvested, or does it simply fail to be detected?

Reflection on genetic and (commercial sector) phenotypic changes that have occurred in beef cattle and sheep (and probably other species as well) over the last 30 years, it appears that most changes have been gradual, which is consistent with all traditional expectations of genetic improvement. However there are signs that as genetic progress accelerates through learning and as wider genetic variation is accessed – an example being Wagyu genetics – that phenotypic changes can actually be quite dramatic. For this to really take effect, feedback of some form is essential, and that feedback needs to encompass the full range of performance. Otherwise, commercial returns will lag well behind what is possible. In industries with considerable diversity of markets and production enterprises, this inefficiency of feedback is almost certainly generating significant opportunity costs. A case could be made for making systems such as MSA universal through industry support, simply in order to maximise the efficiency of information flow.

- Genetic variation is usually the largest single definable source of phenotypic variation, yet we tend to think of it as minor. Does this affect our communications? Our investment patterns?

Most R&D and management investment over the last 50 years has been into non-genetic change, yet, especially once across-flock and across-herd genetic evaluations are in place, the range in genetic merit of bulls or rams available to commercial producers very rapidly becomes the largest proportion of improvement available to the producer. There is a strong case to be made that investment into new phenotypes should be the priority in livestock R&D, and capturing the benefits of superior genetics the focus of management investment.

- Do we need to more directly take account of the fact that the breeding businesses in our industries are both numerous and modest size? And that these structural factors mean that at a broad level, all traits are important?

Australia (and I think New Zealand) are almost unique in the number and average size of breeding sector operations within the beef and sheep industries. One senses that the resulting diversity is brushed aside as an unfortunate cost, but could it perhaps be, a la “Wisdom of Crowds” a resource? Including the possibility that it allows for recording of more traits than would otherwise be supported?

- Much of livestock production has involved applying management skills which entail adjusting the environment to suit the animal, ideally profitably. This needs to be, and will be, turned on its head. Genetically changing the animals to maximise their fit with the physical and consumer environment is not only more profitable in any but the very short term, but can now be achieved rapidly.

This extends the point above about the returns from investment in genetic and non-genetic change. But it links to the important message that for societal acceptance of livestock agriculture to be maintained, it would be far better if animals did not get diseases or require surgical intervention or potentially stressful management treatment. Breeding animals with these characteristics is simply a matter of having the data and selecting on breeding value. Industry should get on with it.

- Despite this last point, we face the problem that genetic improvement is for many an abstract concept, which hinders both adoption and investment. Making genetic improvement concrete, rapid and obvious is worthwhile.

This point needs no explanation, but it is worth remembering that simple, clear communication of the positive stories about sensible genetic improvement will maintain not only the right to farm but also the right and support to do genetics R&D.

- Finally, the field has focussed essentially on the simple mathematics of objectives, evaluation and design, all of which are well-established and in principle straightforward. New thinking is needed on the structural and business innovation needed to convert these principles into practice.

This point is perhaps rather too broad, but there is a link back to my earlier point about the structure and diversity of our extensive livestock industries' breeding sectors. These structures have survived the last 30 years of enormous changes in evaluation methods and rates of progress, but the introduction of genomic tools seems likely to usher in changes in the role(s) of breeders, their clients and their societies. Helping industry find robust, efficient and equitable models for implementation of the new approaches will require bringing economics and genetics R&D together, and likely involve insights from fields such as network analysis, value chain economics, and service model innovation.

Each of these can be explored further, but my overall message is that while we have made significant progress in the tools of genetic improvement, it is not so obvious that putting them into practice has advanced as much. This could be viewed as that we have done better at those parts of the livestock genetic improvement that do not involve humans – in simple terms, the maths and the modelling. And yet all the decisions that generate returns from the science are made by humans.

To date we have relied on simple (and sound) principles of extension, and in this country tended to avoid directly engaging with industry structural issues. This approach could be characterised as slow and gradual change by infiltration – the "virus" of genetic change has been evolved slowly and introduced rather slowly.

This is not really good enough – livestock industries need to be making improvement at 3% or more per year in profitability simply to stay in place, and this is quite achievable, as several examples demonstrate.

The next phase in the Australian livestock genetic improvement venture requires us to match the innovation and radical simplicity of much of the technical work with parallel innovation and radical simplicity in implementation, and indeed, in making the technical and the implementation simply two sides of the same coin. Each of our livestock industries has examples of methods that have achieved this, but we now need to make those examples the core of what we do, rather than limited scale trials.

Genetic improvement is far too important to leave either to chance or to the slow workings of imperfect markets. And "research" and "extension" (or implementation) cannot be separate. The industries that have made the most out of animal breeding science simply make genetic improvement and then harvest its value. For the sake of the Australian community, we must do the same.

TABLE OF CONTENTS

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:1-5

UNDERSTANDING HOW FARMERS LEARN

H.T. Blair[1], A.M. Sewell[1], R.A. Corner-Thomas[1], P. Kemp[1], B.A. Wood[1], D.I. Gray[1], S.T. Morris[1], A.W. Greer[2], C.M. Logan[2], A.L. Ridler[1], R.E. Hickson[1] and P.R. Kenyon[1]

[1]Massey University, Palmerston North, New Zealand.
[2]Lincoln University, Lincoln, New Zealand.

SUMMARY

Changing the behaviour of people is challenging; changing farmer behaviour is possibly even more so. The evidence presented here suggests that a number of widely-used farmer communication methods are poorly thought of by farmers. Information received by farmers from other farmers was regarded as useful, and this information was regarded as being more useful than that from a number of rural professionals. Those wishing to change farmer behaviour need to: invest time to gain trust; involve farmers in the process of learning; use multiple methods to teach and encourage farmers to talk with each other and scientists in a learning community.

INTRODUCTION

The current New Zealand Government expects the New Zealand scientific community to improve the rate of uptake of new knowledge by businesses and thereby improve the New Zealand economy. Similarly, Centres of Research Excellence funded by the Tertiary Education Commission are expected to show how they will translate new knowledge into improved community benefit, and the recent Primary Growth Partnership granted to Beef + Lamb New Zealand (the farmer-owned industry organisation representing New Zealand's sheep and beef farmers) aims to improve access to information by farmers. However, the rate at which behaviour change by business owners is driven through the provision of new scientific evidence is variable and this is particularly so in the agricultural sector. Indeed, Leeuwis and Aarts (2011) suggested that much of agricultural extension falls well short of achieving lasting change in famer practice.

This paper reports on a pilot farmer learning project and a survey of New Zealand sheep farmer opinion with the intent to show how farmers go about learning new technologies, including how they like to receive information and who farmers perceive as providing useful information.

METHODS

An experimental farmer learning project has been underway at Massey University since 2011. The original group of 18 sheep and beef farmers was expanded to 26 in February 2013. The farmers work with an interdisciplinary group of 7 University experts (3 animal scientists, an agronomist, a farm management specialist, an educationalist and a sociologist). The project focused on a University farmlet trial that investigated lamb finishing on herb mix pastures (clover, chicory and plantain). The participants met 4 times per year at Massey University during a 24 hour period from noon to noon. Farmer participants were interviewed pre-project and after each meeting with specific questions about what activities and experiences had supported their learning.

A printed survey was sent to approximately 12,000 sheep and beef farmers whose addresses were on the Beef + Lamb New Zealand database. The survey was included within the 'Heartland Sheep (NZX Agri, Feilding New Zealand) magazine in October 2012. Farmers had the opportunity to either, fill in the survey and return it via a pre-paid envelope, or to fill it in electronically via a website "www.SurveyMonkey.com". A total of 971 surveys were returned (934 by post and 37 completed online).

Part A of the survey asked farmers to identify themselves based on their farm type (ram breeder or commercial farm) and the breed(s) of sheep on their farm. If a farmer indicated they

had both a ram breeding flock and a commercial flock on their farm they were classified as being a ram breeder (94 vs. 844 ram breeder and commercial farmers respectively). In Part B of the survey farmers were asked to indicate the usefulness of information providers and the usefulness of different forms of technology transfer. Scoring used a one to four scale (1 = no use at all, 2 = little use, 3 = useful, 4 = very useful).

The responses were analysed using the Genmod procedure using a binomial distribution and a log-transformation (SAS 2011) and included the fixed effect of farm type. Scores were analysed using the Genmod procedure using a Poisson distribution and a log-transformation and included the fixed effects of farmer age and farm type.

RESULTS AND DISCUSSION

Results from the farmer survey suggested that farmers place value on obtaining information from other farmers more than most other professions, with the exception of veterinarians (Table 1). It was surprising that farm consultants scored poorly. The farmer learning project showed a similar result albeit on a specific question about the use of herb pastures, whereby farmers were the second most useful group after seed merchants (Table 2). Given that the farmer learning project was focussed on the application of herb pastures, it is unsurprising that seed merchants were considered the most useful source of information.

Farmers placed greater emphasis on the print media (books / booklets, farming press, newspapers and fact sheets) than they did on most other means of technology transfer, the exception being field days (Table 3). The electronic media (CDs, DVDs, phone apps and texts) were considered of little use, although email updates and web-based information were considered useful. This may reflect a typically older age group amongst sheep farmers who are less confident with electronic media. They liked receiving a single page of “normal-people notes” written by scientists, but in language understandable by farmers.

The only significant differences in opinion between commercial farmers and ram breeders involved the usefulness of scientists as information providers (Table 1) and the usefulness of scientific seminars (Table 3), whereby ram breeders found ‘science’ more useful.

The 3-year pilot farmer learning project provided an on-going and up-to-date science focus for scientists to share evidence-based ideas about how herb pastures grow and are utilised by animals. This participatory experience not only provided the most up-to-date and unbiased information, it also provided comparative data such as lamb live-weight gains, botanical composition and weed control. While this engagement in science is labour intensive, it is likely that it is also the most effective method of changing farmer behaviour (Rogoff 2003). This need for engagement is likely to explain the low rating given by farmers for some forms of technology transfer in the farmer survey.

Farmers and scientists were both positioned as experts with different skills to share about herb pasture management. When expertise is distributed across a group and different research-based findings shared, new ideas ‘seed’ and can ‘migrate’ to other members of the community who transform them into new understandings (Brown and Campione 1998). The farmer learning project deliberately built responsive, respectful and trusting relationships between farmers and scientists and between farmers and farmers. Sinnema and Aitken (2012) in a meta-analysis of research, found respectful and reciprocal relationships in learning communities to be an effective determinant of learning. The mutual trust, respect, openness and honesty highlighted the importance of farmers and scientists knowing each other and of understanding their farming systems. The relatively low rating achieved by farm consultants in the survey would suggest they might benefit from devoting time to building trusting and respectful relationships with their farmer clients. Indeed, those consultants who achieve repeat visits, and are therefore likely to be

considered "more useful", are known to invest effort in developing relationships (Gray *et al.* 1999).

In the farmer learning project, a wide variety of multi-sensorial experiences replicating reality, were designed to motivate farmers and to provide repeated opportunities to participate in their learning, instead of simply telling them the key ideas. These learning experiences included: observations, listening, talking, tasting, reading, interpreting data, questioning, comparing ideas, challenging ideas, using calculators, transects and visiting different farms. These varied experiences led to engagement, which in turn should lead to learning. There is convincing educational research pointing to the importance of designing experiences that increase engagement, interest and motivation. Learners should experience at least three different sets of complete information about a concept before it becomes embedded in their network of knowledge, doing so provides the opportunity to revisit concepts (Nuthall 2007).

The farmer learning project intentionally positioned farmers as learners. There were no recipes for herb pasture management, hence the importance of farmers learning how to learn. Farmers came to see themselves as learners, indeed as co-learners and inquirers alongside the scientists. They became producers of knowledge with others, rather than as consumers of researchers' knowledge. They saw the gaps in scientific knowledge and were motivated to join with them in on-going research. This joint participation of farmers and scientists moves past the acquisition metaphor of learning that requires an expert to transmit a body of knowledge (e.g. when farmers listen passively to a speaker), to an emphasis on participation where farmers can observe and get involved in new technologies (Sfard 1998). More recently, Paavola *et al.* (2004) identified a knowledge-creation metaphor to emphasise how original ideas are transformed, expanded or "hatched" in an exchange of views, or dialogue, in 'innovative knowledge communities'.

It would seem likely that if those wishing to change farmer behaviour were better versed in how farmers learn, and what works to support their learning, then greater rates of adoption of, for example, animal breeding and genetic technologies might occur.

ACKNOWLEDGEMENTS

The project was supported by funding from the Massey University's Sheep Research Centre, Gravida (a Centre for Research Excellence) and the Massey University – Lincoln University Agriculture Partnership for Excellence.

REFERENCES

Aitken G. and Sinnema C. (2008) Effective Pedagogy in Social Sciences. New Zealand Ministry of Education, p. 83.

Brown A. L. and Campione J. C. (1998) In 'How students learn: Reforming schools through learner-centered education', p160, editors N.M. Lambert and B.L. McCombs, American Psychological Association, Washington.

Gray D., Kemp E., Gardner J., Rogers N. and McCosh K. (1999). *Journal of International Farm Management* **2**:113.

Leeuwis C. and Aarts N. (2011) *The Journal of Agricultural Education and Extension* **17**:21.

Nuthall G. (2007) The hidden lives of learners. NZCER, Wellington.

Paavola S., Lipponen L. and Hakkarainen K. (2004) Review of Educational Research **74**: 557.

Rogoff B. (2003) 'The Cultural Nature of Human Development', p37. Oxford University Press.

Sinnema C., and Aitken G. (2012) 'Effective Pedagogy in Social Sciences', p15. Belley International Academy of Education, International Bureau of Education.

Sfard A., 1998. *Educational Researcher* **27:** 4.

Table 1. Farmer responses to the question: "Please indicate using a 1 to 4 scale the relative usefulness for you of each of the following information providers"

Provider	Commercial[1]	Ram Breeder[1]	Commercial vs. Breeder
Accountants	1.10 ± 0.02 (3.0) ^{h}	1.05 ± 0.06 (2.9) ^{defg}	ns
Agricultural contractors	1.03 ± 0.02 (2.8) ^{fg}	1.06 ± 0.06 (2.9) ^{defg}	ns
Agricultural retailers	1.06 ± 0.02 (2.9) ^{gh}	1.08 ± 0.06 (2.9) ^{defg}	ns
Agronomists	0.92 ± 0.02 (2.5) ^{cd}	0.92 ± 0.07 (2.5) ^{cde}	ns
Banking	1.02 ± 0.02 (2.8) ^{fg}	1.00 ± 0.06 (2.7) ^{def}	ns
Beef + Lamb NZ	0.93 ± 0.02 (2.5) ^{de}	0.92 ± 0.07 (2.5) ^{cde}	ns
Farming consultants	0.67 ± 0.03 (2.0) ^{b}	0.65 ± 0.08 (1.9) ^{ab}	ns
Fertiliser reps	0.98 ± 0.02 (2.7) ^{ef}	0.90 ± 0.07 (2.5) ^{cd}	ns
Meat companies	1.00 ± 0.02 (2.7) ^{f}	0.95 ± 0.06 (2.6) ^{def}	ns
Other farmers	1.10 ± 0.02 (3.0) ^{h}	1.11 ± 0.06 (3.0) ^{fg}	ns
Ram breeders	1.03 ± 0.02 (2.8) ^{fg}	1.09 ± 0.06 (3.0) ^{fg}	ns
Regional council	0.43 ± 0.03 (1.5) ^{a}	0.50 ± 0.08 (1.6) ^{a}	ns
Scientists	0.85 ± 0.02 (2.4) ^{c}	1.01 ± 0.06 (2.8) ^{defg}	*
Stock agents	1.03 ± 0.02 (2.8) ^{fg}	0.96 ± 0.06 (2.6) ^{def}	ns
Veterinarians	1.16 ± 0.02 (3.2) ^{i}	1.17 ± 0.06 (3.2) ^{g}	ns
Wool buyers	0.87 ± 0.02 (2.4) ^{cd}	0.75 ± 0.07 (2.1) ^{bc}	ns

Means within columns with differing letter superscripts are significantly different $P<0.05$
Differences between commercial and ram breeder responses, p-values of $P>0.05$ are indicated by ns, $p<0.05$ by *
[1] Back-transformed %

Table 2. Farmer responses to the question: "Please identify the three people that you've found it most useful to talk to or use so far about herb pastures"

Role	Number	%
Accountant	0	0.0
Banker	0	0.0
Consultant	3	7.7
Contractors	1	2.6
Farmer	11	28.2
Industry good	1	2.6
Merchant (fertiliser)	2	5.1
Merchant (seed)	18	46.2
Other	0	0.0
Scientist	2	5.1
Veterinarian	1	2.6
TOTAL	**39**	**100**

Table 3. Farmer responses to the question: "Indicate using a 1 to 4 scale the relative usefulness for you, for each of the following forms of technology transfer"

Technology Transfer	Commercial[1]	Ram Breeder[1]	Commercial vs. Breeder
Books / Booklets	1.12 ± 0.02 (3.1) [l]	1.06 ± 0.06 (2.9) [gh]	ns
CDs	0.51 ± 0.03 (1.7) [c]	0.50 ± 0.09 (1.6) [a]	ns
Certificate level courses	0.42 ± 0.03 (1.5) [ab]	0.50 ± 0.09 (1.6) [a]	ns
DVDs	0.60 ± 0.03 (1.8) [d]	0.61 ± 0.08 (1.8) [ab]	ns
Demonstration farms	0.84 ± 0.02 (2.3) [fg]	0.84 ± 0.07 (2.3) [cdef]	ns
Diploma level courses	0.46 ± 0.03 (1.6) [bc]	0.48 ± 0.09 (1.6) [a]	ns
Email updates	0.94 ± 0.02 (2.6) [ij]	0.93 ± 0.07 (2.5) [cdefg]	ns
FITT programme	0.72 ± 0.03 (2.1) [e]	0.78 ± 0.07 (2.2) [bc]	ns
Fact sheets (1-2 pages)	0.96 ± 0.02 (2.6) [ij]	0.99 ± 0.07 (2.7) [defg]	ns
Farmer discussion groups	0.97 ± 0.02 (2.7) [j]	1.02 ± 0.06 (2.8) [fgh]	ns
Farming press	1.15 ± 0.02 (3.1) [l]	1.17 ± 0.06 (3.2) [h]	ns
Field days	1.05 ± 0.02 (2.9) [k]	1.08 ± 0.06 (2.9) [gh]	ns
Industry workshops	0.91 ± 0.02 (2.5) [hi]	0.95 ± 0.07 (2.6) [cdefg]	ns
Monitor farms	0.86 ± 0.02 (2.4) [gh]	0.83 ± 0.07 (2.3) [cde]	ns
Newspapers	1.00 ± 0.02 (2.7) [jk]	1.01 ± 0.06 (2.7) [efgh]	ns
Phone apps	0.37 ± 0.03 (1.4) [a]	0.40 ± 0.09 (1.5) [a]	ns
Radio	0.85 ± 0.02 (2.3) [gh]	0.92 ± 0.07 (2.5) [cdefg]	ns
Scientific literature	0.83 ± 0.02 (2.3) [fg]	0.83 ± 0.07 (2.3) [cde]	ns
Scientific seminars	0.60 ± 0.03 (1.8) [d]	0.77 ± 0.07 (2.2) [bc]	*
Television	0.78 ± 0.02 (2.2) [ef]	0.81 ± 0.07 (2.2) [cd]	ns
Tertiary level courses	0.47 ± 0.03 (1.6) [bc]	0.51 ± 0.09 (1.7) [a]	ns
Text updates	0.34 ± 0.03 (1.4) [a]	0.46 ± 0.09 (1.6) [a]	ns
Web based information	0.88 ± 0.02 (2.4) [gh]	0.85 ± 0.07 (2.3) [cdef]	ns

Means within columns with differing letter superscripts are significantly different P<0.05
Differences between commercial and ram breeder responses, p-values of P>0.05 are indicated by ns, p<0.05 by *
[1] Back-transformed %

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:6-10

MODELLING FARMER INFORMATION TRANSFERS WITH NETWORK ANALYSIS: AN EXPLORATORY FARMLET STUDY

B.A. Wood, H.T. Blair, D.I. Gray, P. Kemp, P.R. Kenyon, S.T. Morris, A.M. Sewell

Massey University, Palmerston North, New Zealand.

SUMMARY

Traditional extension models place farmers at the receiving end of knowledge transfer. The agricultural sector would be better served by a learning model that emphasises networking rather than linearity. Farmers are not simply receivers but also routinely act as extension agents who circulate knowledge through their own interpersonal contacts. A Massey University case study demonstrates the power of these relationships to transfer scientific awareness between farmers. The case also suggests that knowledge flows are affected by the sociological traits of farmer networks. Densely connected and occupationally homogenous networks transfer knowledge at a faster rate than do networks that are loosely tied and heterogeneously composed.

INTRODUCTION

An experimental farmer learning project has been underway at Massey University since mid-2011. 25 sheep and beef farmers are working with an interdisciplinary group of 7 University experts (3 animal scientists, an agronomist, a farm management specialist, an educationalist and a sociologist). The project focuses on a farmlet trial that investigates lamb finishing with herb mix pastures (clover, chicory and plantain). The project has used this trial to explore knowledge transfer opportunities in the pastoral sector. These opportunities centre on farmer learning, in particular through improved interaction with the producers of scientific knowledge.

The Massey learning project has two major dimensions: (1) designing effective small group interactions between farmers and agricultural scientists, and (2) exploring the wider circulation of scientific knowledge through farmer networks. This paper concentrates on the latter. Every farmer maintains their own 'network of practice'. This network is a unique collection of agricultural practitioners (e.g. other farmers, consultants, researchers and merchants of various kinds) that the farmer contacts on a regular basis. Although these interactions are regular, they serve multiple purposes and hence the networks they construct tend to be informally, rather than single-mindedly, coordinated (Eastwood *et al*. 2012). A growing body of research emphasises the significance of such networks for the development of agricultural innovation systems (Darnhofer *et al*. 2012).

The Massey learning project brings together a number of agricultural scientists and farmers to test a specific pastoral innovation, the use of herb pastures. The project design has included an exploratory analysis of how the participating farmers circulate knowledge of the Massey trial through their own networking activities. This analysis lends support to the idea that farmers are significant scientific agents rather than simply end-of-the-line recipients.

METHOD

The participating farmers have been interviewed twice, once prior to the start of the Massey trial in mid-2011 and again 18 months later in December 2012. Complete network data has been collected for 17 of the farmers and this is the dataset analysed here. The first round of interviews recorded each of the farmer's existing contacts for herb knowledge. The follow-up interviews recorded the people with whom the farmer had discussed the Massey trial over the preceding 18 months. These people are divided into: (1) those already identified as existing contacts in the first interview, and (2) new people not previously identified as significant herb knowledge contacts. In essence then, the interviews reveal the extent to which the farmers activated and added to their

existing networks. To use a learning terminology, the results show the retention and recruitment dynamics of farmer-sponsored enrolments in the herb trial.

Social network analysis uses various quantitative metrics to map the social structures in which individuals are embedded (Prell, 2012). The following variables are analysed here:

1. size: the number of herb pasture contacts nominated by the farmer, in terms of pre-existing relationships as well as with regard to retention, recruitment and growth from 2011 to 2012.
2. role: herb contacts are classified according to 11 occupational roles.
3. density: the number of actual ties divided by possible ties. For example, a network of 5 actors has (5*4)/2 = 10 possible ties (i.e. herb contact relationships). If there are 5 actual ties then network density is 0.5.
4. heterogeneity: calculated as 1 minus the sum of the squares of the proportions of each value of the categorical role variable in each of the 17 farmer networks. In network analysis this is known as Blau's heterogeneity index, but it is an often re-invented and diversely named measure (Gibbs-Martin, Gini-Simpson, Herfindahl-Hirschman). Varying between 0 and 1, the index measures the mix of occupations held by the farmer's herb contacts. Statistically, it is the chance that two randomly selected individuals from the farmer's network will have different occupations (Harrison and Klein, 2007).

THE HOMOPHILY HYPOTHESIS

We hypothesise (1) that the 17 farmers will retain and recruit other farmers into the herb trial more frequently than any other occupational grouping, (2) that densely interconnected farmer networks will grow at a higher rate than those that are more loosely tied together, and (3) that occupationally homogenous farmer networks will grow at a higher rate than networks which are occupationally heterogeneous.

The rationale for these hypotheses is furnished by the homophily principle, which holds that contact between similar people occurs more frequently than contact between dissimilar people (McPherson *et al*. 2001). Relatedly, the theoretical literature suggests that high density networks create the trust and shared values that enhance knowledge transfer (Nooteboom and Stam 2008). In short, birds of a feather flock together. A wide range of causal factors have been invoked to explain the homophily principle, but such explanatory analysis is not our objective here. Suffice to say that the principle has been so frequently observed empirically that it approximates a sociological law. Accordingly we hypothesise that social similarity has empowered the circulation of knowledge about the Massey herb trial, just as homophily routinely empowers the sharing of many other social resources.

RESULTS

Table 1 cross-tabulates the 17 farmers' network enrolments by occupational role over the period from June 2011 to December 2012. Prior to the trial's launch, the farmers were sharing herb knowledge across a wide range of occupations. However, half (50.4%) of their networkers were fellow farmers, well ahead of any other occupational group. By 18 months after its launch, the farmers had discussed the Massey trial with 63.2% (79) of their existing contacts and with 113 new people not previously identified. There had thus been a significant growth in network reach. By the end of 2012 the 17 farmers had constructed a new network with 192 members, 53.6% larger than when the trial began. As might be expected, seed merchants were consistently highly placed. Consultants, bankers, contractors, industry-good representatives, scientists and veterinarians all exhibited much the same network presence. Fellow farmers, however, had been both retained and recruited at a much higher rate than any of these other groups. While the pre-existing network was half farmers, the new network was two thirds farmers (65.6%).

Table 1. The 17 farmers' herb knowledge network enrolments by occupational role, June 2011 to December 2012

Role	June 2011 network	Retention	Recruitment	Dec 2012 network
Accountant	0	0	1	1
Banker	0	0	5	5
Consultant	8	5	1	6
Contractors	12	4	3	7
Farmer	63	46	80	126
Industry good	4	2	5	7
Merchant (fertiliser)	5	2	1	3
Merchant (seed)	20	12	4	16
Other	2	1	4	5
Scientist	8	5	2	7
Veterinarian	3	2	7	9
TOTAL	125	79	113	192

Table 2. Structural traits of the 17 farmer networks

Farmer	Initial size	Density	Heterogeneity	Growth%
A	7	0.333	0.776	57.14
B	7	0.476	0.694	14.29
C	17	0.544	0.616	58.82
D	11	0.400	0.744	0.00
E	6	0.733	0.500	233.33
F	6	0.733	0.500	233.33
G	17	0.191	0.740	-23.53
H	16	0.617	0.570	0.00
I	9	0.444	0.815	33.33
J	3	0.333	0.444	133.33
K	6	0.667	0.611	33.33
L	7	0.571	0.694	200.00
M	11	0.600	0.545	81.82
N	5	0.700	0.640	100.00
O	8	0.536	0.656	100.00
P	14	0.231	0.714	-21.43
Q	10	0.311	0.680	60.00

Analysis of the network findings presented in Table 2 lends further support to the homophily hypotheses. For example, the linear regression of network growth on density is significant (r^2 = 0.357, p = 0.011, slope = +284.00). So too is the regression of growth on heterogeneity (r^2 = 0.280, p = 0.029, slope = -0.320).

DISCUSSION

A social network analysis of Massey University's herb pasture trial suggests that farmers are well placed to disseminate agricultural science. As yet we know little about the actual content of these interpersonal communications (though some qualitative data have been collected to inform such an analysis). What is evident, however, is both the scale and the farmer-focus of networking by the 17 farmers participating in the Massey trial. This network reach runs along the lines of social homophily. Dense and occupationally homogenous networks seem to spread agricultural knowledge further than do networks that are loosely and heterogeneously composed. This suggests that farmer information transfers are the expressions of social solidarity; they disseminate science in relations bound by interpersonal trust and shared norms.

CONCLUSION

The scientific inflection of New Zealand farming has long been a significant source of competitive advantage in global markets. It is widely recognised, however, that the circulation of agricultural knowledge is at a critical turning point. New Zealand farming is challenged to make rapid and profound transformations, and yet its current knowledge transfer system is, as a major industry organisation candidly concludes, marred by 'too much noise and mistrust' (Beef+Lamb 2013). This paper has suggested that, in the midst of such official and commercial noise and mistrust, pastoral farmers maintain their own informal and practical networks. Those bent on improving the significance of agricultural science among New Zealand farmers would do well to enhance such farmer-to-farmer communication channels. The social embedding of agricultural science in spontaneous, farmer-driven, conversations is a key not only to the past successes but also to the future prosperity of New Zealand pastoralism.

ACKNOWLEDGEMENTS

The project was supported by funding from Massey University's Sheep Research Centre and Gravida (a Centre for Research Excellence).

REFERENCES

Beef+Lamb (2013) *Primary Growth Partnership: Programme Overview*, available online at http://www.beeflambnz.com/Documents/News%20and%20events/Collaboration%20for%20Sustainable%20Growth%20-%20Overview.pdf?_cldee=Yi5hLndvb2RAbWFzc2V5LmFjLm56.

Darnhofer I., Gibbon D. and Dedieu B. (2012) *Farming Systems Research into the 21st Century: The New Dynamic*. Springer.

Eastwood C., Chapman D. and Paine M. (2012) 'Networks of practice for co-construction of agricultural decision support systems: Case studies of precision dairy farms in Australia', *Agricultural Systems* **108:**10.

Harrison D. and Klein K. (2007) 'What's the difference? Diversity constructs as separation, variety, or disparity in organizations', *The Academy of Management Review* **32:**1199.

McPherson M., Smith-Lovin L. and Cook J.M. (2001) 'Birds of a Feather: Homophily in Social Networks', *Annual Review of Sociology* **27:** 415.

Nooteboom B. and Stam E. (2008) *Micro-Foundations for Innovation Policy*. Amsterdam University Press.

Prell C. (2012) *Social Network Analysis: History, Theory and Methodology*. Sage.

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:10-13

THE CALIFORNIA COMMERCIAL BEEF CATTLE RANCH PROJECT

A.L. Van Eenennaam[1], **K.L. Weber**[1] and **D.J. Drake**[2]

[1] Department of Animal Science, University of California, Davis, CA 95616 USA
[2] University of California Cooperative Extension, Yreka, CA 96097 USA

SUMMARY

The overall objective of this project was to develop a genotyped, phenotyped population to permit the assessment of different DNA-enabled approaches for predicting the genetic merit of Angus sires on commercial beef ranches. Approximately 5,400 progeny conceived in natural-service, multiple sire breeding pastures on 3 commercial cow-calf ranches in Northern California from 2009-2011 were assigned to a single herd bull using SNP data. The number of calves born per sire per calf drop varied greatly, ranging from 0 in ~ 7% of bull seasons to 64. The total adjusted 205d weight per bull per calf drop was almost totally explained by prolificacy (R^2=0.98), and showed little correlation (R^2=0.09) with average calf weaning weight per sire. Over 4,000 offspring were followed through the feedlot, and processing plant to obtain carcass data. Progeny data from the bulls' first calf crop were used to calculate commercial ranch genetic (rEBV) evaluations for the bulls. These rEBVs were then compared to breed association (bEBV) EBV and genomic predictions (gMBV). The rEBV was the most predictive of future progeny performance, with the bEBV and Angus-trained gMBV having similar predictive ability.

BACKGROUND

Genomic breeding values have emerged as a promising technology for providing more accurate breeding values for selection candidates in cattle populations. However, relative to successes in the dairy industry, its adoption in the commercial beef industry remains sluggish. While the value proposition associated with using this technology in the stud sector may have some merit in improving the accuracy of selection (Van Eenennaam *et al.* 2011), there are numerous practical difficulties associated with using this technology in commercial settings and the feasibility and value association with collecting DNA samples from commercial beef cattle remains uncertain.

MATERIALS AND METHODS

A project was conducted to derive a population of Angus sires with high density (50K) genotypes, purchased as yearlings and used as herd bulls in multi-sire breeding pastures with predominantly Angus commercial cows in Northern California, concomitant with phenotyped progeny from which to assess the accuracy of Angus genomic predictions for traits measured in a commercial setting. A small number of South Devon, Hereford and Red Angus bulls were also used on these ranches. The cow to bull ratio was approximately 25:1 and breeding took place in fenced pastures. Bulls underwent a breeding soundness exam prior to the breeding season and were then assigned to breeding groups. Bulls remained in the same breeding group unless they were injured or in inadequate condition based upon the judgment of experienced personnel working on the cooperating ranches. Approximately 5,400 progeny born in 2009-2011 on 3 commercial cow-calf ranches were sire-identified to herd bulls using DNA information from tail hair collected at the time of weaning weight data collection. Weaning weights were adjusted for sex, cow age and calf age according to Beef Improvement Association recommendations except that age ranges were wider than this guide due to practical constraints associated with calves going to summer pastures where they were not accessible for weighing. Carcass data and a meat sample were collected on over 4000 carcasses for DNA confirmation of the animal's identity by comparing the genotype of the meat sample to that which was obtained from hair samples

collected on all calves at weaning.

SIRE PROLIFICACY

Birthdate records were collected on 5,940 individually identified calves enrolled in the trial. DNA samples were collected on 5,382 (90.6%) calves and of these 5,272 (98%) were assigned to an individual sire. Bulls present for a full breeding season siring at least one calf (n=263) produced 19.2±13.1 progeny per calf drop, ranging from 1 to 64 (Table 1). Bulls with reduced breeding seasons due to injury or lack of condition (n=33) produced fewer calves (9.1±8.5) compared to full breeding season bulls (19.2±13.3) progeny (P<.01). Prolificacy was by far the main driver of total weight weaned per sire. The total adjusted 205d weight per bull per calf crop was related (P<.01) to the number of calves (220±1.8 kg increase for each calf) explaining 98 percent of the variation in sire weight weaned per calf crop, and showed little relation (R^2=0.09) with mean adjusted progeny weaning weight per sire.

Table 1. Average bull age at the beginning of the breeding season, and number of calves produced per bull on 3 commercial ranches in Northern California during 2009-2011[1]

Ranch	Year	Season	# Bulls/ season	Mean bull age years (± SD)	Total # calves	Per bull		
						Min # calves	Max # calves	Mean # calves (± SD)
A	2009	Spring	18	3.8 ± 1.2	353	3	47	19.6 ± 13.4
		Fall	19	4.7 ± 0.8	113	1	29	16.1 ± 10.0
	2010	Spring	22	3.6 ± 0.9	346	1	47	18.2 ± 14.2
		Fall	19	4.5 ± 1.0	328	1	48	17.3 ± 12.6
	2011	Spring	17	3.9 ± 1.1	402	4	53	23.6 ± 13.6
		Fall	19	5.4 ± 0.7	286	1	33	15.0 ± 9.2
B	2009	Spring	8	4.6 ± 3	141	1	45	17.6 ± 17.0
		Fall	10	5.1 ± 2.5	214	10	50	21.4 ± 11.4
	2010	Spring	8	3.4 ± 1.4	142	3	30	17.8 ± 8.4
		Fall	12	5.1 ± 2.7	247	4	44	20.5 ± 11.4
	2011	Spring	4	4.6 ± 1.7	110	18	42	27.5 ± 11.0
		Fall	12	5.3 ± 2.9	266	3	51	22.2 ± 15.2
C	2009	Fall	30	4.2 ± 1.1	642	2	54	21.4 ± 13.8
	2010	Fall	27	4.6 ± 1.3	567	1	52	21.0 ± 13.0
	2011	Fall	38	5.4 ± 1.8	573	1	64	15.1 ± 16.1
A	2009-11	All	114	4.0 ± .2	2150	1	53	18.8 ± 1.2
B	2009-11	All	54	4.8 ± .2	1120	1	51	20.8 ± 1.8
C	2009-11	All	95	4.8 ± .2	1782	1	64	18.7 ± 1.4
A,B,C	2009-11	All	263	4.4 ± 1.7	5052	1	64	19.2 ± 13.3

[1] Table values are for bulls present for entire breeding seasons. Thirty three additional bulls were used for only a portion of the breeding season due to injury or other issues (data not shown).

Bulls produced a similar average number of progeny across ranches (18.8±1.2, 20.7±1.8, 18.7±1.4, P=.63), years (19.6±0.4, 19.2±1.4, 17.9±1.4, P=.68), and season (20.3±1.5, 18.3±0.9, P=.27). The most calves born from any single bull in one day was 11 from exposure to these naturally-cycling, unsynchronized females. The more prolific bulls sired more early calves (day 1-21 of the calving season) than low prolificacy bulls. There were also bulls that sired no calves. For each time a bull had an opportunity to breed cows in any given season, there was a 7.3% chance that he would produce no calf. This value is similar to the 6.0% reported by Holroyd *et al.* (2002) in an Australian study looking at the prolificacy of natural service *Bos indicus* bulls.

Scrotal circumference (SC) EBVs were positively related to prolificacy (P<0.01). Repeatability of prolificacy for full season bulls with data for more than one breeding season was 0.43 (±0.08). This is similar to the repeatability values of 0.43-0.69 reported by Holroyd *et al.* (2002). There was a trend (P=.14) for older bulls to sire an increased number of progeny. Although most bulls in the study were Angus a small number of South Devon, Hereford and Red Angus bulls were also used on these ranches. The calves that were sired by South Devon (n=217) and Hereford (n=145) bulls were on average 20.4 kg and 16.4 kg heavier than Angus-sired calves at weaning presumably due in part to heterosis in these herds with high-grade Angus commercial females. Irrespective of hybrid vigor, prolificacy was the main driver of total calf weight weaned per sire.

These retrospective data do not indicate specific bull management practices to enhance prolificacy, other than perhaps attention to SC EBV in commercial bull selection. Measuring scrotal circumference has consistently been reported as a useful method for assessing reproductive function in bulls (Burns *et al.* 2011). Previous work suggested a separate multiple-sire breeding pasture for yearling bulls would be advantageous as yearling bulls in mixed-aged sire groups sired few if any progeny (Van Eenennaam *et al.* 2007).

BULL SELECTION: GENOMIC (gMBV), BREED (bEBV), and RANCH (rEBV) EBV

Criteria for bulls used in this analysis (N=89) were Angus breed background, availability of a DNA sample for high density genotyping, and verified progeny phenotyped for weaning weight and carcass traits produced in the first season of data collection and at least one subsequent season. Genomic breeding values (gMBV) for weaning weight, carcass weight, ribeye area, backfat thickness and marbling were derived from single breed (Angus) and multi-breed training populations obtained from collaborators at Iowa State University/University of Missouri-Columbia (Weber *et al.* 2012b), and the US Meat Animal Research Center (Weber *et al.* 2012a), respectively. The gMBV were compared to Angus breed association breeding values (bEBV) available for bulls at the time of purchase, and single-season ranch breeding values (rEBV) derived from one season of progeny phenotypic data (n=1785). Using selection index theory, the distribution of future progeny performance was estimated for each EBV as normally distributed with a mean of one-half the bull's EBV and a variance of $(1-0.25*h^2*r^2)$ times the phenotypic variance of the trait, where h^2 is the trait heritability and r is the EBV accuracy. The likelihood of the observed adjusted progeny performance in future seasons (n=4108) was estimated for each EBV and then EBV were ranked on this likelihood estimate for each sire.

The frequency at which each EBV was found at a given rank was compared against that which would be expected given a random assortment using a chi-square test. EBVs were not ranked randomly (P<0.05); in general the rEBV was the most predictive EBV and differences between the rank of Angus-trained gMBV and bEBV were non-significant. This suggests that commercial ranch genetic evaluations based on a single season of data (i.e. progeny testing) are more predictive than either the currently available gMBV or traditional pedigree-and phenotype-based breed association bEBV. Weaber (2005) found rEBV for herd sires derived from one season of phenotypic progeny testing generated value by improving the response to selection for targeted traits. The return on investment that results from such progeny testing was found to be greatly influenced by the cost of parentage determination. If the cost of SNP genotyping continues to decrease, the use of DNA-based parentage to develop rEBV may offer commercial producers a cost-effective approach to obtain genetic evaluations on commercial and ranch-developed bulls.

HEIFER SELECTION

Commercial producers frequently have no EBV information upon which to base their replacement heifer selection decisions, and DNA testing offers an appealing approach to provide previously-absent selection criteria. Traits that are of the most economic value to self-replacing

herds are low heritability reproductive traits including age at first calving, reproductive success and replacement rate (Roughsedge *et al.* 2005). Research results suggest that large numbers of records will be required to obtain accurate DNA tests for low heritability traits (Goddard 2009). Further, such tests are the most difficult to validate as there is a paucity of cattle populations with sufficient phenotypic data to estimate the accuracy of new genetic tests for those traits. The value of using DNA information in making replacement heifer selection decisions will depend upon the information available at the time of selection, the accuracy (r) explained by the test, and the selection intensity (i.e. proportion of available heifers that are selected). The latter will be dependent upon the calving and replacement rates. In the absence of accuracy estimates it is not possible to model the value such tests might have for heifer selection. In practice, selection for replacement heifers is frequently driven by age and size as heifers that are born later in the calving season are too immature to be cycling in time for the first potential breeding season. This criterion tends to put indirect selection on fertility (i.e. selects for heifers that were conceived in the first estrus cycle). Additionally phenotypic considerations (feet, legs, udders, reproductive tract score, and pelvic area measurements) are likely to enter heifer selection decisions, further reducing selection intensity. In this study calves born during the first 21 days of the calving season were not randomly distributed among sires in multi-sire breeding pastures: highly prolific sires produced more early calves and hence their descendants were overrepresented in replacement heifers.

TECHNICAL DIFFICULTIES

During the course of this trial we encountered numerous technical difficulties of maintaining data integrity. Although in the field or at weaning we married the electronic ID and DNA barcode electronically, errors sometimes occurred emphasizing the need for a single "foolproof" DNA collection and animal identification system. Additionally in five consecutive Ranch A calf cohorts, the carcass misidentification rate in the processing plant ranged from 3.5 to 19.3%, with an average misidentification rate of 10.8% (Weber *et al.* 2012b). In this study paternity assignment of sampled calves using a 99 SNP panel was very high, but despite concerted efforts in working with the commercial producers DNA samples were not collected on 9.4% of the progeny with birth records. These considerations may influence whole herd results in commercial settings.

ACKNOWLEDGEMENTS

The author acknowledges the USDA National Institute of Food and Agriculture National Research Initiative competitive grant no. 2009-55205-05057 for funding support.

REFERENCES

Burns B.M., Gazzola C., Holroyd R.G., Crisp J. and McGowan M.R. (2011) *Reprod. Domest. Anim.* **46**: 534.

Goddard M (2009) *Genetica* **136**: 245.

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.

Roughsedge T., Amer P.R., Thompson R. and Simm G. (2005) *Anim. Sci.* **81**: 221.

Van Eenennaam A.L, van der Werf J.H. and Goddard M.E. (2011) *J. Anim. Sci.* **89**: 307.

Van Eenennaam A.L., Weaber R.L., Drake D.J., Penedo M.C., Quaas R.L., Garrick D.J. and Pollak E.J. (2007) *J. Anim. Sci.* **85**: 3159.

Weaber R. L. (2005) PhD Thesis, Cornell University, Ithaca, NY.

Weber K.L., Thallman R.M., Keele J.W., Snelling W.M., Bennett G.L., Smith T.P., McDaneld T.G., Allan M.F., Van Eenennaam A.L. and Kuehn L.A. (2012a) *J. Anim. Sci.* **90**: 4177.

Weber K.L., Drake D.J., Taylor J.F., Garrick D.J., Kuehn L.A., Thallman R.M., Schnabel R.D., Snelling W.M., Pollak E.J. and Van Eenennaam A.L. (2012b) *J. Anim. Sci.* **90**: 4191.

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:14-17

GENOMICS FOR THE AUSTRALIAN SHEEP INDUSTRY: FROM DESIGN TO DELIVERY

J.B. Rowe[1], S. Gill[1,3], R.G. Banks[1,4], and J.H.J. van der Werf[1,2]

[1]CRC for Sheep Industry Innovation, Homestead, UNE, Armidale, NSW, 2351
[2]School of Environmental and Rural Science, UNE, Armidale, NSW, 2351
[3] Sheep Genetics, UNE, Armidale, NSW 2351
[4]Animal Genetics and Breeding Unit, UNE, Armidale, NSW, 2351

SUMMARY

This paper tells the story of translating a scientific concept to commercial scale use of game changing genomic technology within a period of just 6 years. It has involved close collaboration between researchers and end-users of the technology, facilitated by the resources and cooperative structure of the Sheep CRC. The paper explores some of the challenges encountered and the ways in which these challenges were addressed. Key elements of success were considered to be (i) delivery via existing modes, i.e. using estimated breeding values (EBVs) provided through Sheep Genetics, as breeders were familiar with the terminology and source of information; (ii) the delivery of EBVs for new traits created a huge amount of excitement, especially for terminal breeders (they do not have this in most other places, usually same traits); and (iii) the close engagement of breeders as potential end-users of the new technology.

THE INFORMATION NUCLEUS PROGRAM

The project was designed in 2006 when it became clear that searching for genes of major effect in extensive livestock had been largely unproductive and that a whole genome approach in the form of genomic selection was likely to be a more useful approach. The aim was to create a resource for estimating genetic parameters for new traits that could potentially be added to a breeding program using genomic selection, and at the same time serve as a reference population that could be used to predict genomic breeding values of young industry sires, both for existing traits and for those new traits that were considered. This was a world first design in which progeny of sires selected to maximise diversity and genetic potential were measured for the broadest range of traits considered to have commercial or scientific relevance (Banks *et al.* 2006). The size of the Information Nucleus (IN) was chosen to achieve both clear proof of concept but also to have useful accuracy of genomic prediction – even for new traits that had not previously been measured (Van der Werf *et al.* 2010)

Detailed protocols for measurement of each trait were developed by each discipline group across a range of institutions to ensure that all data were collected using identical procedures. Achieving agreement on protocols, submission of data, checks on accuracy and payment details across 6 organisations required a full time project coordinator as well as a database manager but this degree of engagement turned out to be a key to the success of the project. Rapid transfer of data to a central database, prompt data checking and a well-coordinated team of analysts delivered quick feedback of new information to a range of stakeholders as well as providing the basis for a number of research publications. IN data flowed into the database of Sheep Genetics to become part of the national system for genetic evaluation and to add to accuracy of estimated breeding values (EBVs) of many industry sires. Moreover, within a few years, EBVs for new traits were released, and with genotype information based on the 50K Ovine chip (Illumina Inc., San Diego, CA, USA) genomically enhanced EBVs could be released for new traits that had not been measured before on-farm. Figure 1 summarises the flow of data collection across discipline areas as well as its analysis, interpretation and processing to facilitate commercial delivery of new information to ram breeders. It also shows the extremely important step of a pre-determined 'route

to market' through the existing Sheep Genetics organisation with the infrastructure, expertise and industry credibility to introduce new breeding values and introducing the use of genomic information.

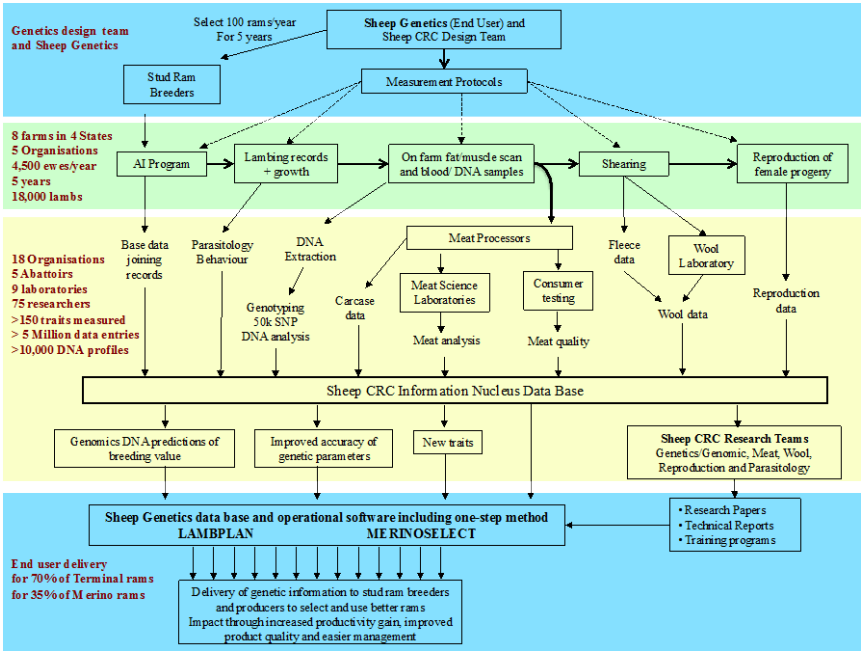


Figure 1. Diagram summarising Information Nucleus program data collection and interpretation

END-USER ENGAGEMENT

A key component of the success of this initiative has been the early engagement with end-users and maintaining their interest and participation throughout the program. The first step was to work with leading ram breeders to encourage them to provide semen from leading young rams representing a wide range of genotypes from all major breeds. Commitment by breeders to collect semen and provide it to the CRC at significantly discounted rates, was matched by an undertaking by the research team to provide accurate estimates of breeding values on these rams based on their progeny for a very wide range of traits. Ram breeders could see information on their rams changing each fortnight through updates to the Sheep Genetics web site. This near real-time availability of new information provided an unexpected benefit in that breeders were able to provide valuable feedback to the researchers whenever new data entry resulted in unexpected changes to breeding values. This resulted often in adjustment of protocols and data checking routines.

Breeders became comfortable about the reliability of the IN data and its value in contributing to EBVs for new traits. Many of the sires used in the Information Nucleus program were then widely used in industry to link into other evaluations such as young sire programs. This in turn brought further information into the national database for the rams being used in the Information Nucleus and rapidly expanded the impact of the program. The fact that genomic breeding value from the IN program were separate from Sheep Genetics estimates of Australian Sheep Breeding

Values (ASBVs) for industry rams allowed unbiased validation and the ability to produce unbiased blended breeding values.

The extensive range of measurements on existing and new traits provided the resource material for researchers working in many areas of biology and genetics. For example, the impact of the current breeding program on meat quality traits became clarified and new EBVs gave tools to the industry to turn around any potentially negative correlated selection responses. An improved understanding of the biology of meat quality and delivery of EBVs for new traits created a lot of excitement - particularly for terminal breeders.

COMMUNICATION

From the commencement of the program there has been a well-structured communication plan to keep all stakeholders well informed of the progress, e.g. via letters to breeders who had supplied rams; quarterly newsletters to all organisations involved in the Sheep CRC; and regular media releases with information on significant developments.

Awareness of the potential use of genomic technologies in the dairy and beef industries was a benefit as well as a potential risk. The communication plan was therefore structured in a way that built on the background knowledge from other industries without over-selling the potential benefits. Communicating a time-frame for delivery of new genomic technologies needed a careful balance between maintaining a sense of urgency for the research team and realistic industry expectations.

Field days and the appointment of industry advisory groups played an important role in connecting local producers and ram breeders with the program.

ESTIMATING BENEFITS AND COSTS

Initial estimates of benefits of genomic selection to the Australian sheep industry were reported by Van der Werf (2009) and Banks and Van der Werf (2009). A project initiated by Meat and Livestock Australia set out to document potential benefits of using genomic technologies to individual breeders. This project developed a number of case studies in conjunction with leading ram breeders. Through workshops with the group and with detailed analysis and modelling with each breeder the case studies provided guidelines for cost effective strategies for genotyping and some insight in the price structure needed to make the new technologies commercially attractive to ram breeders. It also provided guidelines for breeders to consider when planning their investment in genomic testing.

GENOMIC PILOT PROJECTS

In early 2010, just 3 years into the project, it became clear that the genomic predictions of breeding values were sufficiently accurate to warrant industry scale testing of the new technology. A subsidized genomic test ($50/test) was offered to clients of Sheep Genetics and a total of 460 rams were tested. One of the major challenges in this first pilot project was the complexity of end-to-end sampling, analysis and reporting. Setting up new protocols for each component of the process was a big job and there were also delays due to a very extensive review of the new genomic breeding values prior to their release.

In 2011 the second pilot project was advertised at the same price of $50/sample. At this stage there was a greater range of traits able to be predicted and increased accuracy for all traits offered. A total of 860 rams were tested and all results were reported well within the 12 week turnaround time specified at the time the tests were offered. Feedback on the service and on the value of the results to breeders was generally positive. In this second trial, DNA sampling was via blood cards, which proved to be a lot more efficient than earlier ear punches or nasal swaps. The sample tracking system was much smoother, and the analysis pipeline was now setup to communicate

much more efficiently with the large industry phenotype and pedigree database, as well as the CRC database with genotypic and phenotypic information from the reference population.

In 2012 the third pilot project was advertised in two parts, one for small-scale sampling (between 10 – 20 tests per breeder) and a commercial scale genotyping/breeding component in which breeders and groups of breeders working together could order up to 200 tests. The cost per test was again set at $50. The 3,000 tests allocated to the 2012 project were fully subscribed within three weeks. An on-line ordering system worked well and many aspects of the sample handling and reporting were also fully automated.

An important component of the 2012 pilot project has been the technical support provided for the 11 breeders participating in the 'commercial scale' component. Each breeder, or group of breeders, has had access to a range of specialist geneticists during planning workshops as well as one-on-one access to a consultant able to help optimise the design of their breeding program to take advantage of genomic information. These case studies will be available for others to use in modifying current breeding programs to include genomic information.

The main benefit of the pilot projects has been to prepare ram breeders and Sheep Genetics for commercial scale application of genomic technology. A second benefit has been the collection of additional genomic information on a wider range of industry rams that have subsequently been widely used and thoroughly measured. This information has been valuable in terms of improving accuracies for some of the more difficult to measure traits such as number of lambs weaned.

CONCLUSIONS - COMMERCIAL TESTING AND FUTURE CHALLENGES

In 2013 a commercial service is being offered by Sheep Genetics based on a combination of a "standard density" SNP test (12k SNP) in combination with the use of 50k SNP tests on some key flock sires. The commercial cost of the test will be around $50 and is consistent with subsidised prices paid in previous years in the pilot projects.

Sheep Genetics client base directly covers approximately 80% of rams produced for prime lamb production and over 30% of Merino rams. It is expected that the results and benefits of the new genomic technologies extend through a large proportion of the rams evaluated through the Sheep Genetics program and therefore through the Australian sheep industry. In large part the benefit of this initiative is a result of the relatively short lead-time from initiating the research to commercial availability of the new product.

Funding of an ongoing resource flock program to maintain genomic prediction accuracies for difficult to measure traits is currently provided through MLA. There is the possibility of increasing numbers of breeders electing to use only genomic information or reducing the amount of data collected in commercial breeding programs. A sustainable financial model is needed whereby a fair value is determined for physical data collection and/or a margin imposed on genomic tests to cover ongoing maintenance of genomic prediction accuracies.

REFERENCES

Banks R.G., van der Werf J.H.J. and Gibson J.P. (2006) *Proc. 8th World Congr. Genet. Appl. Livest. Prod*., Belo Horizonte, Brazil. 30-12.

Banks R.G. and van der Werf J.H.J. (2009) *Proc. Assoc. Advmt. Anim. Breed. Genet.* **18**: 430

Van der Werf J.H.J. (2009) *Proc. Assoc. Advmt. Anim. Breed. Genet.* **18**: 38

Van der Werf J.H.J., Kinghorn B. P. and Banks R.G. (2010) *Anim. Prod. Sci.* **50**: 1

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:18-21

A SURVEY EXAMINING THE NEW ZEALAND BREED COMPOSITION, MANAGEMENT TOOL USE AND RESEARCH NEEDS OF COMMERCIAL SHEEP FARMERS AND RAM BREEDERS

R.A. Corner-Thomas[1], P.R. Kenyon[1], S.T. Morris[1], A.W. Greer[2], C.M. Logan[2], A.L. Ridler[1], R.E. Hickson[1], H.T. Blair[1]

[1] International Sheep Research Centre, Massey University, Private Bag 11222, Palmerston North, New Zealand 4442

[2] Faculty of Agriculture and Life Sciences, PO Box 85084 Lincoln University, Lincoln, New Zealand 7647

SUMMARY

This survey has highlighted the extent of cross-breeding and wide-spread use of composites, particularly with the use of Finnish Landrace, Texel and East Friesian breeds within the New Zealand sheep flock. Forty percent of the flocks in this survey were composite flocks indicating the level of cross breeding that has occurred in the last 20 years. Overall, there was greater use of management tools by ram breeders than commercials farmers, although the use of some tools was not as great as expected with BVs used by only 22% of commercial farmers and 59% of ram breeders. Lastly, this survey outlined those areas farmers perceive as warranting more research which are primarily those that directly affect farm income.

INTRODUCTION

In recent times there have been considerable changes to the structure and productivity of the New Zealand sheep flock. The New Zealand sheep flock through the 1900's was dominated by the Romney breed. In the late 1980's the national flock, of 60 million sheep, consisted primarily of the Romney, Coopworth and Perendale breeds which made up 46, 13, and 8% of the flock, respectively (Stewart and Garrick 1996). At this time the national lambing percentage was 102% and the average carcass weight was 13.5 kg (NZMWES 1988). In 2011, the national flock had been reduced to 31.9 million but achieved a lambing percentage of 122% and an average carcass weight of 18.25 kg (Beef + Lamb NZ 2013). The net effect of these increases in individual performance traits is that the total amount of lamb meat produced now is very similar to that produced in the late 1980s (Bray 2004). Although the scale of the improvements in productivity in the New Zealand sheep flock during this short time are impressive, this has only been possible due to a multitude of factors. There have been considerable changes in land-use at either end of the spectrum with marginal country either retired through the land tenure review process or planted into forestry while large areas of more fertile land have been converted to dairy, viticulture or consumed within urban sprawl. In addition, there has been a gradual and continued increase in on-farm productivity as a result of improved managerial capability and animal genetic merit. There is an increasing array of managerial tools and access to information. In addition, there has been a considerable increase in the utilisation of cross-breeding following the introduction of 'Exotic' breeds such as Finnish Landrace, Texel and East Friesian in the early 1990s (Blair 2011). However, while these advances have been accessible by farmers, there is minimal information available on the uptake of such managerial tools or the impact cross-breeding has had on the number of composite flocks in New Zealand.

The ultimate goal of sheep research is to provide information and or tools that will assist with improving productivity. The adoption and utilisation of this research is dependent on the perceived benefits accrued by the end-user. However, little information is available on what New Zealand farmers rate as important areas of research. Research programmes are typically based on either the beliefs of scientists, or a few 'focus group farmers', and which may be driven by the

strategic direction of funding bodies. Knowledge of what farmers perceive to be important research areas will allow for better use of available research funds and assist with the development of managerial tools which encourage farmer uptake and provide the greatest benefit.

The purpose of this survey was to determine the current genetic structure of the New Zealand sheep flock, the use of management tools and the perceived research needs of sheep farmers in New Zealand.

METHODS AND MATERIALS

A printed survey was sent to approximately 12,000 sheep and beef farmers whose addresses were on the Beef + Lamb New Zealand database. The survey was included within the 'Heartland Sheep (NZX Agri, Feilding New Zealand) magazine in October 2012. Farmers had the opportunity to either, fill in the survey and return it via a pre-paid envelope, or to fill it in electronically via a website "www.SurveyMonkey.com". A total of 971 surveys were returned (934 by post and 37 completed online).

Part A of the survey asked farmers to identify themselves based on their farm type (ram breeder or commercial farm) and the breed(s) of sheep on their farm. If a farmer indicated they had both a stud flock and a commercial flock on their farm they were classified as being a ram breeder (94 vs. 844 ram breeder and commercial farmers respectively).

In Part B of the survey farmers were asked to indicate if they had used a range of management tools in the last three-years on their farm. In addition, they were asked to indicate on a one to four scale (1 = not important, 2 = little importance, 3 = important, 4 = very important) the relative importance of potential research areas for their sheep enterprise.

The proportion of respondents that selected a particular management tool was analysed using the Genmod procedure using a binomial distribution and a log-transformation (SAS 2011) and included the fixed effect of farm type. A comparison of the score given to each research area was analysed using the Genmod procedure using a Poisson distribution and a log-transformation and included the fixed effects of farmer age and farm type.

RESULTS AND DISCUSSION

Ewe breeds / composites. The 971 farmers, that completed the survey, identified a total of 1306 flocks present on their farms (700 farmers had 1 flock, 161 had 2 flocks, 53 had 3 flocks and 28 had 4 or more). Of these flocks, there were 780 straightbred (26 individual breeds were listed) and 526 composite flocks. Romney was the dominant straightbred breed (n=369, 47%) followed by Perendale (n=114, 15%), Coopworth (n=87, 11%) and Merino (n=36, 5%). This finding is similar to that of Blair (2011) who reported, that of flocks listed on SIL, Romney made up 35%, Perendale 14% and Coopworth 13%. In contrast, Stewart and Garrick (1996) reported that in the 1989 breed census 59% of registered ewes were Romney, 16% Coopworth and 10% Perendale.

Farmers also identified 135 terminal straightbred flocks (12 individual breeds were listed), with the most numerous being Texel (n=43), Suffolk (n=21), and Poll Dorset (n=17). In the present study, the Texel represented 30% of the terminal straightbred flocks, indicating the success of this breed since first being introduced in the early 1990s. By comparison, the prevalence of Finn (n=8) and East Friesian (n=4) breeds reported in this survey was relatively low.

Of the 526 composite flocks 449 were Romney based (these included composites that were Coopworth or Perendale based). The vast majority of the composites could be classed as a maternal type (n=451) compared with terminal type (n=49). Within the composites, 220 (42%) had Texel, 111 (21%) had Finn and 52 (10%) had East Friesian genetics. In addition, of the composite flocks that contained Finn, East Friesian or Texel genetics, 89 had two of these types and 10 had all three. Overall, in this survey 40% of the total flocks listed were composites, indicating their relative importance within the New Zealand sheep industry.

Farm management tools used. The percentage of either commercial farmers or ram breeders

that have used various management tools is given in Table 1. The management tools that were most frequently used were ewe teeth and udder examination, ultrasound pregnancy diagnosis, and weighing of sale lambs, all being used by more than 71% of respondents. In comparison, the tool least used was EID, being 4% of commercial farmers and 15.6% of ram breeders. Overall, for the management tools listed, they were more likely to be utilised by ram breeders than commercial farmers. The only exceptions being ewe teeth examination, ewe feet examination, ewe body condition scoring and ultrasound pregnancy scanning, largely due to the fact that nearly all of these were utilised by a high percentage of both commercial farmers and ram breeders. Increased use of management tools by ram breeders was largely anticipated and presumably reflects the greater collection of phenotypic data to assist with selection decisions compared with commercial farmers. Of particular note was the utilisation of breeding values (BVs). While the utilisation of BVs by ram breeders was nearly three-fold that of commercial farmers, more than 40% of breeders do not use this as a management tool. The reasons for the relatively low use of BVs by ram breeders could not be determined from the current survey but is worthy of further investigation. The lower use of BVs by commercial farmers is either because they do not readily identify the benefit from using BVs when selecting rams or that they rely on their ram breeders to do this for them.

Table 1. The percentage of respondents that indicated they had used the management tools listed on their operation within the previous 3 years (transformed mean ± SEM (back-transformed %))

Management tools	Commercial (n=844)	Ram breeder (n=96)	Commercial vs Ram breeder
Non EID Ear tags	-1.18 ± 0.08 (23.6 [1]) [b]	0.79 ± 0.22 (68.8) [de]	P<0.001
EID ear tags	-3.17 ± 0.18 (4.0) [a]	-1.69 ± 0.28 (15.6) [a]	P<0.001
Ewe teeth examination	1.96 ± 0.10 (87.7) [h]	1.95 ± 0.31 (87.5) [f]	P=0.970
Ewe feet examination	0.26 ± 0.07 (56.5) [e]	0.69 ± 0.22 (66.7) [cde]	P=0.057
Ewe udder examination	1.27 ± 0.08 (78.1) [g]	2.15 ± 0.33 (89.6) [f]	P=0.011
Weigh ewes	-0.62 ± 0.07 (35.1) [c]	-0.17 ± 0.20 (45.8) [b]	P=0.039
Ewe body condition scoring	-0.30 ± 0.07 (42.5) [d]	-0.17 ± 0.20 (45.8) [b]	P=0.541
Weigh sale lambs	0.91 ± 0.08 (71.3) [f]	1.77 ± 0.29 (85.4) [f]	P=0.004
Weigh replacements	-0.32 ± 0.07 (42.1) [d]	0.56 ± 0.21 (63.5) c[d]	P<0.001
Breeding Values	-1.28 ± 0.08 (21.8) [b]	0.38 ± 0.21 (59.4) [cd]	P<0.001
Mating harness	-1.16 ± 0.08 (23.9) [b]	0.17 ± 0.20 (54.2) [bc]	P<0.001
Ultrasound pregnancy scanning	0.93 + 0.08 (71.7) [f]	1.04 + 0.23 (73.9) [e]	P=0.644

Means within columns with differing letter superscripts are significantly different P<0.05

Perceived research requirements. The perceived future research needs of respondents are given in Table 2. Those research areas that affected farmer income directly (improved lamb survival, live weight gain in young stock, and reproduction) or that affected cost and influenced animal performance (health/disease, soils/fertiliser, and nutrition) scored at a higher level. In comparison, those areas that have less direct relevance to farm performance received a lower score (animal welfare/behaviour, economic and systems modelling, environmental/sustainability and forages/agronomy). This information, across a significant number of farmers, could help prioritise future research strategy to better match the perceived needs of the intended end-user.

The only differences in perceived research needs between commercial farmers and ram breeders occurred in the areas of environmental/sustainability and genetics/genetic technologies. The reasoning for the latter can be expected given the greater prevalence of BVs as a management tool by ram breeders and also the expectation that the benefits accrued through utilising BVs will be greater and more apparent to ram breeders than commercial farmers. The reasoning for the perceived differences in need for research in the areas of environmental/sustainability between commercial and ram breeders is unclear.

Table 2. The number of respondents that provided a rating to each research area and the rating given (transformed mean ± SEM (back-transformed mean score))

	Commercial (n=844)		Ram breeder (n=96)		Commercial vs. Ram breeder
Research areas [1]	**n**	**Score**	**n**	**Score**	
Animal Welfare/Behaviour	810	1.04 ± 0.02 (2.8 [1]) [cd]	95	1.06 ± 0.06 (2.9) [ab]	P=0.739
Economic and systems modelling	774	0.90 ± 0.02 (2.5) [a]	92	0.92 ± 0.07 (2.5) [a]	P=0.794
Environmental/ Sustainability	801	1.01 ± 0.02 (2.7) [bc]	93	1.14 ± 0.06 (3.1) [bc]	P=0.038
Forages/Agronomy	787	1.02 ± 0.02 (2.8) [bcd]	90	1.07 ± 0.06 (2.9) [ab]	P=0.495
Genetics/Genetic technologies	792	1.07 ± 0.02 (2.9) [d]	91	1.24 ± 0.06 (3.5) [c]	P=0.005
Health/Disease	819	1.22 ± 0.02 (3.4) [fg]	94	1.27 ± 0.05 (3.6) [c]	P=0.332
Lamb Survival	824	1.25 ± 0.02 (3.5) [g]	92	1.27 ± 0.06 (3.6) [c]	P=0.720
Live weight gain in young stock	815	1.22 ± 0.02 (3.4) [fg]	94	1.24 ± 0.06 (3.5) [c]	P=0.745
Meat yield and quality	804	1.16 ± 0.02 (3.2) [e]	93	1.24 ± 0.06 (3.4) [c]	P=0.224
Nutrition	809	1.19 ± 0.02 (3.3) [f]	94	1.22 ± 0.06 (3.4) [bc]	P=0.649
Reproduction	804	1.19 ± 0.02 (3.3) [ef]	93	1.21 ± 0.06 (3.3) [bc]	P=0.742
Soils/Fertiliser	824	1.24 ± 0.02 (3.4) [fg]	94	1.22 ± 0.06 (3.4) [bc]	P=0.775
Wool	805	0.98 ± 0.02 (2.7) [b]	90	0.95 ± 0.07 (2.6) [a]	P=0.716

Means within columns with differing superscripts are significantly different P<0.05

[1] Back-transformed mean score. Mean value, 1 = not important, 2 = little importance, 3 = important, 4 = very important

ACKNOWLEDGEMENTS

Funding for this work was provided by Massey-Lincoln and Agricultural Industry Trust and the Gravida (the National Centre for Growth and Development).

REFERENCES

Beef + Lamb NZ (2013) Annual Report 2011-2012.

Blair H.T. (2011) *Proc. Assoc. Advmt. Anim. Breed. Genet.* **19:** 407.

Bray A. (2004) *Prim. Indus. Manage.* **7:** 32.

NZMWES (NZ Meat & Wool Boards' Economic Service) (1988) Annual review of the New Zealand sheep and beef industry 1987-1988. Wellington, New Zealand.

SAS (2011) Statistical Analysis System. SAS Institute Inc., Cary, NC, USA.

Stewart J.B., Garrick D.J. (1996) *Proc. N. Z. Soc. Anim. Prod.* **56:** 301.

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:22-24

DR DAVE JOHNSON'S CONTRIBUTION TO DAIRY CATTLE GENETIC EVALUATION IN NEW ZEALAND

B. Harris

Livestock Improvement Corporation, Private Bag 3016, Hamilton, New Zealand

SUMMARY

Dr Dave Johnson's contributions to New Zealand's dairy cattle genetic evaluation over the last 20 years are outlined. Five major research areas are highlighted: animal model introduction, variance component estimation, reliability of breeding values, on-demand test-day model results and genomic selection. A key measure of the success of Dr Johnson's research is that many of the findings and systems are still being used every day in New Zealand and around the world.

INTRODUCTION

In this paper, the contributions of Dr Dave Johnson to New Zealand's dairy cattle genetic evaluation will be discussed. This review will cover the period from 1993 to the present. Over this time there have been major changes in the genetic evaluation of dairy cattle. In the early 1990s many countries, including New Zealand, were adopting the animal model methodology for routine genetic evaluation. This was followed by the broadening of breeding objectives to include fertility, survival and other functional traits, which necessitated the need for genetic evaluations of these traits at a national level. The last decade has seen the adoption of the test-day model methodology for routine genetic evaluation of the production traits in dairy cattle. The latest change has been the adoption of genomic selection in dairy cattle that took place over the last six years. LIC, in collaboration with industry partners has been responsible for the research, development, maintenance and routine operation of the New Zealand's dairy cattle national genetic evaluation over the last 20 years. Dr Johnson has been closely involved in, and has made significant contributions to, all these research areas throughout his career at Livestock Improvement Corporation (LIC). Rather than list all the contributions over this 20-year period, I will highlight five areas that have had a major impact.

ANIMAL MODEL INTRODUCTION

In June 1996, a new animal model was implemented in New Zealand whereby dairy cattle were genetically evaluated using an across-breed animal model (Harris, 1996 and Garrick *et al*., 1997). The animal model used was a single-trait repeatability model. The model allowed for heterogeneous subclass variation, that arose from the inclusion of mixed breed contemporary groups, by including fixed breed and group effects directly in to the additive relationship matrix. Also, a new methodology to predict total lactation yields from individual test-day information was developed to provide phenotypic production records for the mixed model analysis. This method was developed by Dr Johnson during 1994 and 1995. This method accounted for any number of herd tests over any testing frequency and allowed for variable information among herd-mates and for the effects of culling (Johnson 1996). The advantages of this approach were that all the phenotypic records were adjusted to a common 270-day length, additionally, a computationally efficient method to calculate the accuracy of prediction which weighted the records according to the amount of available information was devised. Although this phenotype is no longer used in the production national genetic evaluation it is still used to provide milk volume, fat yield and protein lactation yield phenotypic records are still used in a multiple-trait model to calculate fertility and longevity breeding values (BVs).

VARIANCE COMPONENT ANALYSIS

An algorithm was described that estimated variance components for a univariate animal model using REML in a paper by Dr Johnson and Professor Robin Thompson titled "Restricted maximum likelihood estimation of variance components for univariate animal models using sparse matrix techniques and average information" published in 1994 (Johnson and Thompson 1995**).** Sparse matrix techniques were employed to calculate those elements of the inverse of the coefficient matrix required for the first derivatives of the likelihood. The method made use of an average information (AI) restricted maximum likelihood (REML) algorithm. At that time, variance components were commonly estimated using derivative free (DF) REML methods or expectation-maximization (EM) algorithms. The AI REML procedure was found to be about 5 times faster than DF method and about 15 times faster than EM algorithm. The research lead Dr. Johnson to develop variance component software based on the AI REML algorithm. Both univariate and multivariate analyses could be undertaken with this software. This software was subsequently used to estimate genetic and phenotype parameters from the data recorded in progeny test herds. These analyses were based on 100,000s of records across multiple traits. These types of analyses would have been computationally infeasible with any of the other available software at that time. Many of the estimates of the genetic and phenotype parameters are still in use in the current national genetic evaluation.

CALCULATION OF RELIABILITY OF BREEDING VALUES

The reliability of a BV is a measure of its accuracy. Exact reliabilities can be calculated from the inverse of the mixed model equations. However, in national evaluations the mixed model contains more than 10,000,000 equations, making them computationally infeasible to invert. A method of approximating reliability that was computationally fast and provided estimates with low amounts of bias was needed for the national evaluation. A new method was developed for calculating approximate reliability of breeding values for national systems of evaluation by Johnson and Harris in 1998 (Harris and Johnson 1998a). The method combined the reliability of three information sources: parent average, animal's own records, and progeny records. This method provided good approximation to the actual values with minimal upward bias and was better than the current method used in New Zealand at that time. One of the major advantages of the method was that it could be extended to accommodate more complex models by altering the selection index equations within the method. This method has been now been extended to several complex models including test-day models (Ducrocq and Schneider 2007) and genomic selection models (Harris and Johnson 2010). This method was also extended to the estimation of reliability for Interbull multiple across country sire genetic evaluations (MACE) (Harris and Johnson 1998b). The motivation was to address concerns relating over-estimation of MACE reliabilities and its flow-on effects on the weighting of foreign information being included in national genetic evaluations. The proposed method was found to be a significant improvement over the method used at that time. The method was adopted by Interbull in 1999 and still used.

ON DEMAND TEST-DAY MODEL

In 2007, a testday model (TDM) was developed to provide national genetic evaluation for the production traits (Harris *et al.* 2006). A TDM can simultaneously account for fixed environmental effects such as herd-testday contemporary groups, and genetic, permanent environmental (PE) and temporary environmental (TE) random effects. In the TDM model, each test day is modelled, thereby taking into account the temporary environment, resulting in an improved accuracy of evaluation over a lactation model. Also, the TDM can include functions of the cows' days in milk to account for cow-to-cow differences in the shape of the lactation curve. Dr. Johnson was an integral member of the development team. One aspect of the TDM development that has gone

unrecognised was Dr Johnson's on-demand TDM build for herd-testing customers. Because the national TDM model is only run approximately every 3 weeks, a system to provide updated results at the time of an individual herd-test was required thereby enabling farmers to make breeding and culling decisions based on the most up-to-date information. Dr. Johnson developed a simple system of selection index equations, that incorporated the latest herd-test results into the most recent TDM evaluation. The solution was computationally simple allowing up 1000 herds (300,000 cows) to be processed daily at the peak of the season.

GENOMIC SELECTION

Over the last five years, a considerable amount of research effort was being directed to the application of genomic selection to a national evaluation system for number of livestock species, including dairy cattle. In New Zealand we had the additional complication of requiring an across-breed genomic evaluation system in order to get genomic evaluations on progeny-tested Jersey Holstein-Friesian crossbred sires. Johnson and Harris (Harris and Johnson, 2010) describe a method for the prediction of breeding values incorporating genomic information in an across-breed evaluation. The first stage involved the prediction of genomic breeding values for genotyped individuals. The novel component of this method was the estimation of the genomic relationship matrix in the context of a multi-breed population. The paper outlines a selection index procedure to blend genomic predictions with the ancestral information that is lost between the process of deregression of the national breeding values and subsequent re-estimation using the genomic relationship matrix. Finally, the paper describes how the genomically-enhanced predictions are filtered through to non-genotyped descendants using a regression procedure. This is basis of the current national and LIC genomic selection methods used to select the genomic young bulls to be progeny tested.

CONCLUSIONS

This paper has only covered a small number of the contributions that Dr Dave Johnson has made to New Zealand's dairy cattle genetic evaluation over the last 20 years. However, each of these contributions has had a major impact. The many case the contributions are unique, in that they solve real research problems facing the New Zealand industry, such as the need for an across-breed genetic evaluation system and as such solutions were not readily available from the scientific literature. A key measure of the success of Dr Johnson's research is that many of the findings and or systems are still being used every day in New Zealand and around the world. Another important the contribution that Dr. Johnson has made is to his colleagues' work, most of which is unrecognised through co-authorship, by providing daily guidance across many areas within statistical research. His colleagues have found this guidance invaluable to their own research projects.

REFERENCES

Ducrocq V. and Schneider M. del P. (2007) *Proc. Interbull mtg* **37***:82*
Garrick D. J., Harris B. L. and Johnson, D. L. (1997) *Proc. Adv. Anim. Breed. Gen*. **12**:611
Harris B. L., Clark J. M. and Jackson R. G. (1996) *Proc. N. Z. Soc. Anim. Prod*. **56***:12*
Harris B. L., Winkelman A. W., Johnson, D. L. and Montgomerie W. A. (2006) *Proc. Interbull mtg* **35** *5:27*
Harris B. L. and Johnson D. L. 1998a. *J. Dairy Sci.* **81***:2723*
Harris B. L. and Johnson D. L. 1998b. *Proc. Interbull mtg* 17:81
Harris B. L. and, Johnson D. L. 2010. *J. Dairy Sci.* **93***:1243*
Johnson D. L. 1996. *Proc.N. Z. Soc. Anim. Prod.* **56***:16*
Johnson D. L. and Thompson R. 1995. *J. Dairy Sci.* **78***:449*

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:25-28

EFFECT OF DAUGHTER MISIDENTIFICATION ON DAIRY SIRE EVALUATION

A.M. Winkelman

LIC, Private Bag 3016, Hamilton, New Zealand

SUMMARY

The impact of daughter misidentification on dairy sire breeding value (BV) estimation was investigated by comparing sire progeny group means of DNA-verified cows to sire progeny group means of cows that had paternity determined through mating records. The daughters' BVs were adjusted for the dam contribution prior to the calculation of the means. BVs for milk volume, fat yield, protein yield, somatic cell score (SCS) and liveweight, and a five-trait index (breeding worth (BW)) containing these traits, were analyzed. Comparisons were done within sire breed (Friesian, Jersey and Friesian Jersey (FJ) Cross).

Estimates of progeny group means of the production traits of DNA-verified daughters were, on average, higher than those of daughters for which paternity had been assigned via mating records. The estimates ranged from 4.2 to 10.7 litres for milk volume, 0.16 to 0.30 kg for fat yield and 0.16 to 0.28 kg for protein yield. The differences between the progeny group means was close to zero for SCS, while the differences ranged from -0.076 to 0.18 kg for liveweight. The differences between the progeny group means of the five-trait BW were less than 2 BW points.

The magnitudes of the effect tended to increase with increased genetic merit. Higher genetic merit sires are likely to have greater bias than lower genetic merit sires. There was, however, considerable sire-to-sire variation in the difference between the progeny group means.

INTRODUCTION

Internationally, estimates of the percentage of cows that are misidentified to sire range from 5% (Ron *et al.*, 1996) to 23% (Gelderman *et al.* 1986). LIC proves their young sires in progeny test herds, referred to as Sire Proving Scheme (SPS) herds, prior to widespread use. In SPS herds, 95% of the cows are mated to young bulls and 5% are mated to proven bulls. Results from DNA paternity verification found that the rate of misidentification in SPS herds was 4.7%, 6.6% and 5.5% in seasons 2005, 2006 and 2007, respectively (assessed using 3602, 4427, 5120 sire-daughters tests, respectively, in the seasons) (Ric Sherlock personal communication). The expectation is that the percentage of misidentification is lower in SPS herds than in the non-SPS herds. Hence, the proportion of misidentified progeny is expected to increase from first to subsequent proofs.

A number of approaches have been used to assess the effect of misidentification of sires on genetic evaluation. Van Vleck (1970) used a deterministic model of the sire-daughter inheritance path to assess the effect of sire misidentification on genetic evaluation and estimates of genetic trends. He found that misidentification resulted in biased genetic evaluations and estimates of genetic trends. The bias increased with an increased proportion of misidentified daughters. Geldermann *et al.* (1986) also used a deterministic model (again considering only paternal pedigree errors) to show that a misidentification rate of 15% decreased accuracy of genetic evaluation, decreased estimates of heritabilities and reduced genetic gain. Estimates of the drop in genetic gain ranged from 8.7% to 16.9% for heritabilities of 0.5 and 0.2, respectively. Losses of similar magnitudes were found using stochastic simulation (Harder *et al.* 2005) and by Banos *et al.* (2001). Misidentification is expected to shrink the scale of the estimated breeding values (BVs). This is because the progeny that were incorrectly assigned to superior sires would more likely be the progeny of sires with a lower genetic merit than the top-end bulls. Hence, the top sires' genetic evaluations would be biased downwards. Similarly, progeny incorrectly assigned to sires of low

genetic merit, would more likely be the progeny of sires with a higher genetic merit than the lower-end bulls, thereby biasing the genetic evaluation of these sires upwards.

DNA verification of paternity involves comparing the DNA markers of an animal to those of its putative sire. LIC first offered a SNP-based DNA sire verification service to customers in the mid 1990s. Later, the service was provided by GeneMark. The test is based on approximately 100 SNPs, where the sire was deemed correct if the concordance between him and his progeny was at least 99%. The question arose as to whether data on cows for which paternity had been DNA-verified could be used to assess the impact of misidentification on the genetic evaluation of sires. The purpose of this study was to determine whether sire genetic evaluations based on cows that had paternity assigned via DNA verification differed to evaluations for which paternity was determined using mating records alone.

MATERIALS AND METHODS

The impact of misidentification on sire evaluation was assessed by comparing the sire contribution to daughters' BVs where paternity had been assigned via DNA verification to those that had been assigned using mating records. A sire's contribution to his daughter's BV can be partially determined by removing the dam's contribution to the BV. This approach does not remove the daughter's own Mendelian sampling (MS) contribution to the BV. However, if the MS is assumed to have a progeny group mean of zero, then averaging the sire's contribution across all his daughters within each of his progeny groups (i.e. DNA-verified and otherwise) should be a means of determining the impact of misidentified daughters on the sire's proof. No difference in the progeny group means would indicate that sire evaluation is not affected by misidentification of progeny. If the mean of the progeny genetic evaluation for the DNA-verified group is higher than that of the group that had paternity determined via mating records, then there is evidence that the misidentification is biasing the sire genetic evaluations downwards.

A total of 680,491 cows DNA-verified to sire were extracted from the national database. Of these, 392,677 had herd test records. These cows were the daughters of 4853 sires. All daughters of these sires were extracted from the national database. A total of 11,892,687 daughters had herd test information. Progeny of Friesian, Jersey and Friesian-Jersey (FJ) cross sires were retained for analysis. Edits were done to ensure that sires had at least five daughters in each progeny group (i.e. paternity assigned via DNA-verification or mating records). BVs for milk volume, fat yield, protein yield, somatic cell score (SCS) and live weight (hereafter referred to as milk, fat, protein, SCS and liveweight) were obtained for the daughters as well as their sires and dams. The BVs did not have Interbull or genomic information incorporated. The final data set included 3452 sires (1847 Friesians, 1159 Jerseys and 446 FJ crosses) with a total of 320,663 DNA-verified daughters and 8,618,574 daughters that had paternity assigned via mating records.

Equation [1] shows the components of a daughter's BV. Equation [2] shows the calculation of the daughter BV adjusted for the dam contribution (BVs_{adj}).

$$daughter\ BV = \frac{1}{2}\ sire\ BV + \frac{1}{2}\ dam\ BV + MS \quad [1]$$

where, MS is the Mendelian sampling; E(MS) = 0.

$$BV_{adj} = daughter\ BV - \frac{1}{2}\ dam\ BV = \frac{1}{2}\ sire\ BV + MS \quad [2]$$

The BVs_{adj} were calculated for all daughters for all 5 traits. Additionally, the adjusted five-trait Breeding Worth index (BW_{adj}) was calculated using the BVs_{adj} and the economic weights of published by NZ's Animal Evaluation Unit (AEU) in February, 2013 (Anonymous, 2013). The BVs_{adj} and BW_{adj}, were averaged over each sire and progeny group. Hence, every sire had two progeny means for each trait – one in which paternity was determined via DNA verification (DNA) and the other in which paternity was determined using mating records (REC)).

The effect of progeny group (DNA versus REC) was estimated using the linear regression of progeny mean on the progeny group and sire BV for each trait. A test of whether the magnitude of the estimate of progeny group was affected by the magnitude of the sire BV was done by including the interaction between the progeny group and sire BV.

RESULTS AND DISCUSSION

Table 1 shows the estimates of the progeny group effect within sire breed for milk, fat, protein, SCS, liveweight and BW. The model was parameterized so that the results are relative to the REC group. Hence, estimates greater than zero indicate that the mean of the DNA progeny group was higher than the mean of the REC group. The estimates were greater than zero for milk, fat and protein and close to zero for SCS. The liveweight mean was greater than zero for the Friesians and the FJ crosses but negative for the Jerseys. The estimates for milk follow the expected trend of being highest for the Friesians, lowest for the Jerseys and intermediate for the FJ crosses. The results for the FJ crosses are not intermediate between the Friesians and Jerseys for fat and protein. Nevertheless, estimates greater than zero are an indication that misidentification to sire biases the sires' BVs downwards.

Table 1. Estimates of effect of progeny group for milk, fat, protein, SCS, liveweight and BW[1]

Sire Breed	N	Milk (l)	Fat (kg)	Protein (kg)	SCS (log(cells/ml))	Liveweight (kg)	BW ($)
Friesian	1847	10.681***	0.271***	0.276***	0.00	0.175***	1.635***
Jersey	1159	4.214***	0.307***	0.191***	-0.001	-0.076	1.965***
FJ Cross	446	6.25***	0.164***	0.164***	0.001	0.142*	0.899***

[1]* = P<0.05, **=P<0.01, ***=P<0.001

Table 2 contains the estimates of the interaction of progeny group and sire BV/BW. The model is parameterized so that the values show the difference between the slopes of the sire BV in the DNA and REC groups. The estimates were small but positive indicating that the difference between the progeny group increases with increasing sire BV. The estimates were not significant for BW.

Table 2. Estimates of the interaction between progeny group and sire BV or BW[1]

Sire Breed	Milk (l)	Fat (kg)	Protein (kg)	SCS (log(counts/ml))	Liveweight (kg)	BW($)
Friesian	0.0125***	0.0070**	0.0045*	0.0079	0.0038***	0.0007
Jersey	0.0127***	0.0071*	0.0032	0.0070*	0.0084*	0.0035
FJ cross	0.0178***	0.0138**	0.0120**	0.0163**	0.0161***	0.0037

[1]* = P<0.05, **=P<0.01, ***=P<0.001

While the estimates of progeny group differences were positive, there was considerable variation in the difference between the DNA and REC means within sire. They ranged from ±200 litres for milk, ±10 kilograms for fat, ±8 kilograms for protein, ±0.4 to 0.4 units for SCS and ±4 kilograms for liveweight. Negative differences occurred for all breeds in both high- and low-BV sires. Such differences could arise from the fact that some sires had very few daughters in the DNA group and thousands in the REC group. The mean MS deviation of a small progeny group could differ markedly from zero. Additionally, there is likely sire-to-sire variation in the

percentage of daughters in the REC group that are misidentified. Nevertheless, the differences in the DNA and REC progeny groups suggest that misidentification does bias the sires' estimated BV.

Harris *et al.* (2007) found an annual per-cow genetic trend in NZ following the introduction of the BW was 2.5 kilograms of milk solids. The weighted (over sire breed) averages of the progeny group effect in Table 1 for fat and protein are 0.269 and 0.233 kilograms, respectively. Summing these gives a total of 0.502 kilograms of milk solids. This value is 20% of the annual genetic gain. The expectation is that top-end sires would be underevaluated for fat and protein by more than 0.5 kilograms. The underevaluation and reduced ability to identify extreme sires would have a negative effect on genetic gain.

The question remains as to what proportion of misidentified daughters could result in a sire's proof being underestimated by the amounts found in this study. Johnson (2010) used a simulation study to determine the effect of sire misidentification on the reproof effect. The reproof effect was found to vary with the percentage of parentage errors in the first and subsequent proofs. When the initial progeny test scheme had a 5% parentage error and the data used for the subsequent proof had 30% parentage error, with 80% of the daughters sired by other graduate bulls and 20% of the daughters sired by bulls with genetic merit equal to that of the cow population, the reproof effect for protein was -0.24 kilograms. An estimate of 0.23 kilograms of protein is a difference in a sire BV of 0.46 kilograms. This value is considerably higher than that found for the reproof effect. Sires evaluated following their initial proof may have in excess of 30% misidentified progeny in the commercial population. The next step of the research will involve estimating within-herd heritabilities for each trait and determining the association between the estimates and the level of sire misidentification as outlined by Dechow *et al.* (2007). Negative correlations suggest that the information on within-herd heritabilities can be used to identify herds that provide inaccurate data for sire evaluation.

CONCLUSIONS

A comparison of progeny group means of daughters that had paternity assigned via DNA verification versus mating records found that estimated BVs are, on average, biased downwards when all progeny are not DNA-verified. There is evidence that the effect increases with increasing sire BV. Higher genetic merit sires are likely to have greater bias than lower genetic merit sires.

REFERENCES

Anonymous (2013). http://www.nzael.co.nz/news/2013/2/2141742705/nzael-releases-revised-economic-values. Accessed 15 May 2013.

Banos G., Wiggans, G.R. and Powell R.L. (2001) *J. Dairy Sci.* **84**:2523.

Dechow C.D., Norman H.D., Zwald N.R., Cowen C.M., Meland O.M. (2007) *J. Dairy Sci.* **91**:1640.

Geldermann H., Pieper U. and Weber W.E. (1986) *J. Anim. Sci.* **63**:1759.

Harris B., Pryce J.E. and Montgomerie, W.A. (2007) *Proc. Assoc. Advmt. Anim. Breed Genet.* **17**:434.

Johnson D.L. (2010) A study of prior proofs and subsequent reproofs of graduate progeny-tested bulls. LIC Internal Report.

Ron M., Blanc Y., Band M., Ezra E. and Weller J.I. (1996) *J. Dairy Sci.* **79**:676.

Van Vleck L. D. (1970) *J. Dairy Sci.* **53**: 1697.

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:29-32

THE EFFECTS OF USING MALE AND FEMALE GENOTYPES IN GENOMIC EVALUATIONS

D.L. Johnson

LIC, Private Bag 3016, Hamilton 3240, New Zealand

SUMMARY

The objective of this study was to quantify additional accuracy of genomic evaluation from the addition of female genotypes to a dairy cattle population. The basic training set consisted of 6,150 progeny-tested bulls born prior to 2007 and the validation set consisted of 350 progeny-tested bulls born 2007-2008. Additionally, 36,350 female genotypes were included in the training population. The phenotypes were deregressed breeding values for production traits. Ridge regression was used with two models: (1) common SNP effects fitted for both genders and (2) SNP effects depending on gender with an assumed correlation. Bayes methods B and Cπ were also fitted under scenario (1). The accuracy of genomic evaluation was increased by 5 to 10 percentage points with the inclusion of female genotypes, depending on breed and trait. There was little difference in accuracy among models and methods of analyses.

INTRODUCTION

Genomic breeding values are now being widely used for bull selection in the dairy industry. One factor influencing the accuracy of genomic predictions is the size of the reference or training population. The relationship between predictive ability and the size of reference population has been demonstrated in Daetwyler *et al.* (2008) and Goddard and Hayes (2009). An option to increase accuracy of genomic evaluation is to combine reference populations from different countries (EuroGenomics, David *e tal.* 2010). Another option to boost the reference population is to genotype females. Apart from bull dams, LIC has a program of genotyping daughters of young bulls in the sire proving scheme (SPS) to maintain integrity of bull proofs through parentage testing.

MATERIALS AND METHODS

As at the end of the 2012/2013 season, LIC had a total of 6,500 progeny-tested bulls genotyped on the Illumina BovineSNP50 Beadchip (Illumina Inc., San Diego, CA). The validation population was taken as the 350 bulls progeny tested over the last two seasons (born 2007-2008). The base reference population comprised the remaining 6,150 bulls born 2006 and earlier. From a larger pool of genotyped cows, a total 36,350 with lactation records were included in the reference population. The cows comprised SPS daughters and their contemporaries as well as cows genotyped for other research purposes but excluded daughters of young bulls in the validation population. Most of the cows were genotyped on the Illumina 50K panel with some genotyped on a lower density GGP panel (6.5K) and then imputed to 50K using Beagle 3.3.2 (Browning and Browning 2009). There were 38,808 SNP included in the analyses after removing SNP for low call rates, minor allele frequencies <2%, non-Mendelian inheritance, failed Hardy-Weinberg tests and low imputation accuracy. The bull population was multi-breed comprising mainly Holstein-Friesian (HF), Jersey (JE) and crossbred (FJ) bulls. The bull reference comprised 56% HF, 34% JE and and 7% FJ and the validation population was correspondingly 39%, 30% and 24% reflecting the development of crossbred bulls in recent years. The cow population was a similarly structured multi-breed population with 32% HF, 23% JE and 40% FJ.

The SNP effects were estimated using multiple-regression models where the marker effects are treated as random. The model can be written

$$y=Xb+Zs+e \quad [1]$$

where **y** is phenotype, **b** denotes fixed effects (in this case just an overall mean), **s** denotes SNP effects, **X** and **Z** are design matrices and $E(\boldsymbol{y}) = \mathbf{Xb}, \boldsymbol{var}(\mathbf{s}) = \mathbf{I}\boldsymbol{\sigma}_s^2, \boldsymbol{var}(\boldsymbol{e}) = \mathbf{R}\boldsymbol{\sigma}_e^2$ **.** The mixed model equations (MME) corresponding to [1] are

$$\begin{bmatrix} X'R^{-1}X & X'R^{-1}Z \\ Z'R^{-1}X & Z'R^{-1}Z+\lambda I \end{bmatrix}\begin{bmatrix} \hat{b} \\ \hat{s} \end{bmatrix} = \begin{bmatrix} X'R^{-1}y \\ Z'R^{-1}y \end{bmatrix}$$

where $\lambda = \sigma_e^2/\sigma_s^2$. This ridge regression with known λ is equivalent to GBLUP when solving directly for genomic breeding values g=Zs provided we have the relation $\sigma_g^2 = \sigma_s^2 \sum_j 2p_j\,(1-p_j)$ between genetic variance and common SNP variance, with p_j denoting allele frequency of SNP *j*. This basic ridge regression model was used for the two reference sets; (1) bulls only and (2) bulls plus cows. In addition, for reference set (2), a mixture model approach was used for model [1]. A Bayes B model was fitted assuming that each marker has either a zero effect with known probability $\pi = 0.95$ or a non-zero effect with different λ values (Meuwissen *et. al*., 2001). A Bayes Cπ model was also fitted where one assumes a common λ but unknown π (Habier *et. al*., 2011).

For reference set (2), model [1] was extended to allow for different SNP effects depending on gender.

$$y=Xb+Z_1s_1 + Z_2s_2 + e \quad [2]$$

with

$$var\begin{bmatrix} s_1 \\ s_2 \end{bmatrix} = \begin{bmatrix} I & I\rho \\ I\rho & I \end{bmatrix}\sigma_s^2$$

where ρ denotes SNP correlation between genders and with common variance among SNP within gender and I is the identity matrix. The MME corresponding to model [2] are

$$\begin{bmatrix} X'R^{-1}X & X'R^{-1}Z_1 & X'R^{-1}Z_2 \\ Z'_1R^{-1}X & Z'_1R^{-1}Z_1 + \dfrac{\lambda I}{(1-\rho^2)} & \dfrac{-\lambda\rho I}{(1-\rho^2)} \\ Z'_2R^{-1}X & \dfrac{-\lambda\rho I}{(1-\rho^2)} & Z'_2R^{-1}Z_2 + \dfrac{\lambda I}{(1-\rho^2)} \end{bmatrix}\begin{bmatrix} \hat{b} \\ \hat{s_1} \\ \hat{s_2} \end{bmatrix} = \begin{bmatrix} X'R^{-1}y \\ Z'_1R^{-1}y \\ Z'_2R^{-1}y \end{bmatrix}$$

Two values of ρ equal to 0.7 and 0.9 were assumed. The BLUP models were solved using a conjugate gradient method.

Phenotypes were the the deregressed BV for the three production traits, milk volume, fat and protein yield, hereafter referred to as milk, fat and protein. The deregression procedure was carried out as in Garrick *et. al.* (2009). The elements of the diagonal **R** matrix associated with the error structure were calculated as $\left(c + \frac{1-r_i^2}{r_i^2}\right)\sigma_g^2/\sigma_e^2$ where *c*=0.1 is the assumed fraction of genetic variance unexplained by the markers and the second component is associated with the error variance of the deregressed BV with reliability r_i^2 for individual *i*. The constant c^{-1} also acts as an upper bound for the weighting applied to phenotypes corresponding to highly proven sires.

The validation procedure involved the regression of deregressed BV on genomic BV for the young bulls within breed as per Interbull procedure (Mantysaari *et. al.* 2010). The correlations were summarized as well as the regression coefficients to assess accuracy and bias of prediction. The accuracy attained through selection of the top 20 bulls on genomic BV was also investigated.

RESULTS AND DISCUSSION

The correlations between genomic BV and progeny-test BV for milk, fat and protein, based on the validation population, are summarised within breed in Table 1. The first data column is based on the bull reference while all other results relate to the combined reference. The inclusion of cows

in the reference has increased the correlations by an average of 0.09-0.10 for milk and fat and somewhat less at 0.05 for protein. Comparisons across columns of Table 1 indicate generally small differences among genomic evaluation methods using the combined reference population. In particular there appears little advantage to fitting marker effects by gender.

Table 1. Validation correlations for bull reference and combined (bull + cow) reference populations – RR=ridge regression, ρ is assumed correlation when fitting SNP effects by sex

Reference	bull	combined	combined	combined	combined	combined
Method	RR	RR	RR(ρ=0.9)	RR(ρ=0.7)	BayesB	BayesCπ
Fat						
HF	0.55	0.72	0.72	0.69	0.69	0.71
JE	0.62	0.64	0.66	0.66	0.61	0.63
FJ	0.50	0.60	0.62	0.62	0.60	0.60
Protein						
HF	0.50	0.54	0.56	0.55	0.54	0.56
JE	0.51	0.58	0.58	0.57	0.56	0.58
FJ	0.68	0.71	0.72	0.72	0.66	0.69
Milk						
HF	0.59	0.68	0.68	0.66	0.70	0.71
JE	0.54	0.70	0.67	0.64	0.69	0.69
FJ	0.74	0.76	0.77	0.77	0.76	0.77

The average reliabilities of the phenotypes for production traits were about 0.85 and 0.3 for bulls and cows, respectively. Based on the weighting formula, bulls would get an average weight of 3.6 and cows 0.4. Thus it takes about nine cows at that level of reliability to provide information equivalent to one progeny-tested bull and so 36,350 cows is equivalent to about 4,000 bulls. Based on the formula of Goddard and Hayes (2009), and assuming a heritability of 0.8 and effective population size of 100, the expected incremental change in the accuracy of genomic evaluation due to an increase of 4,000 bull equivalents above a base of 6,000 bulls is approximately 0.08. The results of this study suggest that the advantage of inclusion of the female data is close to expectation.

Table 2 summarises the regression coefficient of phenotype on estimated genomic BV for the ridge regression. The values represent a weighted average across breed. The expectation is unity and smaller values indicate some degree of inflation or bias in the genomic predictions. With the base reference set of bulls, the regressions are close to unity but decrease to about 0.8 for fat and protein when cows are included in the reference, suggesting some inflation. It is important to correct for this bias otherwise overestimation of genomic BVs will erode farmer confidence in genomic evaluations.

Table 2. Validation regression coefficient weighted across breed

Trait	Bull reference	Bull + cow reference
Fat	1.01±0.09	0.80±0.05
Protein	0.94±0.08	0.80±0.06
milk	0.97±0.07	0.95±0.05

The regression yields estimates of population parameters, however it is the animals at the top end of the distribution that are of interest. The top 20 bulls within breed were selected on genomic BV for the two reference populations. The change in average phenotype of these two groups of bulls (reference (2) minus reference (1)) is presented in Table 3 along with the number of bulls common to both groups. For each trait there was a positive change for two of the three breeds. The standard error of the difference (SED) between the averages of the two groups, assuming independence among bull proofs, is expected to be $\sigma_g\sqrt{2(n-m)(1/r-1)}/n$ where n=20 is the number of bulls selected, m is the number of bulls in common and r=0.75 is the daughter-proven reliability of an individual bull. Given genetic standard deviations of 329 litres, 13.6 kg, and 9.6 kg for milk, fat and protein, respectively, the SEDs are included in Table 3. Apart from fat, the evidence of significant improvement using data from the selected bulls is not as strong as that indicated by the population statistics however they are based on small numbers.

Table 3. Number of bulls intersecting the top 20 for genomic BV based on the two reference populations and difference in average deregressed daughter-proven BV

	Fat (kg)		Protein (kg)		Milk (litres)	
Breed	Bulls in common	difference	Bulls in common	difference	Bulls in common	difference
HF	12	6.1±1.6	9	0.4±1.3	9	20±45
JE	16	-1.6±1.1	13	1.0±1.0	15	30±30
FJ	13	2.6±1.5	14	-0.9±1.0	15	-60±30

CONCLUSIONS

There is some evidence to indicate that increasing the size of reference population through inclusion of cow data may lead to an improvement in the accuracy of genomic evaluation. The feasibility of including cow genotypes in a single-step method of evaluation (Aguilar *et. al.*, 2010), which combines information from genotyped and non-genotyped animals, is currently being investigated to confirm results of this study and check validation over a sequence of years. This will provide computational challenges in terms of inversion of the genomic relationship matrix which may become infeasible in the future as the number of genotyped animals increases. Reparameterisation of the MME in terms of marker effects instead of directly as BVs may be a better computational strategy.

REFERENCES

Aguilar I, Misztal I., Johnson D.L., Legarra A., Tsuruta S. and Lawlor T.J. (2010) *J. Dairy Sci*. **93**:743.

Browning B.L. and Browning S.R. (2009) *Amer. J. Hum. Genet*. **84**:210.

Daetwyler H.D., Villanueva B, and Woolliams J.A. (2008) *PLoS ONE* **3**:e3395.

David X., de Vries A., Feddersen E. and Borchersen S. (2010) *Interbull Bull*. **41**:77.

Garrick D.J., Taylor J. and Fernando R.L. (2009) *Genet. Sel. Evol*. **41:**55.

Goddard M.E. and Hayes B.J. (2009) *Nature Reviews Genetics* **10**:381.

Habier D., Fernando R.L., Kizilkaya K. and Garrick D.J. (2011) *BMC Bioinformatics* **12**:186.

Mantysaari E., Liu Z. and VanRaden P. (2010) *Interbull Bull*. **41**:17.

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

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:33-36

IMPROVING THE RELIABILITY OF FERTILITY BREEDING VALUES IN AUSTRALIAN DAIRY CATTLE

J.E. Pryce[1,2], M. Haile-Mariam[1,2], P. Bowman[1,2], T. Nguyen[1,2] K. Konstantinov[1,2,3], G.J. Nieuwhof[1,2,3] and B.J. Hayes[1,2]

[1]DEPI, 5 Ring Road, La Trobe University, Bundoora, VIC, 3083, Australia
[2]Dairy Futures CRC, 5 Ring Road, Bundoora, VIC, 3083, Australia
[3]ADHIS Pty Ltd, 22 William St, Melbourne, VIC, 3000, Australia
jennie.pryce@depi.vic.gov.au

SUMMARY

Over several decades, a decline in fertility in dairy cattle has been observed around the world. Breeding values, known in Australia as ABVs were implemented for fertility in 2003 using a single trait model for calving interval and since then the genetic trend in fertility ABVs appears to have stabilised. In April 2013 a new fertility ABV calculated using a multi-trait model was introduced that includes the following predictors of 6 week in-calf rate: calving interval, lactation length, days to first service, non-return rate and pregnancy rate. The new multi-trait fertility model has increased the reliability of fertility ABVs (compared to the single-trait model it replaced) for bulls born since 2000 by 6.5% and 7.6% for Holsteins and Jerseys respectively. The limitation to realising the full potential of this model is the capture of data. For cows that have calving interval records, 93%, 28%, 21% and 24% have records used for lactation length, days to first mating, pregnancy result and non-return to first service used in ABV calculations (for records collected between 2008 and 2010). The Dairy Futures CRC in conjunction with ADHIS have recently embarked on a co-ordinated effort to capture many more mating and pregnancy records that are electronically recorded on-farm but currently do not contribute to fertility ABVs. We have also set up a genomic reference population of females through identifying herds that have very well recorded data. The introduction of genomic data has been shown to increase the reliability of fertility ABVs of first proof bulls by a further 3%, having a genomic reference population will aid in sustaining the reliability of fertility ABVs. The collection of extra phenotypic data in addition to genomics is expected to increase the response to selection in fertility in the Australian dairy herd.

INTRODUCTION

Fertility in dairy cattle has been declining. Over the last decade the genetic trend for protein yield has been favourable (+1 kg/year) and calving interval has been unfavourable (+0.5 days/year) (Haile-Mariam and Pryce 2012). Selection for fertility can help to stop the downward genetic trend.

To be fertile, a cow needs timely return to cyclicity, display of oestrus, conception and staying in-calf. Fertility is therefore a complex trait that benefits from multi-trait prediction (Haile-Mariam and Pryce 2012). There are several ways in which the reliability of fertility breeding values (known in Australia as ABVs) can be increased. In Australia, the strategies being used to increase the reliability of fertility ABVs include: 1) use a multi-trait model to estimate fertility breeding values that encapsulates as much of the underlying genetic variation as possible; 2) capturing phenotypic data that can be used to predict ABVs; 3) using genomics to increase the reliability; 4) setting up genomic reference populations of cows with excellent fertility records.

The new multi-trait fertility model was launched in April 2013 and includes the following predictors of 6 week in-calf rate: calving interval, lactation length, days to first service, non-return rate and pregnancy rate. The previous model was based on calving interval (CI). However, calving

interval suffers from censoring, because cows with the poorest fertility do not re-calve. A multi-trait prediction is expected to result in higher reliabilities, which is especially valuable for young bulls because first proofs are generally based on low numbers of daughters and are therefore more susceptible to biases introduced by censoring. Furthermore, using additional data should lead to better fertility breeding values that are available earlier. Until recently a limitation to extending the model used to calculate fertility ABVs from single-trait to multi-trait was data availability. The limitation to realising the full potential of this model is still the capture of data, which can be increased by actively seeking many more mating and pregnancy records that are electronically recorded on-farm but currently do not contribute to fertility ABVs. Furthermore, the very best of these herds could contribute further still through genotyping and becoming part of a "female" genomic reference population.

The main objectives of this research were: 1) to compare ABVs from the new multi-trait model to those from the old fertility ABV model; 2) benchmark the amount of fertility data currently used for genetic evaluations, so that the impact of a data collection project can be accurately quantified; 3) describe the process to identify cows for a genomic reference population.

MATERIALS AND METHODS

Multi-trait fertility model. The Australian Dairy Herd Improvement Scheme (ADHIS) calculated fertility ABVs for the "new" multi-trait model and the previous "old" model it replaced. The increase in the number of bulls with a publishable fertility ABV was calculated based on the April 2013 ABV run using only Australian data and conventional breeding values for bulls with first daughter born from 2000 onwards. The correlation between ABVs based on the new multi-trait and the current two-trait model was calculated for these animals.

Data capture. The aim of the data capture project is to increase the reliability of fertility ABVs through actively sourcing many more mating and pregnancy records that are electronically recorded on-farm, but currently do not contribute to fertility ABVs. To be able to assess the impact of this effort, the first step was to benchmark the amount of fertility data captured (required for the new fertility multi-trait model). Data were extracted in August 2012 from the ADHIS database for cows that calved between 2008 and 2011. Data editing rules specific to fertility data (Haile-Mariam and Pryce 2012) were applied to the data in addition to standard ADHIS rules; for example animals without sires, birth-dates or calving dates were excluded from the data extract. Data was extracted again in April 2013 and the increase in data quantified.

Genomic nucleus population. Cows with superior fertility data i.e. high proportions of calving dates, mating dates and pregnancy testing were identified using a scoring system where each cow was awarded points for data that qualified for ADHIS evaluations for an index on yield, fertility, workability, calving ease, cell count, conformation and survival. Extra points were awarded for repeated records. This strategy has already been used to identify cows with valuable data to include in the Australian genomic reference population and is also being applied to recruit the best 100 herds that wish to be part of the experiment (subject to geographical and breed spread) to become part of a genomic reference nucleus to increase the reliability of fertility breeding values further still.

RESULTS AND DISCUSSION

Multi-trait fertility model. The reliability of fertility ABVs for bulls born since 2000 has increased by 6.5%, 8.7% and 7.6% for Holsteins, Red breeds and Jerseys respectively (Table 1). This improvement is mainly because lactation length data is available for cows that did not have calving interval. The advantage of including additional fertility traits, such as pregnancy and calving to first service is principally to improve the timing of fertility ABVs for selection decisions i.e. rather than waiting for 2 consecutive calving dates. Another advantage of using a multi-trait

model is that the standard deviation of bull ABVs increases by about 13% overall (and 6% within the Holstein breed for bulls born since 2000). The threshold for publication of fertility ABVs set by ADHIS is 55% reliability and daughters in at least 10 herds. There are approximately double the number of domestic Holstein and Jersey bulls with publishable fertility ABVs compared to the old model (Table 2). The new fertility ABV has also passed Interbull genetic trend evaluation tests; this means that for the first time in Australia many foreign bulls now have publishable fertility ABVs. The total number of bulls with publishable fertility ABVs has increased 17 and 5 fold for Holsteins and Jerseys respectively (Table 2).

Table 1. Reliability of bulls born since 2000 for the old and new fertility ABV

Breed	Number of bulls	Reliability old ABV	Reliability new ABV	Difference	Correlation
Holstein	2,421	61.9	68.3	6.5	0.87
Red breeds	29	62.1	70.8	8.7	0.81
Jersey	498	62.4	70.0	7.6	0.86

Table 2. Number of Holstein and Jersey bulls with publishable (>55% reliability in at least 10 herds domestically and >55% for Interbull proofs)

	Holstein		Jersey	
	Old ABV	New ABV	Old ABV	New ABV
Domestic	3,711	7,038	746	1,350
Interbull	0	55,362	0	2,466
Total	3,711	62,400	746	3,816

Data capture. In Australia, data is transferred from farm to local Data Processing Centres and then to ADHIS. This works very well for fertility data stored in some software used on farms, but not for others. Tests of manual transfer of data have been successful and automation of this process is currently underway for one software provider. We anticipate this will have a positive impact on data available for ADHIS fertility ABV calculation in the future.

Currently, among cows that had calving interval records in 2008, 2009 or 2010 the average proportions with lactation length, days to first mating, pregnancy result and non-return to first service data used in ADHIS ABV calculations at August 2012 were 93%, 28%, 21% and 24% respectively. By March 2013, for exactly the same period of time (2008-2010 inclusive) the number of cows with data used by ADHIS for fertility ABV calculations has increased from 1,171,287 to 1,384,086, which is an 18% increase. Veterinary clinics are also part of the work being undertaken, veterinarians commonly use DairyData software (written and supported by Warrnambool Veterinary Clinic), and this is proving to be a valuable source of pregnancy test data. One of the challenges to maximise the benefit of data from veterinary clinics in particular is establishing ways in which data can be obtained from non-herd-testing herds.

Genomic nucleus population. The Australian genomic reference population for Holsteins (April, 2013) consists of 3,449 Holstein bulls (with Australian daughters) and 8,691 cows selected based on the quality and quantity of their phenotypes. The genomic reference population for Jerseys currently consists of 946 Jersey bulls (with Australian daughters) and 3,996 females. The contribution of genomic data has been shown to increase the reliability of fertility ABVs of first proof bulls by 5-10% for the old ABV and around 3% for the new ABV (Table 3). National genomic evaluations commenced in 2011.

Table 3. Reliability of conventional and genomic breeding values for 1791 Holstein and 361 Jersey bulls

Breed	Old ABV			New ABV		
	N	Conventional	Genomic	N	Conventional	Genomic
Holstein	1,048	54.3%	65.0%	1,791	70.0%	73.0%
Jersey	208	60.8%	64.6%	361	68.8%	72.3%

As genomic selection replaces progeny testing, the reliability of genomic breeding values is expected to decrease because the distance between the reference and predicted population increases (Lillehammer *et al.* 2010). Strategies to reduce the reduction in reliability include exchanging genotypes between countries, using denser or more informative SNP data and genotyping females to become part of the reference population. The research we are undertaking will attempt to use all 3 strategies, including setting up an industry-resource genomic reference population.

Future. Future research activities to increase the reliability of fertility ABVs include: 1) using sequence and genomic data to increase reliabilities through improved capture of genetic variants responsible for variation in fertility; 2) improving our understanding of non-additive and epistatic genomic effects; 3) testing other predictors of fertility to improve the ABV model: heifer fertility, body condition score, protein percentage and other novel measures; 4) exploring the impact of management on fertility ABVs e.g. synchrony and inductions.

CONCLUSION

A new multi-trait fertility model was introduced in April 2013 that has resulted in increased reliabilities of fertility ABVs and consequently the number of bulls with publishable fertility ABVs. To increase the reliability of fertility ABVs further, extra phenotypic data (especially on mating and pregnancy tests) is actively being sought. In conjunction with our other activities in genomics and more recently sequence data, it is expected that reliabilities of fertility ABVs will continue to improve, giving farmers better choice of bulls and increasing the response to selection of this important trait.

ACKNOWLEDGMENTS

Dairy Futures CRC and ADHIS are gratefully acknowledged for funding this research.

REFERENCES

Haile-Mariam M. and Pryce J.E. (2012) *Proc. 5th Australasian Dairy Sci. Symp.* 298.
Lillehammer M., Meuwissen T.H.E.and Sonesson A.K. (2011) *J. Dairy Sci.* **94**: 493.

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:37-40

NEW APPROACHES TO GENETIC ANALYSIS OF FERTILITY TRAITS IN NEW ZEALAND DAIRY CATTLE

F.E. Bowley[1,2], P.R. Amer[1] and S. Meier[3]

[1]AbacusBio Limited, PO Box 5585, Dunedin, NZ, 9058
[2]Department of Environment and Primary Industries, 5 Ring Road, Bundoora, VIC, 3083
[3]DairyNZ Limited, Private Bag 3221, Hamilton, NZ, 3240

SUMMARY

There is potential for the current New Zealand fertility breeding value (BV) to be improved using additional information and traits. Data from 169 herds were analysed to determine the benefits of utilising alternative phenotypic measures in the calculation of the fertility BV. The heritability of calving season day (CSD; the number of days from the planned start of calving to the actual calving date) and the percentage of cows calving within 42 days of the planned start of calving (CR42) increased modestly (from 0.0206±0.0027 to 0.0213±0.0029 and 0.0087±0.0015 to 0.0092±0.0017, respectively) after accounting for the use of controlled internal drug release (CIDR) treatments and induced calvings. Incidence of either CIDR use or calving induction as a single binomial trait had a heritability of 0.0223±0.0020. The use of pregnancy diagnosis data allowed fertility information that would otherwise be discarded to be included in analyses; when used to assign a prolonged CSD and a value of 0 for CR42 to animals that failed to calve, it increased the heritabilities of both of these traits (to 0.0278 and 0.0114, respectively). As CSD was found to be more than twice as heritable as its binary counterpart, it shows potential to replace CR42 as the calving trait used in the fertility BV. Post-partum anoestrous interval (PPAI), derived using incomplete pre-mating oestrous recording in some herds, had a heritability of 0.0813±0.0110 and hence has potential as a trait to be included in genetic improvement programs, but would require more rigorous recording of oestrous during the pre-mating period to be an effective trait. Due to the increasing economic importance of fertility traits, and low heritabilities requiring large numbers of recorded daughters to get accurate BV predictions on sires, data recorded on-farm will become increasingly important in the genetic improvement of fertility. It is recommended that a system of identifying and incentivising herds with robust data-recording systems be designed and implemented to ensure ongoing collection of comprehensive and accurate data.

INTRODUCTION

The reproductive performance of dairy cows in New Zealand is superior to that in many other countries (Harris *et al.* 2002; Griffiths *et al.* 2007). However, fertility has been steadily declining phenotypically over the past 20 years; for example, Harris *et al.* (2006) reported a 10% decline in the proportion of cows re-calving within 42 days of the subsequent calving period between 1990 and 2004. While the estimated heritabilities of fertility traits are small (often less than 0.05), large additive genetic variation exists, meaning that improvement through genetic selection is possible (Harris and Montgomerie 2001; Harris *et al.* 2002). Fertility was first added as a breeding value (BV) to the New Zealand economic selection index, Breeding Worth (BW), in 2001 (Harris and Montgomerie 2001), and is currently largely based on the percentage of calvings within 42 days of the planned start of calving (CR42; Harris *et al.* 2006). BVs may be improved by re-defining current traits and through the inclusion of new traits. Therefore, potential for further improvement may exist through the use of alternative phenotypic measures in the calculation of the fertility BV.

The purpose of this study was to assess methods of data filtering and modification to improve fertility trait heritabilities, and investigate novel ways of measuring fertility to improve the current New Zealand fertility BV.

MATERIALS AND METHODS

Data and fertility traits analysed. Up to 259,651 records (depending on the trait; Table 1) from 139,134 animals (cows and heifers) in 169 herds participating in a fertility monitoring project across New Zealand were available for analysis (Brownlie *et al.* 2011). SAS (version 9.2) was used to handle, filter and manipulate the data.

Table 1. Fertility traits analysed, their acronyms and descriptions

Trait	Acronym	Description
Post-partum anoestrous interval	PPAI	Days from previous parturition to first observed oestrous (or first mating if oestrous not observed)
Percentage mated 21 days	PM21	1 if first mating occurred within 21 days of start of mating date, 0 if first mating occurred after 21 days, and missing if not mated
Mating season day	MSD	Days from start of mating date to first mating (similar to PM21, but left as a continuous trait instead of being scored 0/1)
Calving rate 42 days	CR42	1 if calved within 42 days of planned start of calving date, 0 if calved after 42 days, and missing if not calved
Calving season day	CSD	Days from planned start of calving date to calving (similar to CR42, but left as a continuous trait instead of being scored 0/1)

Statistical model. Heritabilities were estimated using a univariate animal model in ASReml (version 3), consisting of mean, covariates of age and breed percentage, fixed effects of herd, contemporary group (herd, year and if the animal was a cow or heifer at the time) and interaction between herd and year, and random effects of animal and permanent environmental effect.

Data manipulation and modification.

Data filtering. Various filters were applied to the data to minimise distortion of results. For example, animal records were removed from the dataset used in analyses if the animal was greater than six years old at time of mating, its sire had less than four daughters, there were less than 50 animals in its contemporary group, or it was mated after February or before August (i.e. outside of the normal window for seasonal-calving herds).

Adjustment for fertility treatments. The use of fertility treatments, namely controlled internal drug release (CIDR) and calving induction, results in artificially-altered fertility records. Hence, in order to assess their impact on fertility trait heritabilities, any affected records were set to missing. An additional binomial trait representing the incidence of CIDR use or induction (CIDRIND) was also calculated and analysed. For each record, an animal was scored as 1 if it was treated with a CIDR or induced, 0 if there were records of CIDR use or induction for other animals in that herd-year, and missing if there was no record of CIDR use or induction in that herd-year.

Pregnancy diagnosis data. Some pregnancy diagnosis data, including whether the animal was confirmed pregnant or not, was also available, and was used to test the effect of including knowledge of failed pregnancies on the heritabilities of CR42 and calving season day (CSD). Animals that were diagnosed as not pregnant were given a CR42 record of 0 and a CSD record 10 days later than the last calving day for the year in that herd.

RESULTS AND DISCUSSION

Adjustment for fertility treatments. Accounting for the use of CIDR treatments and induced calvings by setting affected records to missing altered fertility trait heritabilities (Table 2). While the heritabilities of post-partum anoestrous interval (PPAI) and percentage mated 21 days (PM21)

did drop slightly and remained unchanged for mating season day (MSD), this filtering of records modified by fertility treatments had a positive effect on the heritabilities of calving traits. In addition, as the use of such interventions creates fertility records not representative of the true fertility of the animal, removing such records from analyses is the logical option. Attempting to correct records for the effects of these interventions by fitting them as fixed effects in the statistical model did not increase heritabilities.

Table 2. Number of records, adjustments for CIDR use and induction and resulting heritabilities (with standard errors in parentheses) for each of the fertility traits

Trait	Number of records	Heritability (unadjusted)	Set missing if CIDR used?	Set missing if induced?	Heritability (adjusted)
PPAI	31,252	0.0814 (0.0104)	✓	✓	0.0813 (0.0110)
PM21	259,615	0.0335 (0.0035)	✓		0.0352 (0.0037)
MSD	258,854	0.0239 (0.0029)	✓		0.0239 (0.0030)
CR42	218,098	0.0087 (0.0015)	✓	✓	0.0092 (0.0017)
CSD	217,053	0.0206 (0.0027)	✓	✓	0.0213 (0.0029)

CIDRIND was found to have a heritability of 0.0223±0.0020, and it is recommended that this trait be included in genetic evaluations for fertility, particularly since setting other affected records to missing would mean individuals sub-optimal for fertility would not be adequately penalised unless it is incorporated.

Use of pregnancy diagnosis information. In the absence of pregnancy diagnosis data, when an animal fails to become pregnant or sustain pregnancy, CR42 and CSD are recorded as missing. This is effectively a loss of fertility information, because the associated failure to calve is not accounted for when analysing these calving traits, as missing records are excluded from analyses. Since some pregnancy diagnosis information was available, this was used to test the effect of including knowledge of failed pregnancies on the heritabilities of CR42 and CSD. The result was an increase in the heritability of CR42 from 0.0092 to 0.0114 and CSD from 0.0213 to 0.0278. Hence, recording and use of pregnancy diagnosis data has been confirmed as important in the analysis of calving-based fertility traits.

Replacing CR42 with CSD. CR42 is the calving trait currently used in the fertility BV. However, after adjusting for CIDR use and calving induction, and utilising pregnancy diagnosis information, the heritability for CSD (0.0278) was more than twice that of its binary counterpart, CR42 (0.0114), which suggests that genetic progress for fertility may be made at an increased rate if selection was based on CSD instead. Use of CSD allows differentiation between animals calving in the first versus the second 21-day period of the calving season, allowing more refined selection to ensure a compact calving pattern is maintained. It also alleviates problems associated with not being able to determine the planned start of calving date very accurately because of variation in gestation length. However, further research using a larger dataset is required to verify the benefits of replacing CR42 with CSD as the calving trait in the fertility BV.

Potential for use of PPAI. Of the relatively low number of 31,252 "first oestrous" records, only 1,663 (5%) were actually recorded as an observed oestrous; the remaining were proxies based on the first recorded mating for the season in herd-year groups where some oestrous recording was performed. Despite this, PPAI was the most heritable (0.0813±0.0110, after adjusting for CIDR use and induction) of all the traits analysed, suggesting that it may be a valuable indicator of fertility. Currently, one of the reasons for poor PPAI measurements is that oestrous detection only begins 3 to 4 weeks prior to the planned start of mating date at the earliest, and many animals cycle before this time. An analysis of the data revealed that, based on an average PPAI of 40 days,

58% of first oestrous detected would have occurred prior to this and hence not been recorded. Hence, PPAI shows substantial potential as a trait to be included in selection programs for fertility, but it would be necessary to highlight to farmers the benefits of early oestrous recording, both for genetic improvement as well as general herd management.

Collection of farmer-recorded fertility data. Farmer-recorded data are going to become increasingly important in the genetic evaluation of future sires. This importance will be driven by the rapid development of genomic selection methodologies and the associated reduction in the number of progeny-tested sires, as well as the increasing availability of on-farm milk recording and analysis systems reducing the need for herd testing. Currently, data recording on commercial farms is highly fragmented and is not always stringent, particularly in the case of fertility. It is recommended, therefore, that a system of identifying and incentivising selected herds with robust data-recording systems be designed and implemented to ensure ongoing collection of comprehensive and accurate data for use in genetic improvement programs. Ideally, these herds would use a high-proportion of semen from elite young sires identified using genomic information.

CONCLUSION

The results of this study show that there is potential for improvement in the New Zealand fertility BV through the inclusion of additional information. Adjusting for CIDR use and calving induction by setting affected records to missing is not only logical but increased the heritabilities of the calving traits CSD and CR42. Incidence of CIDR use or induction as a binomial trait was also reasonably heritable and should be included in fertility evaluations to ensure animals that are sub-optimal for fertility are adequately penalised. Pregnancy information further increased accuracy of selection for CSD and CR42. As CSD seems to be a more heritable trait than its binary counterpart, further study into replacing CR42 with CSD as the calving trait in the fertility BV is warranted. PPAI shows promise as a fertility trait with relatively high heritability, but would require recording of oestrous during the pre-mating period. In order to facilitate ongoing genetic improvement of fertility, herds with robust data-recording systems need to be identified and used as sources of comprehensive and accurate data to be used in sire genetic evaluations.

REFERENCES

Brownlie T.S., Weir A.M., Tarbotton I., Morton J.M., Heuer C. and McDougall S. (2011) *N. Z. Vet. J.* **59**: 28.

Griffiths L.M., Loeffler S.H., Socha M.T., Tomlinson D.J. and Johnson A.B. (2007) *Anim. Feed Sci. Technol.* **137**: 69.

Harris B.L. and Montgomerie W.A. (2001) *Proc. Interbull Annu. Meet.* **27**: 139.

Harris B.L., Pryce J.E. and Johnson D.L. (2002) "The economics of dairy cow fertility" Livestock Improvement Corporation, Hamilton, New Zealand.

Harris B.L., Pryce J.E., Xu Z.Z. and Montgomerie W.A. (2006) *Proc. N. Z. Soc. Anim. Prod.* **66**: 107.

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:41-44

A BAYES-A LIKE METHOD IN ASREML

Arthur R Gilmour

School of Mathematics and Applied Statistics, University of Wollongong, NSW, 2522
***Cargo Vale* Cargo NSW 2800**

SUMMARY

The paper describes a method implemented in **ASReml** for estimating genomic breeding values and marker effects distributed according to a *t* distribution from a large panel of SNP markers. The method is similar to the MCMC Bayes-A method. It estimates effects in the individual animal space and back-solves to obtain the marker effects.

INTRODUCTION

With the increasing availability of SNP panels for genotype characterization comes the challenge of how best, or at least effectively, to utilize them. Two emphases are common; first to predict breeding values, using the markers to define genetic relatedness more accurately than by using expected average relatedness as predicted from a pedigree, and second to identify loci (markers) of large effect hoping that the large effect is due to a nearby major gene (QTL).

Meuwissen *et al.* (2001) proposed several approaches including methods called GBLUP and Bayes-A. The basic marker model is $\boldsymbol{y} = \mathbf{X}\boldsymbol{b} + \mathbf{M}\boldsymbol{g} + \boldsymbol{e}$ where marker scores (**M**) are used to fit random marker allele effects ($\boldsymbol{g}$) with a common variance σ^2_g. This is equivalent to using the genomic relationship matrix we write as $\mathbf{G} = \mathbf{MDM}'$ used in place of the Numerator Relationship matrix in the animal model to produce GBLUP ($\boldsymbol{u}$) where **M** is the matrix of (centred) marker scores (0/1/2) and $\boldsymbol{D}$=diag(1/*s*), $s = \Sigma\ 2p_i(1\text{-}p_i)$ and p_i is the proportion of the minor allele for marker *i* (Stranden and Garrick, 2009). The animal model formulation is generally more tractable because the number of markers typically far exceeds the number of animals. The link is that $\boldsymbol{u} = \mathbf{M}\boldsymbol{g}$ and $\boldsymbol{g} = \mathbf{M}'\mathbf{G}^{-1}\boldsymbol{u}$. In the mixed model equations, **G** is scaled by a variance parameter which is related to the marker variance component: $\sigma^2_u = s\sigma^2_g$.

However, it is likely that markers are not equally informative, that they have diverse variances. The Bayes-A model assumes a scaled inverse Chi-square distribution for the individual marker variances implying a *t* distribution for the marker effects and uses Markov chain Monte Carlo techniques to estimate the marker effects. Sun *et al.* (2012) propose an EM method based on the GBLUP model but where **D**, initialized at diag(1/s), is updated each iteration using the estimated marker effects as $\mathbf{D} = \text{diag}(\boldsymbol{g}*\boldsymbol{g} + (k\text{-}2)\sigma^2_g)/(k+1)$ where k is the degrees of freedom of the Chi-square distribution and σ^2_g is the marker variance assumed known. The idea here is that if we assume a scaled inverse Chi-square distribution for the marker variances, we can estimate those variances using the estimated marker effects and the scale parameter. We then use the estimated marker variances as weights for the marker covariables and reform **G**. That is then used in the mixed model equations to re-estimate the marker effects. This paper describes an implementation of this method in **ASReml** (Gilmour 2013). The method is called Fast Bayes-A (FBA).

MATERIALS AND METHODS

Data Set The method is demonstrated on a simulated data set (Szydlowski and Paczyńska, 2010) comprising marker scores for 10031 markers on 3226 animals, 2326 of which have phenotype and all of which have 'true' breeding values. The data is supplied in a marker file and a phenotype file. The phenotypic variance for this data is 100.6.

Three Models

For **M**, $\sigma^2_u = s\sigma^2_g$, **g** and *k* as defined above, we define three forms of the genomic relationship matrix **G**=**MDM**' where **D** is a diagonal matrix of relative variances used to weight each marker:
GBLUP: **D** is diag(1/s), **G** formed once.
FBA-F: **D** is diag($\boldsymbol{g}*\boldsymbol{g}/\sigma^2_u + (k-2)/s)/(k+1)$, **G** reformed each iteration with σ^2_u fixed
FBA-G: **D** is diag($\boldsymbol{g}*\boldsymbol{g}/\sigma^2_u + k/s)/(k+1)$, reformed each iteration with σ^2_u updated.

The animal model is then fitted using the **G** matrix and marker effects (**g**) are predicted from the animal effects. For GBLUP and FBA-G, the genetic variance (σ^2_u) is estimated; for FBA-F it must be held fixed. FBA-F is the model described by Sun *et al.* (2012) but with **D** multiplied by the σ^2_u rather than applying σ^2_u as a scale factor for **G**. Further, Sun *et al.* (2012) used the σ^2_u estimated from the GBLUP (equal marker variances) model as the known prior variance.

The difference between FBA-F and FBA-G is that σ^2_u is estimated in the latter, and is scaled according to the RHS constant in the expression **D**. So, if *s* is set to one, the variance parameter is related to the marker variance σ^2_g, not the genetic variance; using *k-2* instead of *k* results in a value *k/(k-2)* larger; use of *k/s* results in a variance estimate comparable to the GBLUP value.

Meuwissen *et al.* (2001) used a value of *k* close to 4 which pulls the marker variances toward *(k-2)/(k+1)*=0.4 of the average value under GBLUP. The distribution is less skewed under the FBA-G model and so it does not follow the nominal inverse Chi-square distribution.

Models fitted. These three models were fitted to the simulated data and the FBA models evaluated with *k* at 4.2, 3.8 and 3.5. The FBA-F model was fitted assuming the variance ratio (Genetic/Residual) obtained from the GBLUP fit, although it could have easily been evaluated with σ^2_u =44.0. The FBA models identify a few markers of large effect and we examine the impact of fitting 4 of these as fixed covariates (putative QTL).

RESULTS

The primary results are summarized in Table 1. For the GBLUP model, the genetic variance (ratio) was estimated at 44.03 (0.808) corresponding to a marker variance component of 0.01177 and the Log Likelihood was -6077.5. The largest marker effect was -0.165 for marker 4480.

The number of markers having a large effect was strongly influenced by the value of *k*, with large consequent jumps in the Log Likelihood. However, further reducing *k* to 3.2 gave a poorer fit, especially for the FBA-F model (values not given). From these and other models fitted, we see a jump in Log likelihood for each marker of large effect identified: -6049, -6033, -6012, -6007 for 1, 2, 3 and 4 markers with large effect. Each large marker is effectively fitted as a fixed effect (having a relatively large individual effect variance).

The marker variances are less skewed under FBA-G than FBA-F and so fewer large markers are detected for a given value of *k*. Indeed a plot of marker effects with variance fixed and *k*=3.8 against marker effects estimated when the variance is estimated and *k=3.5* shows very close agreement except for the 2 largest effects which are 20% larger under the latter model (Figure 1).

The accuracy is the correlation between the BLUP values predicted for the 900 individuals without data and the 'true' breeding values of these individuals. It increases with increasing Log Likelihood.

Having identified markers of large effect, these can be fitted as separate covariates. Table 2 shows the Wald F statistics and effects of the 4 markers having largest effect; they explain 40% of the genetic variance. Markers 952 and 954 are neighbours and each is as effective as the other when fitted singly but they also complement each other.

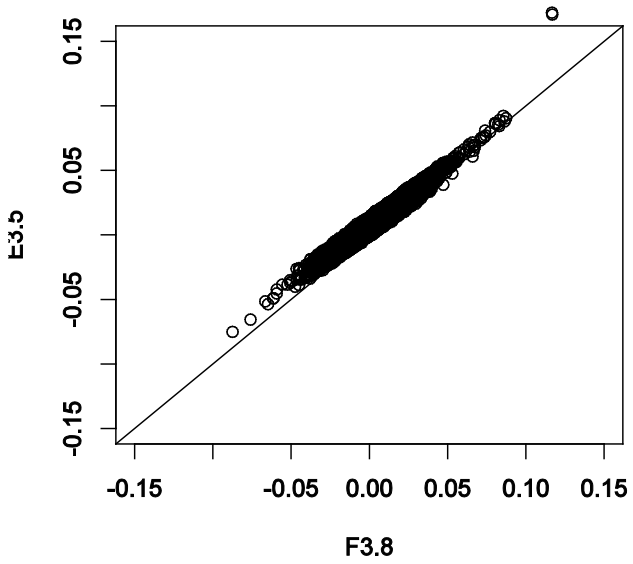


Figure 1. Comparison of marker effects estimated with k=3.8 and fixed known variance with values estimated with k=3.5 and genetic variance estimated (ignoring 4 markers of large effect).

Table 1. Comparison of model statistics from the GBLUP model and for the Fast Bayes-A like model for 3 levels of degrees of freedom and holding the variance ratio fixed, or estimating it.

Degrees of freedom (*k*)	LogL	Residual Variance	Genetic Variance	Accuracy	Large Markers
GBLUP	-6077.5	54.5	44.0	0.611	
FBA-F Genetic Variance ratio fixed at 0.808					
4.2	-6042.6	57.9	46.8	0.635	954/4480
3.8	-6008.1	52.5	42.4	0.656	952/954/4480/5488
3.5	-5995.9	52.0	42.0	0.659	145/952/954/ 2719/4480/5488
FBA-G Genetic Variance estimated					
4.2	-6050.3	53.7	49.7	0.636	4480
3.8	-6033.0	53.9	43.4	0.645	954/4480
3.5	-6004.3	53.8	44.6	0.655	952/954/4480/5488

Table 2. Wald F statistics and fixed effects for 4 markers of large effect in a GBLUP model where the genetic variance was estimated at 26.35 and residual variance at 54.17. The Incremental (Marginal) F reflects the variation explained as markers are added in order (after all others).

Source	Incremental F	Marginal F	Fixed Effect
snp(952)	55.84	12.80	2.24 ± 0.63
snp(954)	14.42	12.90	-2.19 ± 0.61
snp(4480)	60.69	61.51	3.50 ± 0.45
snp(5488)	46.57	46.57	-3.33 ± 0.49

DISCUSSION

ASReml has been widely used for fitting GBLUP models where users have supplied the **G** matrix. Now it can directly make a common form of the **G** matrix, and report marker effects as well as animal effects.

The Bayes-A like models give a better fit to the genetic relationship matrix than the GBLUP model, as indicated by the Log likelihoods, and identify markers of large effect. The number of large effects identified is related to the peakedness of the *t* distribution which is controlled by the degrees of freedom, *k*. There is currently no formal method to choose a value for *k* in this implementation. Sun *et al.* (2012) used 4.2 but 3.5-3.8 seems more appropriate here.

The GBLUP model runs much faster than the FBA model because the **G** matrix is only formed and inverted once whereas in the FBA model it must be formed and inverted each iteration, increasing the time in this example from 40s to 180s per iteration. Therefore, it will generally be more efficient to follow the path Sun took and estimate the genetic variance under the GBLUP model and then use that value as the fixed prior for the FBA-F model. Furthermore, the FBA runs typically required from 20 to 40 iterations for the marker effects to stabilize while the GBLUP run took about 8 to 10 iterations to estimate the variance parameters.

ASReml can fit identified markers as separate (fixed or random) effects. Including the 4 markers identified in the FBA-F model with *k=3.8* as fixed effects and estimating the remaining genetic variance under the GBLUP model produced an estimate 40% lower than obtained in the original GBLUP model.

The FBA implementation is restricted to a single marker matrix on a single trait but the **G** matrix formed can be saved for use in more complex models.

REFERENCES

Gilmour A.R. (2013) *Pers. Comm.* Technical Report, University of Wollongong.

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

Stranden I. and Garrick D.J. (2009) *J Dairy Sci* **92**: 2971.

Sun X., Qu L., Garrick D.J., Dekkers J.C.M., and Fernando R.L. (2012). *PLoS ONE* **7(11)**, e49157.

Szydlowski M., and Paczyńska P. (2010) http://www.biomedcentral.com/1753-6561/5/S3/S3.

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:45-48

A VERY SIMPLE MODEL FOR EXAMINING POTENTIAL IMPACTS OF VALUE CHAIN PARAMETERS ON DIRECTION OF SELECTION AND GENETIC CHANGE

R. G. Banks[1], P. R. Amer[2], and J.H.J. van der Werf[3]

[1]Animal Genetics and Breeding Unit, UNE, Armidale, NSW, 2351
[2]AbacusBio, Dunedin, New Zealand
[3] Animal Science, UNE, Armidale, NSW 2351

SUMMARY

The Australian beef (and meat sheep) industries face a significant strategic challenge around optimizing the joint improvement of qualities demanded by the consumer and production attributes affecting on- and off-farm enterprise profitability. A very simple example based on beef cattle value chains is used to show that this joint optimization may not be trivial. Possible responses to this challenge are discussed, with an emphasis on the need for coordination across the sectors involve in the value chain.

BACKGROUND

Beef (and to a lesser extent meat sheep) breeding and production in Australia is characterized by:

- diversity of production environments
- diversity of market demands, especially in relation to "quality" traits
- typically 5 links in the value chain from breeder to consumer (breeder, producer, feedlotter, processor, retailer, consumer)
- very low levels of vertical integration through ownership
- limited and/or diverse flow of price signals across the links in the chain

These circumstances generate significant challenges for development and implementation of breeding objectives and corresponding selection indexes. Given these characteristics the question is apparent should whole of chain approaches be used, and if so, how will they be accepted by breeders, and how will the breeding sector respond in terms of investment in the recording necessary to underpin selection (including genomic selection)?

To the extent that these issues are forms of market failure, they have contributed to the establishment of collective levies for a range of industry investments, including R&D. There is potential for such funds to assist with funding the recording of Hard-to-Measure traits in reference populations, but no systematic model has been developed as yet for sourcing or allocating funds for this purpose.

This paper outlines how selection index methods coupled with very simple economic models, could contribute to the development of such systematic models.

A SUPPLY CHAIN MODEL WITH A VERY SIMPLE BREEDING OBJECTIVE

A beef value chain with 5 sectors is modelled (breeder, producer, feedlotter, processor, retailer and consumer), with a breeding objective that includes 4 traits (reproduction rate, sale weight,

carcase yield and eating quality). The objective is very loosely based on those developed for Angus cattle in Australia, but very much simplified here.

The economic modelling consists of 2 steps:

- estimate the expression of genetic change in each of the 4 traits in each of the 5 non-breeder sectors,
- estimate the extent to which a price signal is passed for each trait, from each sector to the one above it

Together these steps allow both the total value of genetic change in the 4 traits across the chain, and how much of that total value is passed back to the breeder to be modelled. The basic data and results are outlined in Table 1.

Table 1: Basic parameters of the economic and selection index model

		Repro'n Rate %	**Sale Weight kg**	**Carcase Yield %**	**Eating Quality Marble Score**
Whole Chain Objective Index economic weights ($)		$7.50	$1.75	$10.50	$300.00
Predicted genetic gain using Whole Chain index (trait units)		0.31	3.13	-0.13	0.36
Predicted genetic gain using Whole Chain index ($)		$2.30	$5.49	-$1.36	$106.89
Sector	**Parameter**				
Breeder	total benefit received	$0.23	$0.50	$0.00	$0.00
Producer	predicted expression of genetic gain	$2.30	$5.02	$0.00	$0.00
	% transmission to next sector up	10%	10%	1%	10%
	total benefit received	$2.30	$5.02	-$0.06	$0.03
Feedlot	predicted expression of genetic gain		$0.16	-$0.26	
	% transmission to next sector up	0%	0%	15%	10%
	total benefit received	$0.00	$0.16	-$0.39	$0.32
Processor	predicted expression of genetic gain		$0.31	-$0.91	
	% transmission to next sector up	0%	0%	15%	10%
	total benefit received	$0.00	$0.31	-$0.91	$3.21
Retailer	predicted expression of genetic gain				$10.69
	% transmission to next sector up	0%	0%	0%	10%
	total benefit received	$0.00	$0.00	$0.00	$32.07
Consumer	predicted expression of genetic gain				$71.26
	% transmission to next sector up	0%	0%	0%	30%
	total benefit received	$0.00	$0.00	$0.00	$71.26

The transmission rates used here are estimates based on observation of industry.

The 2nd row of the table shows the economic values for the whole chain. The row labeled "Breeder – total benefit received", shows the economic values that the breeding sector would apply based on price signals passed back to that sector. If these economic values are applied to index calculations using the same genetic parameters, the trait and $ value outcomes are as shown in Table 2 (over page).

Table 2: Key results – economic values, trait and $ responses.

Trait	Whole Chain			Breeder		
	Economic Value	Trait Response	$ Response	Economic Value	Trait Response	$ Response using Chain EVs
Repro'n Rate	$7.50	0.31	$2.30	$0.23	0.15	$1.14
Sale Wt	$1.75	3.13	$5.49	$0.50	13.97	$24.45
Yield	$10.50	-0.13	-$1.36	$0.00	0.08	$0.86
EQ	$300.00	0.36	$106.89	$0.00	0.00	$0.00
Total			$113.32			$26.44

Examination of Tables 1 and 2 shows very clearly that:

- on a whole chain basis, eating quality is a very important trait, contributing 94% of the total value of genetic change
- selection on the basis of price signals received by the breeding sector generates genetic change that is radically different in both direction and value from that based on whole chain value. In this case, almost all genetic change is in sale weight, there is little or no change in yield or eating quality, the value of genetic change for the whole chain is less than one quarter of that in the whole chain case, and no benefit is delivered to the consumer sector. However, benefits to the breeding and production sectors, through improved reproduction rate and sale weight, and the returns from them passed back to breeders from producers are very much greater than in the whole chain case.
- The differences in economic outcomes are substantial; $26.44 per cow joined per generation (approximately $5 per cow joined per year) compared with $113 (approximately $21 per cow joined per year). On a whole of industry basis, this difference would total $80m pa.

This very simple model case, using price signal parameters drawn from observation of industry, suggests a clear conflict of interests between the on- and off-farm sectors. How might industry respond to this situation?

POTENTIAL RESPONSES

The first response examined is developing and implementing selection indexes. To date in the Australian beef industry, the whole of chain approach has been adopted (Barwick, pers. comm.) If the genetic and price parameters modelled here are broadly relevant to the beef industry, then this approach has been good for the whole chain, at least to the extent that selection has been guided by the "whole of chain" indexes. However, the breeding and production sectors may have experienced lower direct benefits than had indexes based on on-farm returns been applied.

For this approach to be sustained for the long-term, breeders must either be altruistic, or estimate the opportunity cost to the industry, and hence to them, of ignoring improvement in quality traits. Depending on the elasticity of demand for meat of different quality, that opportunity cost would ultimately be in declining real price for meat.

This raises the question of whose interests should breeders, or industry as a whole, seek to maximize when investing in genetic improvement. The observation that a considerable portion of benefits from genetic improvement accrues to consumers is often interpreted as justifying government support for industry, usually through funding for R&D. However, it can be argued that it is the responsibility, or at least sensible aim, of an industry to maximize consumer benefits

in its own interest. Under this view, breeders should aim to select to maximize benefits for each of their value chain partners. Resolving, or at least balancing these two perspectives, is an important strategic question for industry consideration.

An important aspect of this industry perspective is the importance of obtaining reliable estimates of consumers' preferences for quality levels, and of the transmission parameters between sectors. At present, the core research for estimating these is done by a very small number of individuals in the breeding sector or working with it, with very limited access to industry data. A whole of industry contribution to obtaining and sharing this data would seem sensible.

A second area of response available to industry lies in taking steps to improve the flow of price signals. The Australian beef industry through Meat and Livestock Australia is investing in infrastructure which could assist (Meat Standards Australia, National Livestock Identification Scheme, Livestock DataLink), but integration across value chains is largely an opportunity for individual chains, rather than a whole of industry imperative.

A third area of potential response lies in co-investment in reference population data. The beef and sheep industries have started down this path via the Beef and Sheep Information Nucleus programs, currently with different funding models but both drawing substantially on Commonwealth funds. The example outlined here suggests a mechanism whereby the co-investment could be balanced with interests:

- traits and trait economic values determined for a breed-production system-market combination
- sectoral expression of genetic change in each trait calculated using index methods
- transmission parameters for each trait x sector estimated
- recording costs for each trait for an appropriate reference population estimated
- sectoral contributions to those costs on a trait basis calculated according to the relative benefits captured by each sector. The actual contributions could well be adjusted according to realized genetic change in each trait, thus increasing the incentive not just for recording but for genetic improvement as well.

This draft approach would require coordinated examination of the potential for change involving all sectors, coupled with the application of selection index calculations. The latter is straightforward; the former is challenging, but failure to resolve this challenge guarantees massive opportunity costs.

It is reasonable to ask whether the problems highlighted here would disappear in the event of industry re-structuring, in broad terms involving some form of vertical integration. Were this is improve the clarity and flow of price signals, then the problems would surely be reduced, at least for those value chains so re-structured. This raises the question of whether and how widespread such restructuring might be: to date, despite some obvious economic attractions, vertical integration remains very limited in both industries.

CONCLUSIONS

In a diverse, heterogeneous and multi-sectoral industry such as the Australian beef industry, optimizing genetic improvement is a complex case of the coordination problem. Technically, the issues and how to solve the relevant calculations are straightforward, but implementation requires inputs and actions from a range of interest groups.

Currently, the approach to solving this coordination problem is a mixture of "leaving it to the market" coupled with dependence on a combination of foresight and altruism on the part of the breeding sector. The example used here, while very simple, shows that the costs of this strategy can be both significant and unevenly distributed amongst sectors. For sustained viability of the industry, a better approach to balancing the interests of sectors needs to be developed and applied.

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:49-52

BUSINESS METRICS FOR SHEEP IMPROVEMENT LIMITED (SIL) RAM BREEDERS

T.J. Byrne[1], N. B. Jopson[1] and M.J. Young[2]

[1]AbacusBio Limited, P O Box 5585, Dunedin 9058, New Zealand
[2]SIL, Beef + Lamb New Zealand Limited, P O Box 39-085, Christchurch 8545, New Zealand

SUMMARY

While there is not a clear distinction between the information used in potential ram breeding business metrics and flock genetic improvement, the objective of this research was to look for opportunities to report statistics back to breeders each year that would help guide their ram breeding business. A consultation process showed that breeders are interested in generating more progeny of higher genetic merit and reducing the number of progeny culled because they do not meet criteria for sale to ram buyers. Breeders put less value on metrics related to how much clients pay for rams, or how many years clients had been buying rams.

This paper describes the data requirements for, and calculation of, business metrics that utilise information held in the SIL database together with sale data ram breeders can collect. These metrics include measures of the 'impact' a given sire has on the ewe flock (through selection and persistence of his daughters), the proportion of a ram's sons sold, average price of a ram's sons, and average price per unit of estimated breeding value, sub-index or overall index. Examples are provided for those metrics where information is already available on the SIL database. Calculation of the reported metrics is straightforward. However, more complete data are required in existing SIL record fields in addition to the need to include new data fields in SIL, in order to produce robust and informative metrics for ram breeding businesses.

INTRODUCTION

Sheep Improvement Limited (SIL) has a wealth of data and information on individual animal and flock performance (Newman *et al.,* 2000). Many ram breeders also collect data about client expectations, sale prices, general signals coming from the commercial market, and financial performance of their business. Leading ram breeders strive to increase the profitability of their business. In order to do so, information beyond that associated with flock genetic improvement is required.

There is an opportunity for wider use of the SIL database, such that some data held by ram breeders could be analysed along with data in the SIL database to generate metrics that better characterize the ram breeding business. The Irish Cattle Breeding Federation (ICBF) and Sheep Ireland database is used extensively in the provision of information beyond that associated with genetic improvement (Wickham, 2012). This paper describes the data requirements for, and calculation of, business metrics that utilise information held in the SIL database together with information ram breeders can collect.

CONSULTATION

This work was initiated by surveying a small number of large-scale breeders since they are more likely to yield robust metrics and to see the benefits of these to their ram breeding business. Dual Purpose (DP or ewe breed) flocks were chosen because of the importance of maternal traits in defining genetic merit and because it was considered that Terminal Sire (TS or meat breed) flock metrics would be a subset of those studied for DP flocks. Three flocks belonging to different breeders provided the information on which this report is based. The consultation process involved individual meetings with the breeders. From these meetings a list of potentially useful ram breeder business metrics was compiled based on suggestions from breeders and from discussions related to

metrics offered to the breeders at that time. Later, the breeders completed a survey which offered them paired alternatives from this list of different ram breeder business metrics to determine their preferences (Byrne *et al.*, 2012).

This consultation process showed that breeders are interested in generating more progeny of higher genetic merit and producing progeny that make it through commercial culling criteria. This essentially reflects a desire to reduce 'wastage', i.e. rams bred that cannot be sold to commercial farmers. Breeders generally maintain strong relationships with their clients, have a good knowledge of client requirements and value clients that buy lower priced rams as much as those buying higher priced rams. Hence there appears to be little business value in knowing how much different clients pay for rams.

METRIC CALCULATION

The priority metrics for breeders were; 1) measures of the 'impact' a given sire has on the ewe flock (through selection and persistence of his daughters), 2) the proportion of a ram's sons sold, 3) average price of a ram's sons, and 4) average price per unit of estimated breeding value (eBV), sub-index or overall index. Each of these is considered in more detail below.

The 'impact' a given sire has on the ewe flock. Impact can be calculated such that a) the number of daughters born to each sire (this represents the total opportunity the sire has to contribute to the flock through his daughters), b) the subsequent proportion of each sire's daughters that enter the flock, and c) the subsequent survival of daughters of each sire in the flock are all taken into account.

Data requirements for this metric include a count of the number of daughters born to a sire, the number of daughters born to a sire lambing in each cohort in the flock, and the total number of ewes lambing in each cohort in the flock. Progeny born to a sire (attribute = number of progeny) and daughters with a lambing record (attribute = number of daughters lambing) are already held in the SIL database. Number of daughters lambing in each cohort in the flock is also in the database through NLB records (NLB1, NLB2 etc.).

Cohort impact (*CDI*) can be reported as a deviation from that which would be expected, in terms of percentage contribution to a lambing cohort, based on the number of daughters born, calculated as:

$$CDI_{sc} = \left(\left(\frac{DS_{sc}}{TS_c} - \frac{DB_s}{TB}\right) \times 100\right)$$

where for sire s in lambing cohort NLB1 to NLBc, where c is the total number of lambing cohorts, *DS* is the number of daughters selected (i.e. lambing in the cohort), *DB* is the number of daughters born, *TS* is the total number of daughters selected and *TB* is the total number of daughters born.

Sire impact (*SDI*) can be reported as the mean of the cohort impact deviations across lambing cohorts weighted by the total number of daughters in each cohort, calculated as:

$$SDI_s = \frac{\sum_{i=1}^{c}(CDI_{si} \times TS_i)}{\sum_{i=1}^{c} TS_i}$$

The following example describes the impact of 3 sires. Each of the sires has had daughters lambing in 3 different lambing cohorts, NLB1, NLB2, and NLB3. Table 1 presents the number of daughters have born (*DB*) and the number of daughters selected (*DS*) for sires A, B, and C respectively. The total number of daughters selected (*TS*) for each lambing cohort and the total number of daughters born (*TB*) is also presented.

Table 1. The number of daughters born (*TB*) and selected (*TS*) for sires A, B, and C respectively, the total number of daughters selected (*TS*) for each lambing cohort, and the total number of daughters born (*TB*)

	Number of daughters born (*DB*)	Number of daughters selected (*DS*)		
Sire		NLB1	NLB2	NLB3
A	288	102	85	61
B	140	62	51	39
C	113	35	30	15
	Total daughters born (*TB*)	Total daughters selected (*TS*)		
	541	199	166	115

Applying the formula, impact (*SDI*) for sires A, B, and C can be calculated as -1.57%, +5.79%, and -4.22%, respectively, as weighted average deviations from what would be expected, in terms of percentage contribution to a lambing cohort, based on the number of daughters born.

The proportion of a ram's sons sold. The idea of this metric is to capture how efficient sires are at producing sons that sell. Data requirements for this metric include a count of the number of sons born to a sire, and the number of sons sold. Data for sons born to a sire are already in the SIL database through the pedigree. Sale information would be obtained through existing SIL status codes. This sale percentage metric (*SP*) would be calculated for each sire as a proportion and reported as a percentage of sons sold as:

$$SP_s = \left(\frac{SS_s}{SB_s}\right) \times 100 ,$$

where for sire s , *SS* is the number of sons sold and *SB* is the number of sons born over the sire's lifetime.

The following example calculates the proportion of a ram's (sire A) sons sold. Assuming sire A has produced 150 sons over his lifetime and 60 have been sold; the proportion of his sons sold is 0.4. This can be compared, for example, to sire B who has produced fewer sons (65) over his lifetime but 45 have been sold; a proportion of 0.69.

The average price of a ram's sons. The idea of this metric is to capture how efficient sires are at producing sons that sell at high prices. Data requirements for this metric include a count of the number of sons sold from a sire, and the individual ram sale price. Sale information would be obtained through existing SIL status codes but additional information on sale price would need to be added to the database.

The average price metric (*AP*) would be calculated for each sire, and reported as the average price of sons sold as:

$$AP_s = \frac{\sum P_s}{N_s} ,$$

where for sire s , *P* is the sale price of sons sold, and *N* is the number of sons sold. An addition to this calculation would be the total earnings per sire or the earnings per sire adjusted for the number of progeny (earnings per progeny born). The total earning per sire (*TE*) metric would be calculated as:

$$TE_s = \sum P_s$$

where for sire s , *P* is the sale price of sons sold. The earnings per progeny born (*TEP*) metric would be calculated as:

$$TEP_s = \frac{\sum P_s}{NP_s},$$

where for sire s , *P* is the sale price of sons sold, and *NP* is the number of progeny born.

The average price per unit of estimated breeding value, sub-index or overall index. The idea of this metric is to capture the gross income received by the breeder per unit of eBV, sub-index or index. Data requirements for this metric include individual eBVs, sub-index or overall index for each ram sold and individual ram sale price. Estimated breeding values, sub-indexes or overall indexes for each ram sold would be obtained from SIL genetic evaluations, and sale information would be obtained through status codes. Price information would be required.
The index price (IP) metric would be calculated for each year cohort of sold rams, and averaged over cohorts, as:

$$IP_y = \overline{\left(\frac{I_y}{P_y}\right)},$$

where for sale year y , *I* is the eBV, sub-index, or index of rams sold and *P* is the sale price of sons sold.

DATA AVAILABILITY

Key pieces of data required for the calculation of each metric have been detailed. Counts of animals born and daughters lambing, statuses and ewe exit fate codes can be obtained from the SIL database. Sale prices are not currently recorded on SIL. However, if the metrics described here were available, it is expected this would be the incentive to record such data on SIL.

CONCLUSION

Breeders are interested in generating more rams that are of higher genetic merit and most importantly are of sale quality. There is value in producing metrics that assess the relative merit of sires in producing progeny (male and female) that deliver more value genetically and financially. Implicitly this includes commercial culling criteria such as physical soundness, not just genetic information from SIL. These metrics offer a practical way to include those criteria with an appropriate emphasis in the ram selling business. Breeders appeared to put much less business value on metrics related to how much clients pay for rams, or how many years clients had been buying rams.

Calculation of these metrics is simple. However, in order to calculate metrics described in this report, and for the results to offer value to ram breeders, the following requirements must be met:

- Accurate and complete recording of status and exit fate for all rams and ewes
- Recording of sale price individually for all rams sold

The findings of this study need to be extended by collecting price data and surveying of more breeders.

REFERENCES

Byrne T.J., Amer P.R., Fennessy P.F., Hansen P., and Wickham B.W., 2012. *Animal* **6:**778.

Newman S.A., Dodds K.G., Clarke J.N., Garrick D.J., McEwan J.C., 2000. *Proc. N.Z. Soc. Anim. Prod*. **60:**195.

Wickham B.W., 2012. *International Committee for Animal Recording*, Cork, Ireland. http://www.icar.org/Cork_2012/index.htm.

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:53-56

CAN SELECTION FOR BRIGHTER, WHITER MORE PHOTOSTABLE WOOL REPLACE OXIDATIVE BLEACHING DURING WOOL PROCESSING?

S. Hatcher[1,2], J.W.V. Preston[1,2] and K.R. Millington[1,3]

[1]CRC for Sheep Industry Innovation, Homestead Building, UNE, Armidale, NSW, 2350
[2]NSW DPI, Orange Agricultural Institute, Forest Road, Orange, NSW, 2800
[3]CSIRO Materials Science and Engineering, Geelong Technology Precinct, Deakin University, Waurn Ponds, VIC, 3216

SUMMARY

Responses to selection were predicted for brightness, whiteness and photostability of Merino fleece wools using five common sheep breeding objectives based on Merino or Dual Purpose production systems. Genetic parameters for brightness, whiteness and photostability estimated from the Cooperative Research Centre for Sheep Industry Innovation's (Sheep CRC) Information Nucleus (IN) flock were used in the predictions. Breeding objectives with a high emphasis on reducing fibre diameter will generate small correlated improvements in both brightness and whiteness (0.15 and -0.16 T units respectively) however the responses achieved in 10 years are considerably lower than those required to render oxidative bleaching during processing unnecessary. Including brightness and whiteness as selection criteria produced slight improvements in the predicted response although the increases being just 0.01 across all indexes are unlikely to be of practical significance. Each of the indexes produced correlated improvements in photostability but are again of too low an extent (i.e. -0.05 to -0.22) to overcome the 3 T unit detrimental impact of bleaching on photostability, even when photostability was included as a selection criterion. Based on these predictions, the responses in brightness, whiteness and photostability achieved though the use of common Merino selection indexes are not sufficient to replace the routine use of oxidative bleaching of wool during processing.

INTRODUCTION

Bright, white and pastel shade products are essential for the growing markets for casual clothing, trans-seasonal knitwear, sports and leisurewear (Millington *et al.* 2013). Cotton and polyester fibres are both significantly whiter than Merino wool with typical yellowness (Y-Z) values of 2 and -4 respectively (Millington and King 2010) compared to a range of 6 – 11 for Merino wool (Millington *et al.* 2013). Brightness values > 71.5 and whiteness < 8.5 are considered desirable targets for Merino wool (Wood 2002), and as a result wool destined for use in bright white and pastel shade garments is always bleached with hydrogen peroxide during processing to improve its whiteness (Millington *et al.* 2013). The bleaching process improves brightness to between 80 to 85 T units and whiteness to between 3 to 3.5 T units (Millington *et al.* 2013). However peroxide bleaching has detrimental effects on photostability, as treated wool photoyellows to a greater extent (about 3 T units higher) than untreated wool (Millington *et al.* 2013), and handle and softness also deteriorate (become harsher) with treatment (Millington and King 2010). On-farm selection for brighter, whiter wool may provide an alternative to peroxide bleaching, though selection responses in the order of 5-10 T units for brightness and 5 T units for whiteness are necessary to produce similar brightness and whiteness values as the process of bleaching. This paper reports predicted responses to selection in brightness, whiteness and photostability for common Merino breeding objectives. The aim of the study was to determine whether sufficient improvement in both brightness and whiteness can be achieved through the use of common Merino selection indexes, making oxidative bleaching during processing unnecessary.

MATERIALS AND METHODS

Responses to selection were predicted using selection index methodology (Cameron 1997) following the procedure outlined by Mortimer *et al.* (2010). Three common breeding objectives based on Merino wool production systems were used (Merino 14%, Merino 7% and Merino 3.5%) as well as two used in dual purpose wool and meat enterprises (Dual Purpose 3.5% and Dual Purpose 7%). The different breeding objectives are based on micron premiums (calculated as the increase in the price of wool associated with a reduction of 1 micron in fibre diameter, expressed as a percentage) with higher percentages indicating a greater emphasis placed on reducing fibre diameter relative to wool production. The Dual Purpose breeding objectives have a higher emphasis on liveweight and reproduction than the Merino breeding objectives (Mortimer *et al.* 2010). The relative emphasis of each trait in each of the selection indexes is outlined in Mortimer *et al.* (2010). Predicted responses to selection were calculated for each of the then standard MERINOSELECT™ indices (Brown *et al.* 2007). Selections were based on 4 selection criteria (greasy fleece weight GFW; fibre diameter FD; fibre diameter coefficient of variation FDCV and; liveweight LWT) with responses predicted in the breeding objective traits as well as brightness, whiteness and photostability. All predicted responses were calculated over a 10 year period for a typical ram breeding flock without introductions of outside genetics. Responses in brightness, whiteness and photostability were monitored by assuming a relative economic value of zero.

The correlation structure for the breeding objective traits was based on Mortimer *et al.* (2009), preliminary estimates of genetic parameters for brightness, whiteness and photostability together with the measurement and analysis protocol are provided by Hatcher *et al.* (2010) and Hatcher *et al.* (2011). Updated estimates of the phenotypic variance, heritability and correlations (phenotypic and genetic) for brightness (σ^2=0.77, h^2=0.51), whiteness (σ^2=2.85, h^2=0.33) and photostability (σ^2=0.23, h^2=0.17) used in this paper are taken from the latest analysis incorporating data from the 2008 – 2011 shearing of the Sheep CRC IN (van der Werf *et al.* 2010) Merino yearlings (Hatcher and Preston unpublished). Scoured wool colour measurement on IN wool samples was carried out according to the standard IWTO test method (IWTO 2003). Photostability was measured using a method based on exposure to UVB (280-320nm) radiation for 4 hours (Millington and King 2010).

RESULTS AND DISCUSSION

The genetic variation in brightness, whiteness and photostability indicates that each trait will respond to selection. The high heritabilities of brightness and whiteness (0.51 and 0.33 respectively) indicate that significant progress could be made if single trait selection was undertaken. However, as the key economic traits of importance to Merino wool production are clean fleece weight (CFW), FD and staple strength (SS) (Swan *et al.* 2007) it is unlikely single trait selection for brightness, whiteness and photostability, which are essentially traits of secondary economic importance, will occur. In this context it is the predicted responses in the colour traits resulting from the use of selection indexes commonly used in the Merino industry that will determine the likely rate genetic progress under current market conditions.

Breeding objectives with a high emphasis on reducing FD generate correlated improvements in brightness over a 10 year period. However the changes are small, just 0.11 and 0.15 Tristimulus units (T units) for the Merino 7% and 14% indexes respectively (Table 1). Breeding objectives with a high emphasis on CFW will generate little correlated change in brightness with an improvement of 0.03 T units in 10 years. However, both the dual purpose breeding objectives will result in deterioration in brightness, by -0.18 and -0.06 T units respectively. Including brightness as selection criteria in the breeding objective did generate a slight improvement in brightness, though the effect was nominal just 0.01 T units across the range of indexes.

Only the Merino 14% index generated a correlated improvement in whiteness, with wool predicted to become whiter by -0.16 T units following 10 years of selection (Table 1). The other

indexes would all lead to a deterioration in whiteness (i.e. an increase in yellowness) of between 0.08 (Merino 7%) to 0.48 T units (Dual Purpose 7%) after 10 years. Including whiteness as a selection criterion increased the correlated improvement in whiteness by the Merino 14% index, by just 0.03 T units, although there was little impact on the whiteness in any of the other indexes.

Table 1. Predicted responses in brightness, whiteness and photostability (Tristimulus Units) from selection over 10 years, using a) standard MERINOSELECT™ indexes and b) including brightness, whiteness and photostability as selection criteria

	Merino			Dual Purpose	
	3.5%	7%	14%	3.5%	7%
a) Brightness, whiteness and photostability <u>not included</u> in the selection criteria					
Brightness	0.03	0.11	0.15	-0.18	-0.06
Whiteness	0.35	0.08	-0.16	0.48	0.20
Photostability	-0.16	-0.11	-0.05	-0.22	-0.17
b) Brightness, whiteness and photostability <u>included</u> in the selection criteria					
Brightness	0.04	0.12	0.16	-0.17	-0.06
Whiteness	0.36	0.07	-0.19	0.50	0.20
Photostability	-0.16	-0.11	-0.04	-0.22	-0.17

Each of the indexes generated correlated improvements in photostability (i.e. reductions in photoyellowing). The correlated improvements were greatest for the dual purpose and Merino breeding objectives with a higher emphasis on CFW due to the favourable negative genetic correlation between CFW and photostability (-0.36, Hatcher and Preston unpublished), such that higher CFW is genetically associated with a reduced propensity of wool to yellow following exposure to UVB radiation. Despite this favourable genetic relationship, including photostability as a selection criterion into the breeding objective did not significantly change the predicted response after 10 years of selection (Table 1). Given that oxidative bleaching decreases photostability by 3 T units (Millington *et al.* 2013), the small correlated improvements in photostability achievable through using common Merino selection indexes are likely to be of little commercial significance.

Selection responses of between 5 to 10 T units for brightness and 5 T units for whiteness are necessary to remove the requirement for peroxide bleaching of wool during processing (Millington *et al.* 2013). The average brightness, whiteness and photostability of the Sheep CRC IN yearling progeny shorn from 2008 to 2011 was 70.0, 8.2 and 4.5 T units respectively. Based on these averages, the correlated responses in the colour traits achieved through using common Merino selection indexes would need to generate improvements of 10 to 15 T units for brightness and 4.7 to 5.2 T units for whiteness to make oxidative bleaching unnecessary. The results of this paper indicate that correlated improvements of such magnitude are not possible from commonly used Merino selection indexes. Increasing the selection emphasis applied to the colour traits in these selection indexes will elicit a greater genetic response, however, this is likely to decrease the potential gains CFW, FD and staple strength which are key drivers of profitability in Merino enterprises. However these responses were modelled on a 'typical' ram-breeding flock with no use of outside sires. Using introduced sires with significantly brighter, whiter and more photostable fleeces is an option to speed up the rate of genetic improvement. The range of estimated breeding values (EBV) of sires used in the IN was 2.8 for brightness (-1.4 to 1.4), 3.4 for whiteness (-1.8 to 1.6) and 0.9 for photostability (-0.4 to 0.5), indicating that selection of sires with brighter and whiter wool is possible for brightness and whiteness but less so for photostability due to the narrow EBV range for that trait. The issue then becomes one of identifying whiter sires in industry

which is problematic given the low adoption rate of colour measurement in the Australian sheep industry (Millington *et al.* 2011) and the routine use of subjective assessment of greasy wool colour by breeders for ram selection (Brown 2006) rather than the more expensive direct objective measurement of brightness and whiteness on scoured wool. The moderate genetic correlation between greasy colour and whiteness (0.52) (Hatcher *et al*. 2011) among IN progeny indicates that greasy wool colour is a useful alternative selection criterion for whiteness, nevertheless the accuracy of identification of whiter sires may be compromised if whiteness is not directly measured given the narrow range of sire EBVs.

Genetics does play a role in controlling the brightness, whiteness and photostability of wool, as the heritability of these traits ranges from 0.17 (photostability) to 0.51 (whiteness). However many other variables including fibre diameter variation, differences in grazing environments (UV intensity, soil type, soil chemistry and ambient temperature), time of shearing, sebum production in the skin and the presence of bacteria mediate the expression of these traits (Hatcher *et al.* 2010) and may place an upper threshold on the achievable selection response. Sheep coats are an economically viable on-farm management intervention to increase whiteness, by 1 unit (Hatcher *et al.* 2003) depending on the length of time they are worn following shearing (Hatcher *et al.* 2008). As the current whitest Merino wools have whiteness of 6 T units (Millington *et al.* 2013), using sheep coats in conjunction with selection is unlikely to generate sufficient additive improvement in brightness or whiteness to make the use of routine oxidative bleaching of wool during processing unnecessary. Chemical bleaching of wool fibre during processing will remain necessary for wool to gain and maintain market share in the trans-seasonal knitwear, sports and leisurewear markets.

ACKNOWLEDGEMENTS

We acknowledge the dedicated team of scientists and technical staff at each of the 8 INF sites who are responsible for management of each of the flocks and implementation of the comprehensive sampling, assessment and measurement protocols.

REFERENCES

Brown D.J. (2006) *Int. J. Sheep Wool Sci.* **54**: 1.

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.

Cameron N.D. (1997) *Selection indices and prediction of genetic merit in animal breeding.* Edinburgh, UK: CAB International.

Hatcher S., Atkins K.D. and Thornberry K.J. (2003) *Aust. J. Exp. Agric.* **43**: 53.

Hatcher S., Atkins K.D. and Thornberry K.J. (2008) *Aust. J. Exp. Agric.* **48**: 762.

Hatcher S., Hynd P.I., Thornberry K.J. and Gabb S. (2010) *Anim. Prod. Sci.* **50**: 1089.

Hatcher S., Thornberry K.J. and Gabb S. (2011) *Assoc. Advmt. Anim. Breed. Genet.* **19**: 235.

International Wool Textile Organisation (2003) Method for the measurement of colour of raw wool IWTO-56-03 Brussels IWTO.

Millington K.R., del Giudice M., Hatcher S. and King A.L. (2013) *J. Text. Inst.* In press.

Millington K.R. and King A.L. (2010) *Anim. Prod. Sci.* **50**: 589.

Millington K.R., King A.L., Hatcher S. and Drum C. (2011) *Col. Tech.* **127**: 297.

Mortimer S.I., Atkins K.D., Semple S.J. and Fogarty N.M. (2010) *Anim. Prod. Sci.* **50**: 976.

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.

Swan A.A., van der Werf J.H.J. and Atkins K.D. (2007) *Assoc. Advmt. Anim. Breed. Genet.* **17**: 483.

van der Werf J.H.J., Kinghorn B.P. and Banks R.G. (2010) *Anim. Prod. Sci.* **50**: 998.

Wood E. (2002) *Wool Tech. Sheep Breed.* **50**: 121.

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:57-61

CHARACTERISTICS OF EXTENDED LACTATION AND PERSISTENCY IN AUSTRALIAN DAIRY COWS

M. Abdelsayed, P.C. Thomson, and H.W. Raadsma

ReproGen Animal Bioscience Group, Faculty of Veterinary Science, University of Sydney, Camden, NSW, 2570.

SUMMARY

The aim of this research is explore variation in Australian dairy cattle in their lactation curves over an extended lactation and obtain derived traits that could be used in genetic analyses. The Wood model was fitted to milk yield records from a random subset of 6,018 pure Holstein cows with 244,183 test-day records (29,882 lactations). Two traits of interest, namely persistency and extended lactation, were quantified and relevant descriptive traits derived. Variation among cows in their ability to maintain high production over a longer period of time was evident and a representation of the shape of average lactation curves in Australian dairy cows is presented. Findings showed that milk production during extended lactation phase (from day 305 to day 610 of lactation) is on average equivalent to 40% of the production of the first 305 days of lactation (standard lactation) with an average milk yield over the extended lactation of 8,887 L. Preliminary estimates of heritability for traits describing milk yield under extended lactation and persistency are in order of 0.10 and 0.09. This research will provide dairy farmers with a breeding tool to select cows that are best suited to milk for longer than the traditional 305 days.

INTRODUCTION

There has been a shift especially in Victoria, Australia for herds having a seasonally concentrated calving pattern from 63% in 2004 to 41% in 2006 (Dairy Australia, 2006). The reason for such a shift is from improved feeding of cows and the introduction of new germplasm from North American Holstein Friesian animals into some Australian dairy herds. Such impacts have resulted in an increase in the genetic potential of cows to produce more milk, while at the same time causing a decrease in reproductive performance largely as a consequence of changes in metabolic and physiological lactation requirements. The consequence is an ongoing trend of cows being milked beyond the traditional 305 day system to manage decreasing fertility through retaining high productive cows milking for longer, resulting in healthier, more productive cows and more profit for the dairy producer.

Lactation curve models are useful in helping to define and estimate lactation characteristics of individual cows for genetic selection (Dekkers *et al*. 1998; VanRaden *et al*. 2006), predicting milk yields and milk components, analyses responses of yield to environmental and management changes, and identify opportunities for maximising net value effectively (Dematawewa *et al*. 2007; Dijkstra *et al*. 2010). Thus a fundamental aspect of evaluating extended lactation is the modelling of extended lactation and persistency traits in Australian dairy cattle based on herd recording data.

While trait definitions in the literature differ, heritability estimates for extended lactation are in the range of 0.19 to 0.30 (±0.02) and for persistency traits 0.03 to 0.30 (±0.03). These moderate heritability estimates suggest that these traits are likely to respond well to selection. Limited information available for Australian dairy cows (Haile-Mariam and Goddard, 2008) suggests there are both phenotypic and genetic differences in the ability of cows to continue to milk for long periods. Furthermore predictions of which cows are better at milking for longer can be made based on their previous traditional 305 day milking performance. This current project examines the genetic differences observed in Australian dairy cows that can be successfully milked for longer than 305 days. Estimated breeding values (EBVs) for these cows will be derived, which to date is

not available. For this paper the main focus is on preliminary exploration of phenotypes and the variation that exists between cows in their milk yield profiles (lactation curves) over an extended lactation and heritability estimates.

MATERIALS AND METHODS

Data were obtained from the Australian Dairy Herd Improvement Scheme (ADHIS) including approximately 158 million test-day records from 1985 to 2010 derived from around 7 million cows. Extended lactation milk traits include milk yield, fat, protein, lactose percentage, Australian Selection Index (fat + protein – volume) and energy outflow of fat, protein and lactose, as a measure of energy per lactation. This paper presents some preliminary analyses for milk yield using a random subset of 6,018 pure Holstein cows with 244,183 test-day records and 29,882 lactations after data editing.

Extended lactation curves have been modelled using the Wood (1967) model, the output from this model then being used to quantify the variation in the shapes of cows' lactation curves, particularly in relation to extended lactation and persistency. The model contains three parameters namely (*a*) an overall scaling factor, parameter (*b*) related to the rate of increase prior to the peak yield and parameter (*c*) related to the rate of decline after the peak (Wood, 1967). The model was fitted to the milk-yield data using the nlme nonlinear mixed model package in R (Pinhero and Bates, 2000), but specifically, each cow-lactation returned a set of (k=$\log_e a$, b, c) parameter estimates; further specific details can be found in Hall (2008) and Jonas *et al*. (2011). Other yield and milk component traits will be examined subsequently. In the context of this study, persistency (r_{305}) is defined as the ratio of model-based milk yield at day 305 (y_{305}) to model-based milk yield at peak (y_{max}) (Hall, 2008; Jonas *et al*., 2011), and extended lactation (Ext Lac, XLAC) is defined as the ratio of expected milk yield from day 305 to day 610 ($Area_B$), given cows are in lactation for 2 years, relative to the cumulative yield up to day 305 ($Area_A$). Figure 1 illustrates the model approach. The cumulative yields comprising $Area_A$ and $Area_B$ are obtained from mathematical expression based on the ($k = \log_e a$, b, c) for each cow-lactation (Hall, 2008; Jonas *et al*., 2011).

Genetic parameter estimates and estimated breeding values for these traits were derived using linear mixed animal models using the ASReml-R statistical program. However, for this paper only heritability and repeatability estimates have been reported. Various combinations of fixed effects were fitted and the best model includes herd, year and season and parity group which was analysed as a two category factor, parity 1 = Maiden, Parity two or more = Adult. These effects were fitted additively instead of combining herd, year and season as a proxy to identify the effects of each on extended lactation.

The animal model fitted to the phenotypic data was:

$$Y = \mu + H + TY + CM + PG + A + CowID + \varepsilon$$

where Y = extended lactation (XLAC), persistency (r_{305}) or cumulative 610 milk yield; the fixed effects in the model were National.Herd.ID (H), testYear (start year when test day records were taken) (TY), calveMonth (month of calving/season) (CM) and parity group (2 categories parity 1= Maiden, parity 2 and above (Adult) (PG). The random effects in the model were Animal (A) (polygenic term incorporating pedigree structure) and National.cow.ID (CowID) to account for repeated lactations per cow as well as ε, a random error term.

RESULTS AND DISCUSSION

There is variation in the shape of lactation curves of different cows with extended lactation (beyond the traditional 305 day lactation) as shown in Figure 2. This is also supported with all measures expressing a high degree of variation (CV range 8% - 72%, Table 1). Some cows have a steeper rate of decline with a rapid drop in milk production straight after peak lactation while other cows have a slower rate of decline in milk yield after peak lactation. The latter are more persistent

cows and tend to have flatter curves than traditional 305 day lactation curves. Figure 2 also shows two different lactation curves, one illustrating a non-ideal lactation curve (worst cow: bold line A) and the other illustrating an ideal lactation curve (best cow: bold line B) in terms of high persistence while maintain peak production over a longer period of time. The best cow has a more persistent lactation ($\boldsymbol{r_{305}}$ = 0.92) curve where peak production is maintained for a longer period of time, c = 0.001847, which is a lower than average rate of decline after the peak. For the worst cow the rate of decline is higher (c = 0.004361) than all the other curves, and its persistency is higher than average but lower ($\boldsymbol{r_{305}}$ = 0.55) when compared to the best cow presented in Figure 1. It is evident that some cows that have extended lactation may not necessarily be highly persistent and vice versa, which is illustrated by the two curves labelled C and D in Figure 2 where one of the cows is highly persistent with high yield (C) while the other has a rapid decline from peak production (low persistency) and has lower yield (D).

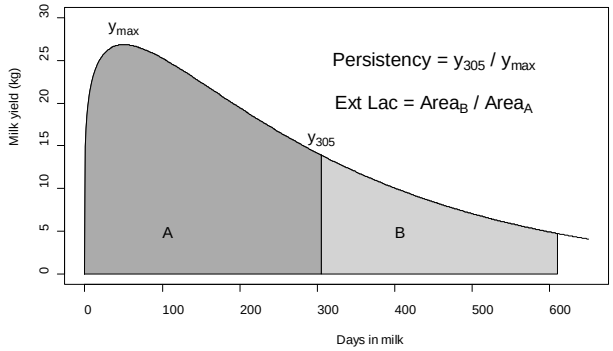


Figure 1. Definition of extended lactation and lactation persistency as a ratio of yields

Table 1 below shows summary statistics from the estimated Wood model parameters and derived persistency and extended lactation traits from a subset of 6,018 pure Holstein cows, from which the mean represents the full dataset (population). Milk production during extended lactation phase (from day 305 to day 610 of lactation) is on average equivalent to 40% of the production of 305 days lactation (standard lactation) with an average milk yield over the extended lactation of 8,887 L. The average persistency is 0.465, indicating that cows can maintain on average, almost 50% of their peak production up to day 305 of lactation.

Heritability estimates (Table 1) for extended lactation, persistency of milk yield and cumulative yield up to day 610 were 0.10 ± 0.03, 0.09 ± 0.03, 0.13 ± 0.05 with repeatabilities of 0.20, 0.19 and 0.42 respectively.

CONCLUSION

Overall there is considerable variation between cows in the Australian dairy herd for persistency and extended lactation. There are certain cows that have higher persistency than others and who are able to maintain production over a longer period of time (extended lactation). The derived traits adequately describe such differences between cows and could be used as input variables in genetic analyses. Genetic parameters such as heritability, genetic, phenotypic, environmental correlations and more importantly breeding value estimates can now be derived for extended lactation and lactation persistency. Thus the findings of such research will provide dairy farmers with a breeding tool to select cows (as well as bulls for breeding) that are best suited to milking for longer than the traditional 305 days.

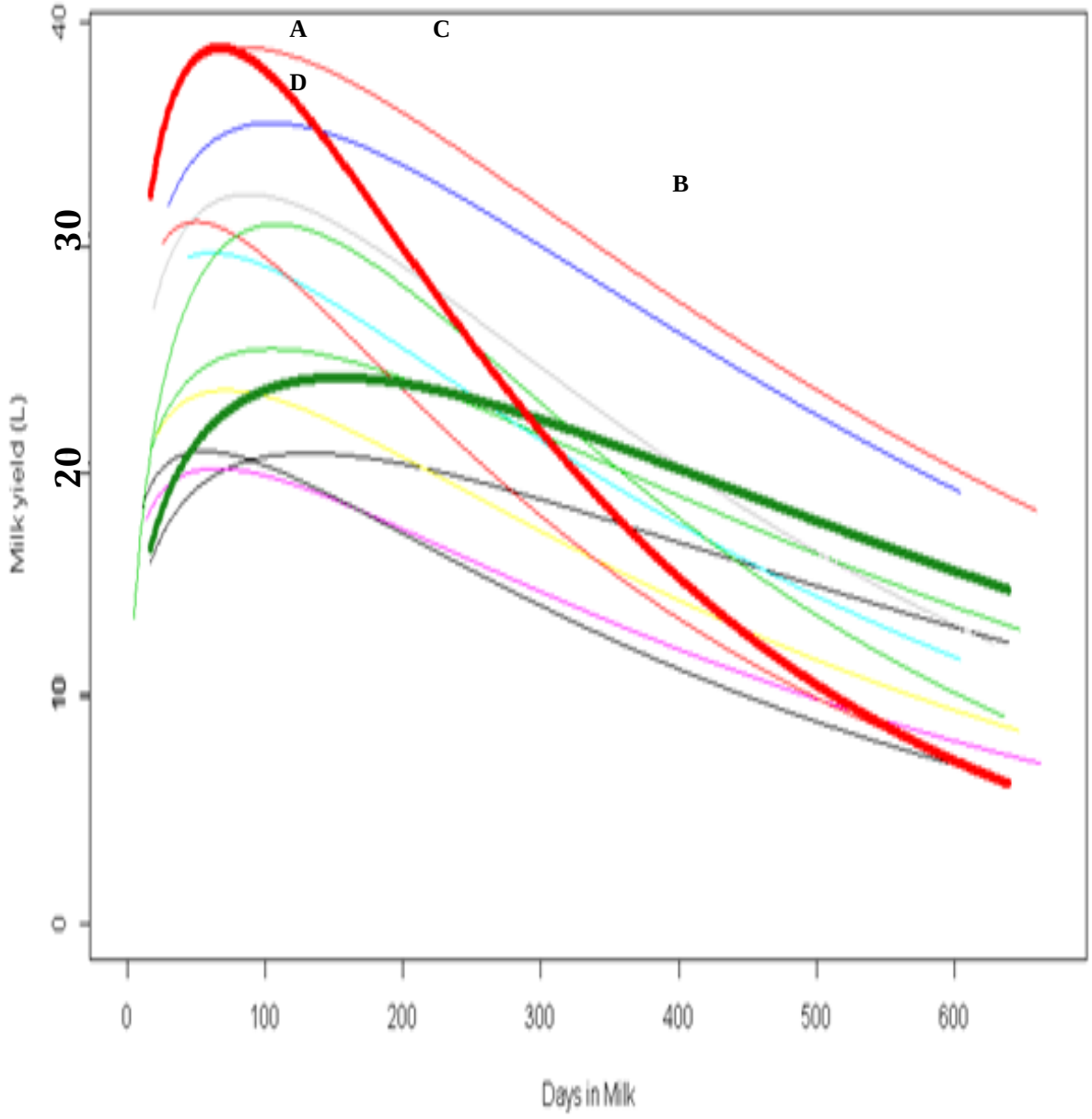


Figure 2. Illustration of the variety of shapes of lactation curves of a sample of 13 cows selected with extended lactations (610 days) modelled using the Wood model. There are cows with more persistent flatter curves, slower rate of decline in milk production after peak milk yield. The best (bold green) cow and worst (bold red) cow in terms of lactation persistency and extended lactation are highlighted.

Table 1. Summary statistics for Wood model parameter estimates (k(=$\log_e a$),b,c) of milk yield for subset of 6,018 pure Holstein cows, derived traits, heritability and repeatability for persistency (r_{305}) extended lactation (XLAC) and cumulative yield up to day 610

Trait	Mean	SD	Min	Max	CV (%)	h^2 ± SE	r
***k* (=log$_e$*a*)**	18.32	1.468	4.11	86.57	8		
b	0.14280	0.103	-0.268	0.555	72		
c	0.0040	0.0016	0.0001	0.0100	40		
r_{305}	0.465	0.148	0.073	0.999	32	0.09±0.03	0.19
XLAC	0.3917	0.147	0.082	0.9756	37	0.10±0.03	0.20
CUMYT610	8,887	2,634	2,891	19,631	30	0.13±0.05	0.42

Abbreviations: r_{305} = Persistency, XLAC = extended lactation, CumYT610 = cumulative yield (L) total up to day 610, CV = coefficient of variation (%), h^2 ± SE = heritability ± standard error and r = repeatability

ACKNOWLEDGEMENTS

The authors would like to thank the Australian Dairy Herd Improvement Scheme (ADHIS) for providing the dataset for this study, Dairy Futures CRC and the University of Sydney for the funding for this project.

REFERENCES

Dairy Australia (2006) Situation and Outlook. Available at: http://www.dairyaustralia.com.au

Dekkers J.C.M., Ten Hag J.H., and Weersink A. (1998). *Livstock. Prod. Sci.* **53**: 237.

Dematawewa C.M.B., Pearson R.E., and VanRaden P.M. (2007). *J. Dairy Sci.* **90**: 3924.

Dijkstra, J., Lopez S., Bannink A., Dhanoa M.S., Kebreab E., Odongo N.E., Fathi Nasri M.H., Behera U.K., Hernandez-Ferrer D., and France J. (2010). *J. Agri. Sci.* **148**: 249.

Haile-Mariam M., and Goddard M.E. (2008). *Animal* **2**: 325.

Hall E.J.S. (2008). PhD Thesis, University of Sydney.

Jonas E., Thomson P.C., Hall E.J.S., McGill D., Lam M.K., and Raadsma H.W. (2011). *Genet. Sel. Evol*. **43**: 22.

Pinheiro J.C., and Bates D.M. (2000). 'Mixed Effects Models in S and S-PLUS' Springer, New York.

VanRaden, P.M., Dematawewa, C.M.B., Pearson, R.E., Tooker, M.E. (2006). *J. Dairy Sci*. **89**: 3213.

Wood, P.D.P. (1967). *Nature* **216**: 164.

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:62-65

MODELLING VARIATION IN BOVINE MILK FAT COMPOSITION PREDICTED USING MID-INFRARED SPECTROMETRY

G. Johnstone[1], M. Coffey[2] and E. Wall[2]

[1]AbacusBio Limited, Dunedin 9016, New Zealand [2]Scottish Rural University College, The Roslyn Institute Building, Bush Estate, Midlothian, EH26 0PH, United Kingdom

SUMMARY

Applying calibration equations to mid-infrared (MIR) spectrometry is an efficient method to predict fatty acid composition in bovine milk. This investigation applied calibration equations to weekly MIR data of approximately 240 cows of mixed age and mixed calving date over a seven month period. Cows belonged to either a high yielding line selected for maximum milk fat and protein production or a control line representing the moving UK average for milk production. Random regression models with Legendre polynomials were fitted for the major fatty acids and fatty acid classes. Correlations were assessed between fatty acids and related milk production traits using bivariate regression. Results showed significant effects of genetic selection line on the fatty acid profile, whereby there was a greater proportion of saturated fatty acids in the milk fat of the high yielding line. Diet, parity number, week of lactation, date of lactation and several of their interactions also had significant affects on the fatty acid profile. Correlations were observed between fatty acids and production traits. The research demonstrates the practical application of this technique in the British dairy industry and highlights the effect of current selection practices on saturated fatty acids as an area of interest for future research.

INTRODUCTION

Research into the fatty acid profile of bovine milk has been limited in the past due to the expensive and time consuming nature of gas chromatography analysis. Mid-infrared (MIR) spectrometry is the method of choice for quantification of milk fat, protein, and lactose content of milk samples. Applying the Soyeurt *et al.* (2011) calibration equations to the MIR spectrum enables efficient analysis of fatty acid profiles for large datasets, not previously possible.

Fat content and composition of food is of growing consumer concern as related health problems continue to be a heavy burden on society. Cardiovascular disease, which is linked to high intake of saturated fatty acids (SFA), is the most common cause of death in the UK (Townsend *et al.* 2012). Milk and milk products are nutrient rich foods but suffer significant criticism as the main source of saturated fat in the diet, contributing approximately 31% and 22% of SFA in the British diet for 4-10 year olds and adults (19-64), respectively (Bates *et al.* 2011). Modifying the fatty acid profile of milk could be an effective way of reducing the SFA intake of the British population without requiring a change in dietary patterns. Four studies have reported a reduction in total and LDL-cholesterol, both risk factors for cardiovascular disease, following consumption of milk and dairy products with modified fatty acid profile (Livingstone *et al.* 2012).

Notably, the fatty acid profile of milk is highly variable, with large potential to make changes through animal nutrition (Ashes *et al*,1997). There is also evidence supporting the potential to change fatty acid profiles through genetic selection, with heritabilities reportedly as high as 0.43 for SFA (Bastin *et al.* 2011). Fatty acid profiles are also influenced by animal health, physiology and energy balance. Therefore, fatty acid profiling could potentially serve as an early indicator of these and related traits for use in cow and herd management.

This paper reports on novel methodology to model variation in the fatty acid profile. Development of appropriate models is important for more extensive genetic and phenotypic analysis.

MATERIALS AND METHOD

Animal data. Data recording, including spectral analysis, was performed weekly on morning, noon and evening milkings of approximately 240 cows of mixed age and calving date. Cows were based at the Dairy Cattle Research Centre at Crichton Royal Farm, Dumfries, Scotland. They belong to either the Langhill high-yielding line selected for maximum milk fat and protein production (select) or a control line representing the moving UK average for milk production (control).

The cows were fed on either a home-grown forage diet (home-grown) or on a bought-in by-product feed (by-product). Over summer, the animals on the home-grown forage diet were at grass during the day and overnight they were being offered a feed of appropriate home-grown ingredients to balance the high protein and relatively low neutral detergent fibre of the grass. The winter ration consisted of grass silage, maize silage, lucerne, red clover silage, field beans, crimped wheat and vitamins and minerals fed as a complete total mixed ratio (TMR).

The by-product diet was based on ingredients available following a primary production process and not normally used for human food. The ration consisted of straw, Vitagold, sugar beet pulp, biscuit meal, feed-grade breakfast cereal meal, soya bean meal, wheat distiller's dark grains, molasses, megalac (calcium soap-bound palm fatty acid distillate) plus vitamins and minerals.

Milk samples were analysed using a calibrated mid-infrared FOSS MilkoScan FT6000 spectrometer. Soyeurt *et al.* (2011) multivariate calibration equations were run on this data to quantify the fat and fatty acid composition of the milk. Notably, 4 milk samples from the Crichton dairy herd were used to add variability to the validation set of samples, and 102 milk samples from this herd were used in the cross-validation dataset to evaluate the efficacy of the calibration equations. The cross-validation demonstrated high predictive accuracy (cross-validation coefficient of determination, $R^2cv > 0.95$) for most of the SFAs, the main monounsaturated fatty acids (MUFA) and most of the main fatty acid groups, including SFA, MUFA, unsaturated, short chain, medium chain and long chain fatty acids.

From the fatty acid predictions (in g/dL of milk) the amount of the individual fatty acids and fatty acid groups as a percentage of total fatty acids was calculated. Fatty acid percentages were averaged across the morning, midday and evening milkings for each cow on each test day.

Cow information, daily milk yield (kg), fat yield (kg) and protein yield (kg), average weekly live-weight (kg) and body condition score were extracted from the routinely collected RobustMilk Langhill database for the test cattle.

Editing of data. Before statistical analysis, records were removed where all 3 milkings were not performed on a test day, where records were taken after the 45^{th} week (310 days) of milking, where total daily milk-yield was less than 4L and where total fat content was less than 1.5g/dL of milk or greater than 9g/dL of milk. Records were eliminated for production traits and fatty acids if the value given was negative or outside 3 standard deviations of the mean.

Data analysis. The fatty acids and classes were analysed as repeated measures per cow by week of lactation using random regression models with legendre polynomials using the ASReml programme (Gilmour *et al.* 2006). Fixed effects included week of lactation, parity number, genetic group, feed group and all combinations of their interactions. Additionally, date-of-milking and date-of-milking-by-feed, animal age at milking and calving month were fitted in the model as (co)variance components. The best order polynomial was selected for the lactation cycle based on the log-likelihood ratio tests and significance of the curves at node points fitted for each trait, with the majority either fitting to order 4 or 5, plus intercept. The final models also contained the random effect of week of lactation (to the same order of polynomial, less the intercept) by individual animal.

Correlations were assessed between fatty acids and between fatty acids and related milk production traits using bivariate random regression, based on the models derived. The polynomial order of the random effect was limited to quadratic due to calculation constraints.

RESULTS AND DISCUSSION

Random regression models. Week of lactation, feed type, genetic line and date of milking had a significant effect on all of the fatty acids and fatty acid classes. Parity number had a significant effect on all fatty acids except the PUFA and trans fatty acids.

The effect of genetic line suggests that selective breeding for increased fat and protein yield could be altering the fatty acid profile of milk in an unfavourable direction. The select line cows had, on average, a 1.8% (95% C.I. 0.6-2.9) higher percentage of SFA in total fatty acids.

However, as expected, feed had a greater effect on FA profile than genetic line. There is, on average, 10.6% (95% C.I. 9.2-12.1) less SFAs as a percentage of total fatty acids in the cows fed the by-product diet. This feed is supplemented with calcium bound palm fatty acid distillate (Megalac®) which is high in both C16:0 and C18:1 fatty acids. Furthermore, feeding of supplemental fat inhibits *de novo* synthesis of the short to medium chain saturated fatty acids in the mammary glands (C4:0 to C16:0) so the percentage of C18:1 in the milk increases and the percentage of saturated fatty acids will, on balance, be lower. This is consistent with results reported previously for Megalac® (Fearon *et al*. 1994).

While there was no significant feed-by-line interaction for saturated fatty acids as a group, this affect was significant for most of the short and medium chain saturated fatty acids when analysed individually. This gene-by-environment interaction suggests the effect of diet on mammary gland synthesis differs between the control and high yield genetic lines.

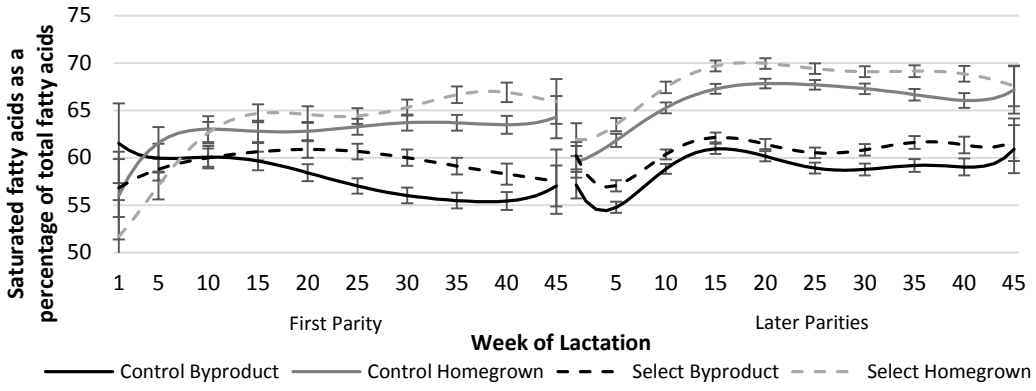


Figure 1: Predicted saturated fatty acid content as a percentage of total fatty acids, for first and later lactations across 45 weeks in milking.

Significant and nutritionally interesting variation is also seen for the main *trans* fatty acid in milk, where by the select line cows express significantly more C18:1 *trans* fat than the control line cows. *Trans* fatty acids, like SFA, have an adverse effect on cholesterol levels and cardiovascular disease. However, there was also a significantly greater proportion of beneficial omega-3 polyunsaturated fatty acids (PUFA) in the milk of the select line cows. Note that the accuracy for the omega-3 PUFA callibration equations falls below the statistical standard required for predictions (RPD 2.0, R^2cv 0.75). This may nevertheless warrant further investigation because,

given that omega-3 fatty acids cannot be synthesised *de novo* by mammals, this suggests that the select line cows have greater uptake of the omega-3 fatty acids from the feed into their milk.

There was significant variation in profile over the course of lactation, especially evident in later parities where patterns are more settled (Figure 1). The low SFA percentages in early lactation corresponded to when cows were producing peak milk yield (Coffey *et al.* 2004) and therefore were typically in negative energy balance. To compensate for this imbalance, fat stores in adipose tissue, which are rich in C18:1*cis9* MUFA, are mobilised and released into the udder and make up a greater proportion of the milk fat (Stoop *et al*. 2009). Notably, the by-product control line gets as low as 54.4% (s.e. 0.60) which could have a significant effect on the functional properties of the milk e.g. in butter or cheese (Ashes *et al.* 1997).

Correlations. There were some significant correlations between fatty acids and milk yield (kg/day), fat yield (kg/day), fat content (g/dL of milk) and condition score across lactation cycle. Standard errors were typically high in early and late lactation, corresponding to low data counts.

Fat content and percentage of saturated fatty acids were generally positively correlated - between week 11 and 40 of lactation there was a significant positive correlation of between 0.45 and 0.56. This corresponds with the higher saturated fat content in the cows selected for increased fat yield.

CONCLUSIONS

Best-fit models have been developed for fatty acids and related traits based on results of calibration equations applied to MIR spectral data. The models highlight that selection for fat and protein yield had a detrimental effect on saturated fat content. Given the relationship between saturated fat intake and human health, further research into this area is warranted.

ACKNOWLEDGEMENTS

This work was funded by the Scottish Government. National Milk Records, UK undertook the MIR analysis of the milk samples. Thanks to Ian Archibald and Alan Mason for maintaining the Langhill database and Ainsley Bagnall and farm staff at Crichton Royal Dairy Research Centre.

REFERENCES

Ashes J.R., Gulati S.K., and Scott T.W. (1997) *J. Dairy Sci.* **80**: 2204.

Bastin C., Berry D.P., Soyeurt H. and Gengler N. (2011) *J. Dairy Sci.* **94**: 4152.

Bates B., Lennox A., Bates C. and Swan G. (2011) "National Diet and Nutrition Survey" Department of Health, London.

Coffey M. P., Simm G., Oldham J. D., Hill W. G. and Brotherstone S. (2004) *J. Dairy Sci.* **87**: 4318.

Fearon A.M., Charlton C.T. and Kilpatrick D.J. (1994) *J Sci Food Agric.* **66**: 247.

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.

Livingstone K.M., Lovegrove J.A. and Givens D.I. (2012) *Nutr. Res. Rev.* **1:** 1.

Soyeurt H., Dehareng F., Gengler N., McParland S., Wall E., Berry D.P., Coffey M. and Dardenne P. (2011) *J. Dairy Sci.* **94:** 1657.

Stoop W.M., Van Arendonk J.A.M., Heck J.M.L., Van Valenberg H.J.F. and Bovenhuis H. (2008) *J. Dairy Sci.* **91:** 385.

Stoop W.M., Bovenhuis H., Heck J.M.L. and Van Arendonk J.A.M. (2009) *J. Dairy Sci.* **92**: 1469.

Townsend N., Wickramasinghe K., Bhatnagar P., Smolina K., Nichols M., Leal J., Luengo-Fernandez R. and Rayner M. (2012) "Coronary Heart Disease Statistics" 2012 ed. British Heart Foundation, London .

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:66-69

STRATEGIES TO OBJECTIVELY GROUP MERINO FLOCKS IN SHEEP GENETICS

D.J. Brown[1], A.A. Swan[1], J.S. Gill[2] and R.B. Banks[1]

[1]Animal Genetics and Breeding Unit[*], University of New England, Armidale, NSW 2351
[2]Meat and Livestock Australia, University of New England, Armidale, NSW 2351

SUMMARY

The Merino breeding population in Australia exhibits considerable diversity in objectives and breeding philosophies. It could be beneficial for both the analysis and reporting of the national genetic evaluation to objectively group flocks into logical subsets. This study evaluates techniques to cluster flocks into logical groups based on either estimated breeding values or genomic information. Principle component analyses were conducted using flock mean breeding values and the genomic relationship matrix. Using the flock mean breeding values, 6 clusters of flocks were identified with the first 2 principle components explaining 73% of the variation between flocks. The first principle component separated flocks based on overall productivity, with approximately equal emphasis across all traits. The second component separated flocks based on fleece weight, wrinkle and staple length. Less separation between flocks was apparent for flocks with below average fibre diameter. The principle components of the genomic relationship matrix were also strongly correlated with mean breeding values across the flocks. The lack of accurate Australian Sheep Breeding Values (ASBVs) for some traits and genomic information across some of the flocks is a limitation of this approach as it makes allocation of some flocks challenging.

INTRODUCTION

Over recent decades mixing of animals between strains of Merino sheep has become more widespread. The industry is made up of flocks with a range of breeding objectives and many breeding philosophies with varying levels of objective measurement and visual selection in use. Sheep Genetics is a genetic evaluation service which provides Australian Sheep Breeding Values (ASBVs) to sheep breeders (Brown *et al.* 2007). Sheep Genetics uses a "Type" classifier to allow separation of Merino flocks for reporting and comparison of genetic trends. However, at times this classifier is too simple to adequately group flocks. This type classification is mostly based on ASBVs ranges for key traits and breeder perception of their wool and sheep type.

The aim of this study was to evaluate techniques to objectively group flocks based on either their average breeding values for key traits or genomic information.

MATERIALS AND METHODS

Estimated breeding values. ASBVs were extracted from the Sheep Genetics MERINOSELECT 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. Traits extracted were Yearling live weight (Ywt), clean fleece weight (Ycfw), fibre diameter (Yfd), fibre curvature (Ycuv), staple length (Ysl) and breech wrinkle (Bwr). For each flock the average ASBV for each trait was calculated using all animals born since 2006. A summary of the data used for each key trait is shown in Table 1. There were 256 flocks with ASBVs available for all traits available.

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

Table 1. Summary of the variation in flock mean ASBVs for the 256 flocks used in this study

Trait	Mean	SD	Min	Max
Ywt (kg)	1.75	2.59	-4.81	9.40
Ycfw (%)	5.48	7.92	-17.53	25.85
Yfd (µm)	-1.05	0.68	-3.14	1.30
Ycuv (deg/mm)	-0.76	4.67	-11.58	17.34
Ysl (mm)	2.15	4.21	-12.52	16.71
Bwr (scores)	-0.05	0.29	-1.13	1.04

Genomic information. 50K Ovine SNP chip (Illumina Inc., San Diego, CA, USA) genotypes were available on 6230 merino animals from 203 flocks. Most of the animals genotyped were part of a reference population in which sires from these flocks were mated to ewes representative of several industry types. While 137 flocks had both ASBVs and at least one animal genotyped, only 51 flocks had ASBVs and 5 or more genotyped animals. In the main however, the set of animals included in ASBV means were independent from the set of animals genotyped.

Models of analysis. Principle component analysis of the standardised flock means for all six traits was conducted using the princomp procedure in R (R Development Core Team 2012) with the resultant 6 principle components (PC1 to PC6) used in the kmeans procedure to cluster flocks. As there are significant differences between traits in their means and variance, all traits were standardised (mean = 0, standard deviation = 1) to allow equal contribution of all traits to the principle component analysis. There is normally a reduction of the within-cluster variation as more clusters are used. However a balance is required between this and the number of flocks within each cluster and the interoperability of the clusters formed. After examining the results from using 4 to 7 clusters, 6 was chosen as the optimal number to adequately separate flocks into groups.

The genomic relationship matrix was calculated for all animals in the Sheep Genetics database with 50K SNP genotypes following the methods of Van Raden (2008) and Yang et al. (2010) and scaled so that the average diagonal element was 1. Principle components of this relationship matrix were then estimated using the singular value decomposition method in a purpose written program using the LAPACK numerical computation libraries. This program was used rather than R for speed of computation.

The first five principle components of the genomic relationship matrix were used for this study. The principle components were standardised prior to being averaged over flocks with animals represented. Flocks with fewer than 2 animals were removed leaving 87 flocks with genomic and ASBV based principle components.

RESULTS AND DISCUSSION

ASBV based principle components 1 to 6 explained 58%, 16%, 10%, 8%, 6% and 2% of variation respectively. The first 2 principle components explain most of the variation (74%). Table 2 shows the correlations between traits and principle components. The flock means were moderately correlated across flocks with most traits except wrinkle having correlations greater than 0.5 with the other traits. PC1 was highly correlated with all ASBVs and thus represented a discriminator of overall production across the Merino industry. The second principle component was more related to wrinkle, staple length and fleece weight, and a little with curvature and staple length, and is likely to separate flocks on style and breeding philosophy. The remaining principle components appeared to concentrate on individual traits.

Table 2 Correlations between traits (standardised) and with principle components

	Ywt	Ycfw	Yfd	Ycuv	Ysl	Bwr
Ywt		0.62	0.60	-0.46	0.54	-0.32
Ycfw	0.62		0.50	-0.72	0.47	-0.11
Yfd	0.60	0.50		-0.63	0.56	-0.25
Ycuv	-0.46	-0.72	-0.63		-0.74	0.28
Ysl	0.54	0.47	0.56	-0.74		-0.42
Bwr	-0.32	-0.11	-0.25	0.28	-0.42	
PC1	-0.78	-0.78	-0.80	0.87	-0.83	0.45
PC2	-0.02	-0.40	-0.08	0.14	0.18	-0.85
PC3	-0.58	-0.08	-0.12	-0.39	0.29	0.05
PC4	-0.02	-0.40	0.54	0.08	0.08	0.17
PC5	0.18	-0.19	-0.25	0.11	0.41	0.21
PC6	-0.09	0.15	0.06	0.25	0.11	-0.02

Cluster means for each trait are generally the best way to visualise the characteristics of the clusters formed. Table 3 shows these cluster means for each trait. For example cluster 3 appears to be the traditional superfine flocks which are low for live weight, fleece weight, fibre diameter and staple length but higher for curvature and wrinkle.

Table 3. Cluster means for each trait (standardised)

Cluster	Flocks	Ywt	Ycfw	Yfd	Ycuv	Ysl	Bwr
1	31	0.61	0.31	1.00	-1.17	1.56	-0.66
2	7	-0.25	-0.35	0.09	0.14	-0.13	-0.06
3	35	-1.22	-1.33	-1.34	1.39	-1.17	0.58
4	20	-0.38	0.44	-0.21	-0.18	-0.48	1.13
5	30	0.85	-0.13	0.15	0.47	-0.13	-1.67
6	14	0.72	1.01	0.40	-0.59	0.37	0.03

The principle components of the genomic relationship matrix were also moderately to highly correlated with the mean ASBVs for the flocks (Table 4). In particular the first genomic principle component appears to separate flocks based on overall production level. The second component appears to separate on fleece weight, curvature, staple length and wrinkle which may relate to flocks using traditional versus skin related breeding philosophies. This is consistent with the clustering based on flock mean breeding values.

Table 4. Correlations between the ASBV flock means and averaged principle components from the genomic relationship matrix (n=137)

	Ywt	Ycfw	Yfd	Ycuv	Ysl	Bwr
genPC1	-0.69	-0.69	-0.70	0.77	-0.65	0.36
genPC2	-0.03	0.21	0.01	-0.34	0.16	-0.18
genPC3	-0.15	0.00	0.12	-0.14	-0.13	-0.01
genPC4	0.11	-0.04	0.04	-0.15	0.36	-0.28
genPC5	0.18	0.00	0.04	0.06	-0.02	0.02

Principle component analysis relies on having all traits observed which reduces the number of flocks which we can cluster. There is also significant variability in the number of animals recorded

and the ASBV accuracy for some traits such as staple length and breech wrinkle. Furthermore the number of animals genotyped across the flocks varies greatly thus influencing the accuracy of the genomic relationship between flocks.

CONCLUSIONS

The results from these analyses clearly show that the Merino industry no longer has distinct types and for all traits considered there is a continuum of performance. Despite this the analysis was able to separate the flocks into 6 distinct groups. While clusters may have been similar for some components they were distinctly divergent for at least one principle component. Clustering based on both flock mean breeding values and on genomic relationships was consistent, with two main clustering dimensions emerging as being the key production traits and plainness of body. This is not altogether surprising since the genetic analysis upon which the breeding value clustering is based, includes genetic relationships. The consistency suggests a reasonably high level of pedigree accuracy in the Sheep Genetics data.

The main reason for exploring this approach was to consider how best to group flocks for both analysis and reporting of results. Sensible clusters would help breeders interpret and filter the mass of breeding value information to make appropriate selection decisions. Both the ASBVs and genomic-based clusters appear intuitively sensible. Further work and consultation with breeders is warranted to determine how best to combine the two sources of clustering information, how to use the resulting clusters in analysis, and whether this approach offers an improvement in clarity for breeders.

ACKNOWLEDGMENTS

This research was funded by Meat and Livestock Australia and Australian Wool Innovation, and Sheep Genetics is made possible through the support of the Australian sheep industry. The authors acknowledge the contributions of the Sheep CRC Information Nucleus, the Australia Merino Sire Evaluation Association and industry-funded research flocks.

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. Ass. Adv. Anim. Breed. Genet.* **17**: 187.

R Development Core Team (2012) Statistical Computing, http://www.R-project.org.

VanRaden P.M. (2008) *J. Dairy Sci.* **91**: 4414.

Yang J., Beben B., McEvoy B.P., Gordon S., Henders A.K., Nyholt D.R., Madden P.F., Heath A.C., Martin N.G., Montgomery G.W., Goddard M.E. and Visscher P.M. (2010) *Nature. Genetics.*, **42:** 565.

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:70-73

ANALYSES OF EWE STAYABILITY IN FLOCKS OF NEW ZEALAND SHEEP

M.A. Lee[1], N.G. Cullen[2], S.A. Newman[1], J.C. McEwan[1] and G.H. Shackell[1]

[1]AgResearch, Invermay Research Centre, PB 50034, Mosgiel 9053, New Zealand
[2]AgResearch, Ruakura Research Centre, PB 3123, Hamilton 3240, New Zealand

SUMMARY

A major determinant of profit for sheep farmers is ewe efficiency. A component of efficiency is the length of time a given ewe remains in a flock compared to her contemporaries. A number of terms (e.g. stayability (STAY), productive life and replacement rate) have been used to describe this trait. Breeding to improve this trait may be of significant economic value to New Zealand sheep breeders.

As an adjunct to the development of genomic selection for this trait, a series of quantitative genetics analyses were performed on a large data set derived from industry and research flocks. After quality control, a total of 697,174 animals, from 241 flocks, that were recorded between 1990 and 2009 were available. A subset of the data was analysed based on culling decisions made from the perspective of a commercial farmer or a ram breeder. The results are consistent with a higher risk of culling in ram breeder flocks. The value of STAY as a trait for selective breeding is discussed in view of the analyses.

INTRODUCTION

Productive life and STAY are likely to be important to the profitability of a breeding flock of ewes as these traits affect the costs such as the breeding of replacements (Byrne *et al*. 2012). Breeding to improve these traits may be of significant economic value to New Zealand sheep breeders while also reducing methane emissions per unit of product (Cottle and Conington, 2013). Analyses of STAY have been published using data from a sheep flock managed commercially in the United States (Borg *et al*. 2009) and from a research flock in New Zealand (McIntyre *et al*. 2012). The costs and benefits may be quite different when comparing a ram breeding flock to a commercial flock. A ram breeder has a primary objective of making genetic gain, and the commercial producer aims to maintain productive ewes in the flock for as long as possible. Therefore, ewe culling decisions may be quite different.

As an adjunct to the development of genomic selection for this trait, a series of quantitative genetics analyses were performed on data from industry and research flocks. Some outcomes from these analyses were used to gain insight into the value of STAY as a trait use to breed ewes to increase profitability.

MATERIALS AND METHODS

Data from a total of 4,030,417 animals from ram breeders was used in this study. The breeders were participants in a research program managed by Ovita and recorded on Sheep Improvement Limited. Ewe-records suitable to estimate age of ewe and culling date were identified, where typically the date of the last record for a ewe was the assumed cull date. Flocks with low numbers of animals or minimal recording of traits, ewes moved between flocks over a life-time and records shown as hogget-mating, were ignored. After the data were edited, the analyses focused on 697,174 ewes from 241 flocks born between 1990 and 2009. The mean number of ewes in each flock was 2,893 with a range of 159 to 25,970. Exit codes were assumed to be defined according to information on Sheep Improvement Limited (Walker and Young, 2009). Five flocks (total n=41,317) had sufficient ewe exit code recording, which was culling based on commercial reasons (C, n=8,375), culling on knowledge reasons (K, 10,053), or unknown/missing (U, n=1,592;

missing data=21,297). These five flocks, excluding the animals with unknown/missing records, were used to investigate STAY from the perspective of a commercial producer (cSTAY). Results from one of these flocks have been described previously (McIntyre *et al*. 2012). Linear mixed models were fit with flock and birth year as fixed effects. Traits as analysed were S(3|2) to S(6|2) following (McIntyre *et al*. 2012), where S(3|2) is the probability a ewe will remain in the flock at age 3 given she was present at age 2. Other trait estimated breeding values eBVs (ewe mature weight, number of lambs born and ewe fleece weight) were derived from standard SIL models.

Statistical analyses for linear mixed models were performed in ASREML3 (Gilmour *et al*. 2009) and all other analyses including Kaplan Meier analysis and simulation were undertaken in R (R Development Core Team~~,~~ 2012).

A microsimulation model, where survival was modelled as a Markov process, was developed to begin to assess the value of cSTAY, relative to other traits known to affect ewe profitability. Only animals with eBV accuracies >0.29 for S(3|2) were used in simulations (n=1,917). This model simulated the lives of ewes that had eBVs for a range of traits including traits in Sheep Improvement Limited's terminal sire index and cSTAY, ewe mature weight, number of lambs born and ewe fleece weight. As an animal passes through the model, revenue and costs are calculated. Typically, each animal was simulated for 5,000 iterations and the mean from these iterations used to calculate outcomes (revenues and costs discounted at 8% per annum). The survival of a given year was estimated from a Kaplan Meier function and this was used to estimate mean population transition probabilities. For an animal in a given year the transition probability was the sum of cSTAY and the population mean for that year. Some key assumptions were that revenue from ewes was assumed to be lambs at a value of $100 plus one-half the terminal sire index value calculated from the ewe. Similarly, mean wool weight per ewe (4.8kg), number of lambs born (1.4), and salvage cost of ewe ($2.65/kg carcase weight) were adjusted according to eBVs. Dry matter intake at a cost of $0.12/kg was estimated from NRC equations based on ewe live weight adjusted with an eBV for ewe mature weight. Other costs for ewes included shearing and crutching, animal health costs, and ewe replacement costs (Byrne *et al*. 2012).

RESULTS AND DISCUSSION

Analyses to compare the difference in culling for commercial (C) or ram breeder (K) flocks were performed. For Kaplan Meier survival analysis the status of a ewe from a commercial flock for a given year in her life-time was assumed to be culled if her exit code was C and censored if K, whereas, for a ram breeder flock an exit code of C or K was assumed to be culled. The results from this analysis are given in Figure 1. The results indicate that the survival of a cohort of commercial ewes (cSTAY) and ram breeder ewes (bSTAY) was respectively 37.7% and 8.7% after five years. These observations are consistent with a ram breeder culling policy based on knowledge of ewes such as breeding or index values.

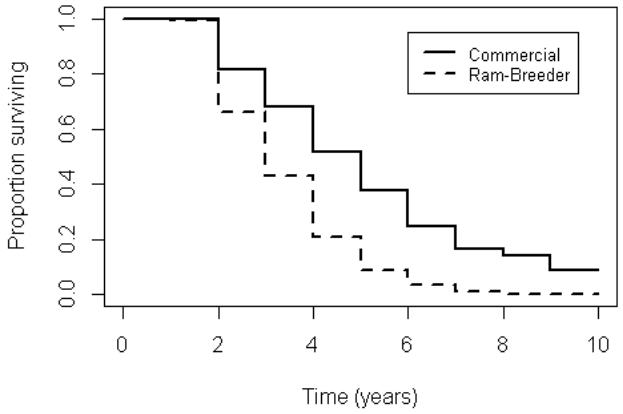


Figure 1 Kaplan Meier survival function of ewe survival in commercial versus Ram breeder flocks.

Linear mixed models were used to investigate genetic parameters. The heritability (standard error) estimated for cSTAY and bSTAY for S(3|2) was 0.048 (0.008) and 0.087 (0.002) respectively, and for S(6|2) 0.082 (0.012) and 0.071 (0.002). The bSTAY estimates are consistent with those from (McIntyre *et al*. 2012). Similar estimates for S(6|2) for cSTAY were described in Borg *et al*., 2009, but for S(3|2) their estimates were zero. Between country differences in policies, for culling ewes after their first mating season, may account for this observation.

The profit for 1,917 animals was calculated using the bioeconomic model described and the distribution of profits given in Figure 2. This is an estimate of the variation in profitability attributable to genetic variance. The results suggest there is significant variation amongst animals with a mean profitability of about $94 and range of -$42 to $272.

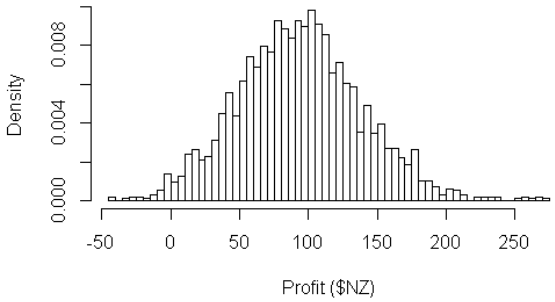


Figure 2 Distribution of live-time profits estimated for ewes. The profit of 1,917 ewes over their lifespan, as a function of estimated breeding values, was estimated by microsimulation. The average profit from 5,000 iterations for each ewe is given as a histogram.

The relative contribution of different traits to profitability was calculated by regressing scaled profit on the scaled trait eBVs S(3|2), S(6|2), ewe mature weight, number of lambs born and ewe fleece weight. The traits S(4|2) & S(5|2) were omitted as they were highly correlated to S(3|2) and

S(6|2) respectively. Economic weights were calculated by varying each trait and calculating the profit attributed to a one unit increase for a given trait. Selection index traits, that take into account the covariance between traits, were calculated by regressing profit on the different trait eBVs. These estimates are given in Table 1. The results suggest that, of the traits analysed and based on the assumptions used in the model, number of lambs born has the most economic value. cSTAY is of more value early in a ewe's life (e.g. the economic weight associated with S(3|2) and S(6|2) was $161.9 and $38.4 respectively). Ewe mature weight contributes negatively to profit through increased feed cost and ewe fleece weight contributed little to profitability. Refinement of this model will enable the calculation of economic and selection index weights to base selection. However, more data on cSTAY may be needed in order to get better estimates of genetic parameters. Moreover it will be of interest to model the value of STAY in different farming environments (Conington *et al*. 2004). Genomic selection may be useful for this trait as it is sex limited, of low heritability, and phenotypic information is recorded late in an individual's life.

Table 1. Estimated economic weightings for maternal traits and relative contributions for each trait

Trait	Economic weight ($/trait unit)	Selection Index weight ($/trait unit)	Relative contribution (%)
S(3|2)	161.90	301.18	16.0
S(6|2)	38.40	170.02	17.6
Number of lambs born	228.20	230.18	64.2
Ewe fleece weight (kg)	24.48	0.59	0.2
Ewe mature weight (kg)	-2.96	-0.38	-2.0

CONCLUSIONS

These analyses suggest that the inclusion of cSTAY in breeding indexes to optimise profitability of ewes in New Zealand will be beneficial and warrants further investigation. However, given that the heritability is low genetic progress will be slow. Further refinement of this model and inclusion of other traits will be needed to better understand its value from a breeding perspective and in relation to other traits.

ACKNOWLEDGMENTS

This work was funded by Ovita. The authors are appreciative to Ram Breeders and industry that have generously contributed their data to allow these analyses.

REFERENCES

Borg R.C., Notter D.R. and Kott R.W. (2009) *J. Anim Sci.* **87:** 3515.
Byrne T.J., Ludemann C.I., Amer P. R. and Young M.J. (2012) *Livestock Science* **144:** 20.
Conington J., Bishop S.C., Waterhouse A. and Simm G. (2004) *J. Anim Sci.* **82**: 1290.
Cottle D.J., and Conington J. (2013) *The Journal of Agricultural Science* **FirstView**: 1-17.
Gilmour A.R., Gogel B.J., Cullis B.R. and Thompson R. (2009) ASREML user guide release 3.0. VSN International Ltd., Hemel Hempstead, UK.
McIntyre S.B., Newman S.A.N., Young E.A. and McEwan J.C. (2012) In: Proceedings of the New Zealand Society of Animal Production. p 152.
R Development Core Team. (2012) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.
Walker G. and Young M.J. (2009) SIL Technical Note - Advanced draft: Ewe longevity (or stayability). Sheep Improvement Limited, New Zealand.

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:74-77

MODIFICATION OF LACTATION YIELD ESTIMATES FOR IMPROVED SELECTION OUTCOMES IN DEVELOPING DAIRY SECTORS

D. M. McGill[1], P. C. Thomson[2], H. A. Mulder[3], and J. J. Lievaart[1]

[1] School of Animal and Veterinary Sciences, Charles Sturt University, NSW 2650, Australia
[2] Faculty of Veterinary Science, The University of Sydney, Camden NSW 2570, Australia
[3] Animal Breeding and Genomics Centre, Wageningen University, The Netherlands

SUMMARY

Animals are extremely important to the livelihood of smallholder dairy farmers in developing countries. However, due to limited resources, investment into long-term genetic improvement is rare and herd recording is minimal. Therefore, obtaining adequate performance records for genetic evaluation and selection is difficult and hence it is important to optimize the selection outcomes from any data that are collected. The aim of this study is to determine a robust and efficient method for estimating lactation yield for low producing dairy cattle and their subsequent genetic evaluation. Using Sahiwal cattle as an example, simulated data sets, based on lactation data from Pakistan, were used to compare different methods of lactation yield estimation (i.e., test-interval method, and three nonlinear models). Furthermore, these estimates were analysed to explore their implications on the subsequent estimated breeding value (EBV) ranking and selection outcomes. Utilising these results, different test-day sampling schedules were compared to investigate possible recording regimes involving few records that can accurately estimate lactation yield without significantly affecting selection. Results indicate that the lactation models proposed by Wood (1967) and Wilmink (1987) yield similar selection outcomes to the recommended test-interval method. These results provide opportunities for further research into test-day scheduling which could reduce the number of records required and have considerable implications on progeny testing systems of low producing dairy cattle and developing dairy sectors.

INTRODUCTION

Breed improvement and selection in developing dairy systems can be challenging as field conditions are generally constrained by a lack of infrastructure for regular test-day recording. For this reason, regular twice daily recording of milk yield for entire lactations is not feasible (Khan *et al.* 2008). The limited resources and data exacerbate the need to utilise each record efficiently to maximize their contribution to the evaluation process (Bajwa *et al.* 2002). Therefore, for any developing dairy sector there is a need to develop a genetic evaluation system which optimizes selection outcomes given the current resources for test-day recording.

There are numerous methods for genetically evaluating milk production based on test-day records. In developed nations complex methods such as test-day models are commonly used. These models require accurate estimates of genetic and phenotypic parameters based on many daily milk yields from large populations of animals which are unlikely to be available in a developing dairy sector (Ilatsia *et al.* 2007). For this reason, simpler methods such as a two-step approach can be used. This is where test-day records are first used to estimate lactation yield and then these values are used as the phenotype for genetic evaluation. Methods of lactation yield estimation from test-day records are well researched. In a developing country scenario, the Test-Interval Method [TIM] (Sargent *et al.* 1968) is recommended by the International Committee for Animal Recording (ICAR 2009). Other approaches involve fitting a mathematical model to lactation data and using the model outputs to estimate yield. Many models have been proposed for describing the lactation curve of dairy animals (Dongre *et a.l* 2011). A handful of studies have investigated the ability of lactation curve models to depict Sahiwal cattle lactation data. Kolte *et*

al. (1986), found that the inverse polynomial function proposed by Nelder (1966) was the superior fitting model, followed by the gamma function proposed by Wood (1967). Contrary to this, Rao and Sundaresan (1979), reported that Wood's (1967) function was the most appropriate. The Wood (1967) model is one of the most widely accepted lactation models and is commonly used in research (Swalve 2000). Similarly, the Wilmink (1987) model is frequently used within test-day evaluations to model the lactation curve of dairy cattle (Naranchuluum *et al.* 2011).

This current study is concerned with Sahiwal cattle in Paktistan and will focus on how different lactation models behave when fitted to the lactation characteristics of this particular breed. Specifically, this study aims firstly to determine which lactation model is the most robust at modelling the lactation curve of Sahiwal cattle at different test-day recording schedules. The second aim is to discuss what implication this may have on the future of test-day sampling in Pakistan and how it can be used to improve their current progeny testing program.

METHODS AND MATERIALS

Lactation Estimation Models and Methods. The lactation estimation methods used within this study were:

1. The test-interval method (TIM) described in Sargent *et al.* (1968) which is based on an approximation of the area under a curve
2. The inverse polynomial model proposed by Nelder (1966): $\mathbf{yield}_i = \frac{1}{a+b*\mathbf{dim}_i+c*\mathbf{dim}_i^2}$
3. The gamma function proposed by Wood (1967) $\mathbf{yield}_i = a * \mathbf{dim}_i^b * \mathrm{e}^{-c*\mathbf{dim}_i}$
4. The lactation model proposed by Wilmink (1987): $\mathbf{yield}_i = a + b^{\mathbf{dim}_i} + c * \mathrm{e}^{-0.05*\mathbf{dim}_i}$

where *a, b* and *c* are different parameters to be estimated separately within each model and **dim** are the days in milk (i = 1, …, 280) for a lactation length of 280 days.

Data. Weekly test-day Sahiwal lactation records from 839 lactations from 464 dams, collected during 2005-2010 from the Livestock Production Research Institute (LPRI), Bahadurnagar Okara, were used as the basis for data simulation in this study. Data were simulated using three different lactation models (Wood, Wilmink and Nelder). Variance and covariance matrices of the parameters (*a, b* and c) and a residual variance of each of these models was determined based on the raw Pakistani data. Using these variance structures and the pedigree relationship matrix (**A**), phenotypic lactation yields were simulated for entire lactations for all the dams in the population. This was repeated 100 times for each of the simulation models to yield three batches of one hundred data sets for comparison. Data were simulated using three different lactation models because it allows for a more thorough comparison of lactation yield estimation methods as it gives an indication of their robustness across different lactation curve shapes.

Model Comparison. Four lactation yield estimation methods were used to fit and calculate the lactation yield for every dam for each set of simulation data. These included the recommended TIM as well as three lactation models, Wood, Wilmink and Nelder, fitted and estimated using the nonlinear mixed effects (NLME) model function in R Version 2.13.0 (R Development Core Team 2008) following a similar process outlined by Raadsma *et al.* (2009). This was carried out for four different test-day scheduling regimes (weekly, monthly, five test-days; random selection and five test-days; stratified selection). For each method, the percentage of models which successfully converged was recorded as well as the lactation yield estimates. The lactation yield estimates were compared with the true simulated lactation yield and summed to calculate a mean square error (MSE) of estimation for each simulated data set. The MSE was then used to directly compare between the lactation yield estimation methods. Lastly, the lactation yield estimates for each simulated data set were used to calculate estimated breeding values (EBVs) for each of the animals

in the data set using ASReml-R Discovery Edition 1.0 (Butler *et al* 2009). The outputs of this analysis allowed further comparison between models to determine if the lactation yield estimation method had any effect on the ranking and subsequent selection of animals.

RESULTS AND DISCUSSION

The robustness of each of the lactation models for fitting Sahiwal test-day data can be determined by comparing the percentage of success rates of each model's ability to be fitted to the different simulated data sets (Table 1). These results show that overall the Wood model is superior to the Wilmink and Nelder models as it generally has higher rates of success, most importantly when fitting data from both a random and stratified selection of five test-day records. This has an important practical implication, as in the field conditions of Pakistan, test-day recording is likely to be irregular and infrequent.

Table 1. Percentage of lactation yield estimation models that were successfully fitted to each set of simulated lactation data at each of the four different test-day recording regimes (weekly, monthly, 5 test-days: random sample and 5 test-days: stratified sample).

Data Simulation Model	Fitted Model	Test-Day Recording Regime			
		Weekly	Monthly	Random	Stratified
Wood	Wood	100	100	82	92
	Wilmink	100	100	76	83
	Nelder	100	100	70	72
Wilmink	Wood	100	100	88	86
	Wilmink	100	100	74	83
	Nelder	100	100	78	75
Nelder	Wood	97	98	60	67
	Wilmink	94	100	69	82
	Nelder	75	71	82	83

Using the MSE values from each of the lactation yield estimation methods we can directly compare between models for the same simulated lactation. The average MSE values across lactation yield estimates can be seen in Table 2. These are presented for only two of the data simulation methods (Wood and Wilmink). The results from the Nelder simulated data are not reported here as the number of failed models caused unreliable values. From Table 2 the MSE values show that the Wilmink and Wood models were superior to the TIM and Nelder methods. Furthermore, the Wilmink model has a lower average MSE than the Wood model in both sets of simulated data (5,124,550 vs 5,327,934 for the Wilmink simulated data and 5,234,436 vs 5,235,715 for the Wood simulated data). This suggests that the Wilmink model is superior to the Wood model in its ability to accurately estimate lactation yield on different types of lactation data.

Despite the differences in the MSE seen in Table 2, the important outcome of this analysis relates to the animals in the top proportion of the population that would be selected for breeding and how they compare with the true (simulated) superior animals. For the different methods of lactation yield estimation, using the Wood simulated data, the average number of corresponding animals with the true top fifty superior animals were; TIM 39.2±2.22, Wood 39.7±2.27 and Wilmink 39.6 ±2.28. For the Wilmink simulated data sets, the results were very similar; TIM 36.8±2.33, Wood 37.1±2.38 and Wilmink 37.1 ±2.41. The results show that the average number of corresponding animals with the true top fifty were all within one animal of the other estimation

methods. This suggests that these methods of estimating lactation yield, for a given test-day scheduling regime, are each capable of selecting the superior animals from a given population.

Table 2: Average Means Squared Error values (± st.dev) for four different methods of lactation estimation (TIM, Nelder, Wilmink and Wood) when calculated using monthly records from two methods of data simulation (the Wilmink and Wood models)

Model used for lactation yield estimate	Model used for data simulation			
	Wilmink		Wood	
	Average MSE (±sd)		Average MSE (±sd)	
TIM	6,102,273	(±385,457.5)	6,143,607	(±377,174.1)
Nelder	5,962,774	(±413,556.1)	5,696,461	(±387,866.5)
Wilmink	5,124,550	(±311,253.8)	5,234,436	(±347,630.8)
Wood	5,327,934	(±342,708.3)	5,235,715	(±368,825.4)

CONCLUSIONS

The benefit of modelling test-day yields is the ability to subsequently estimate lactation yield on fewer records. This then provides an opportunity to record more animals fewer times which will help to improve the accuracy of evaluations as well as increase the population of animals from which selection can take place. The outcomes of this study show that although the Nelder model is capable of fitting and modelling low producing dairy cattle like the Sahiwal, it is unreliable with different lactation curves and test-day sampling regimes. The results from the other lactation models tested, the Wood or Wilmink, show that they are both robust in different scenarios with the Wood model better fitting irregular and infrequent test-day recording regimes. Despite this, both the Wood and Wilmink models provide an opportunity to further investigate their use in estimating lactation yield in Sahiwal cattle and the possibility of reducing the number of required test-day records whilst maintaining the accuracy of selection outcomes.

REFERENCES

Bajwa I. R., Khan M.S., Ahmad Z. and Gondal K.Z.. 2002. Pages 1-56 in Proc. 7th World Congr. Genet. Appl. Livest. Prod. , Montpellier, France.

Butler D. G., Cullis B.R., Gilmour A.R. and B. J. Gogel . 2009. Department of Primary Industries and Fisheries, Queensland.

Dongre V. B., Gandhi R.S., Avtar S., and Atul G. 2011. *Intl. J. Agric. Res. and Rev.* **1:** 6

ICAR. 2009. International Committee for Animal Recording, Niagra Falls, USA.

Ilatsia E. D., Muasya T.K., Muhuyi W.B., and . Kahi A. K. 2007. *Animal* **1**:185

Khan, M. S., Bilal G., Bajwa I.R., Rehman Z. and Ahmad S. 2008. *Pakistan Veterinary J.* **28**:131

Kolte D. V., Gore A.K., and Deshmukh S.N. 1986. *PKV Research Journal* **10**:145.

Naranchuluum, G., Ohmiya H., Masuda Y., Hagiya K., and Suzuki M. 2011. *Anim. Sci.***82:** 383

Nelder J. A. 1966. *Biometrics* **22**:128.

R Development Core Team 2008. R. F. f. S. Computing, ed, Vienna, Austria.

Raadsma H. W., Jonas E., McGill D., Hobbs M., Lam M.K., and Thomson P.C. 2009. *Genet. Sel. Evol.* **41**

Rao M. K. and Sundaresan D. 1979. *J. of Agric. Sci.* **92**:393

Sargent F. D., Lytton V.H., and Wall O.G. Jr. 1968. *J. Dairy Sci.* **51**:170

Swalve H. H. 2000. *J. Dairy Sci.* **83**:1115

Wilmink J. B. M. 1987. *Livest. Prod. Sci.* **17**:1

Wood P. D. P. 1967. *Nature* **216**:164

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:78-81

INCREASING PROLIFICACY OF AWASSI AND ASSAF BREEDS BY INTROGRESSION OF THE *FECB* (BOOROOLA) MUTATION: ACHIEVEMENTS AND CHALLENGES

E. Gootwine[1], A. Rosov[1], M. Abu Siam[2] and E. Seroussi[1]

[1]The Volcani Center, PO Box 6, Bet Dagan 50250, Israel
[2] Siam Veterinary Clinic, PO Box 519, Rahat 85357, Israel

SUMMARY

Awassi and Assaf are local Israeli sheep breeds with prolificacies of ~1.30 and ~1.65 lambs born/ewe lambing (LB/EL), respectively. Introgressing the *B* (Booroola) allele of the *FecB* locus into these breeds led to the formation of prolific strains designated 'Afec Awassi' and 'Afec Assaf', in which lamb production under both intensive and semi-intensive conditions is higher than in the respective local breeds by ~0.8 LB/EL and ~0.5 live lambs born/ewe lambing. Lower survival rate at birth in multifetal pregnancies reduces the ability to fully exploit the economic potential of the Afec strains. A genome-wide association study revealed QTLs on ovine chromosomes 1, 8, 10, 26 and X associated with lamb survival rate at birth as a maternal trait.

INTRODUCTION

Since the beginning of the last century, the Awassi—a low-prolificacy fat-tail sheep breed and the most common breed in the Middle East—has undergone consecutive genetic changes in Israel aimed at improving milk and lamb production (Gootwine 2011; Fig. 1). Within-breed selection for high milk production resulted in the formation of the Improved Awassi dairy strain. Later, crossing the Improved Awassi with the East Friesian breed led to the formation of the Assaf. Today, sheep production in Israel (about 0.5 million head) is managed under a wide range of conditions—from extensive production where the local Awassi is kept for lamb production to the highly intensive dairy and non-dairy flocks where Assaf and Awassi sheep are managed.

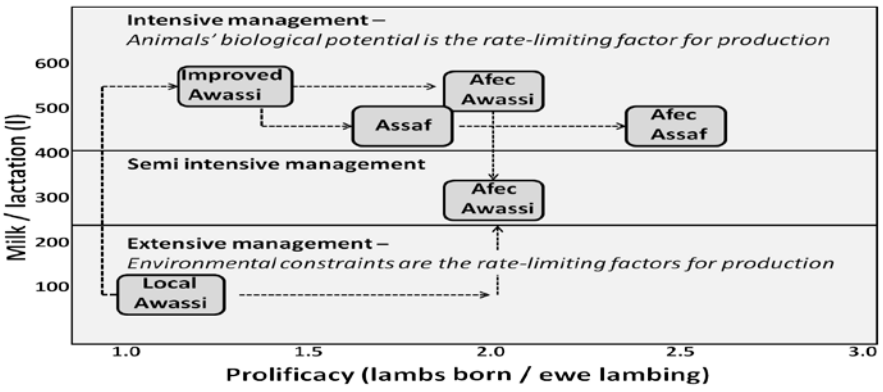


Figure 1. Schematic presentation of the breeding history of Awassi sheep in Israel

DEVELOPING THE AFEC STRAINS

Lamb production has been an important source of income in non-dairy and dairy flocks. The average prolificacy of the Awassi and Assaf is about 1.30 and 1.65 lamb born/ewe lambing (LB/EL), respectively. In 1986, the Volcani Center launched breeding programs to increase the prolificacy of the Improved Awassi and the Assaf by introgression of the *B* (Booroola) allele of the

FecB locus (Piper *et al.* 1985). Five homozygous *BB* Booroola Merino rams obtained from the Invermay Agricultural Centre, New Zealand, served as the source for the mutation. Through the backcrossing and intercrossing phases, selection for carriers of the Booroola mutation was carried out first by monitoring induced ovulation rate in ewe lambs and later, by direct genotyping for the *FecB* locus (Wilson *et al.* 2001). The breeding activity resulted in the formation of highly-prolific strains designated Afec Awassi and Afec Assaf, which carry the Booroola mutation and have average prolificacies of about 1.9 and 2.5 LB/EL, respectively (Gootwine *et al.* 2008).

The improved Awassi and Assaf fat-tail dairy breeds diverge a great deal from the non-dairy, small-body-size Booroola Merino breed. The genetic backgrounds of Afec sheep and Awassi and Assaf sheep (n = 176) were compared in 2012, using the 50K ovine single-nucleotide polymorphism (SNP) beadchip (Illumina). Results showed that throughout the introgression process, the original genetic background of the local breeds was retained almost completely with a main selection signature on ovine chromosome 6, where the *FecB* locus is mapped (Fig. 2).

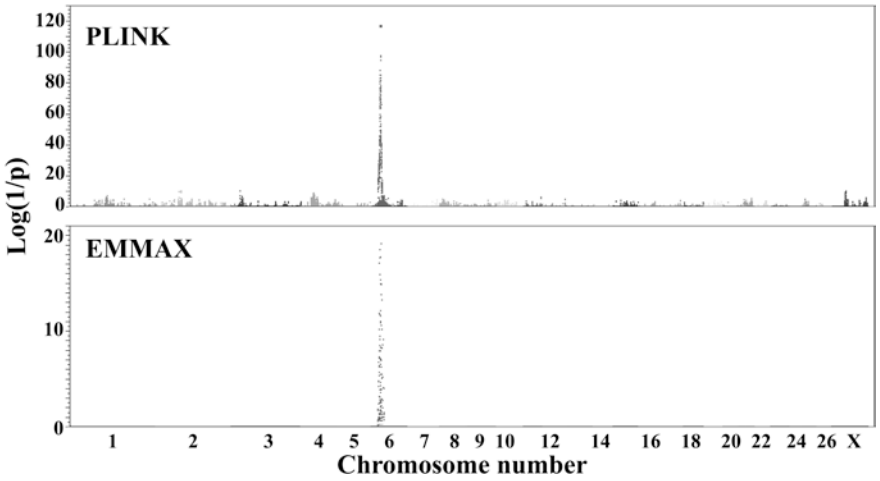


Figure 2. Association analysis between SNP haplotypes and carrying the Booroola mutation in local Israeli sheep breeds. Using the 50K ovine SNP beadchip (Illumina), a selection signature was observed by haplotype analysis of a 10-SNP sliding window using the PLINK package (Purcell *et al.* 2007) and by SNP association using EMMAX software (Kang *et al.* 2010), the latter correcting for population relatedness. While the most significant association ($p < 10^{-116}$) detected by the PLINK analysis pointed to the haplotype spanning the *FecB* locus, EMMAX detected the most significant ($p < 10^{-23}$) SNP 2.3 Mb telomeric to the mutation.

Dissemination of the Afec sheep from the breeding nuclei at the Volcani Center and the Kibbutz Ein Harod Awassi flock has been achieved mainly by selling homozygous *BB* rams to mostly non-dairy intensive commercial flocks where animals are fed to meet all of their metabolic needs. The desired genotype at the *FecB* locus for Afec ewes is *B+*, as homozygous *BB* ewes bear some disadvantages in terms of prolificacy and growth (Gootwine *et al.* 2008). Selection of replacements in the commercial flocks is based on genotyping for the *FecB* locus and in recent years, about 4,000 genotypings for *FecB* have been carried out annually with 0.51 of the genotypes being *B+* (unpublished results). Introduction of the Afec strains to commercial flocks was followed by implementation of managerial means to support the maintenance of highly prolific sheep, including a new treatment for pregnancy toxemia (Zamir *et al.* 2009).

Introgression of the Booroola mutation into local Awassi flocks. About half of the national sheep flock in Israel belongs to the local Awassi and is kept by Bedouin farmers in the Negev—the arid southern part of the country—under traditional semi-extensive management, where animals rely for about half the year on seasonal pasture. Decreases in recent years in the availability of grazing land have forced Bedouin growers to spend more on feeding their animals by purchasing costly grains and fodder, making sheep production nearly unprofitable. To overcome the new economic constraints, we investigated improving local Awassi flocks' productivity by introducing the Afec-Awassi strain. The question arose as to how the Bedouin farmers would be able to change their traditional management to support highly prolific ewes. Since 2007, controlled dissemination of the *FecB* mutation in Bedouins' Awassi flocks has been carried out by distributing *BB* Afec-Awassi rams. It is estimated that in 2013, about 20,000 Afec-Awassi sheep will be successfully bred by Bedouin farmers who appreciate the economic advantage of the genotype (Fig. 3).

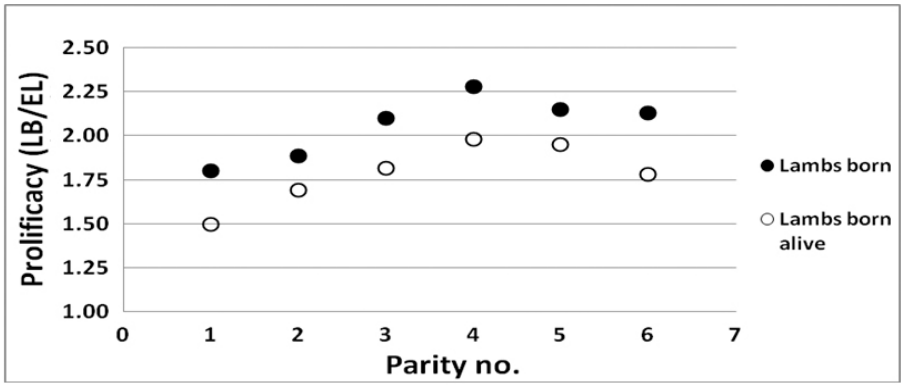


Figure 3. Prolificacy of *B*+ Afec-Awassi ewes (n = 930) in Bedouin flocks in the Negev, Israel, according to parity number. Average prolificacy of mature local Awassi ewes was estimated to be 1.14 lambs born per ewe lambing (LB/EL) based on 4,692 lambing records.

RESEARCH INTO LAMB SURVIVAL RATE AT BIRTH

Larger litter size is associated with lower lamb survival rate at birth (LSRAB), which drops from about 0.95 in the case of singletons to about 0.5 in the birth of sixplets (Gootwine *et al.* 2008). The full economic advantage of high litter size at birth is not captured because of lower LSRAB in the Afec strains, as litters of 4 or more lambs comprise about 15% of the litters in the Afec Assaf (Gootwine *et al*. 2008).

Multifetal pregnancy affects maternal metabolism (Moallem *et al*., 2012) and fetal body weight in a manner comparable to the adverse effects of severe experimental protocols aimed at restricting fetal growth such as maternal undernutrition or carunclectomy (Gootwine 2013). Research into morphometric parameters of newborn Afec-Assaf lambs (n = 957) which account for the effects of crop, sire, litter size, parity number, sex and lamb viability at birth showed that while liveborn and stillborn lambs were similar in their crown rump length, being on average 51.4 ± 0.4 cm, stillborns were significantly (p<0.0001) lower in birth weight (4.1 ± 0.1 and 3.5 ± 0.1 kg for liveborns and stillborns, respectively). This indicates that fetal death in multifetal pregnancies occurs on average some 7–10 days before lambing.

LSRAB can be considered both a maternal and a fetal trait. To investigate the effect of the maternal genome on LSRAB, a whole-genome association analysis utilizing the ovine 50K beadchip (Illumina) was performed on 71 ewes with an average prolificacy of 3.04 LB/EL (4–8 parity records) and with LSRAB values ranging from 0.00 to 0.95. EMMAX (Kang *et al.* 2010) and PLINK (Purcell *et al*. 2007) haplotype analyses indicated a total of 14 regions on chromosomes 1, 8, 10, 26 and X associated ($p < 0.05$) with LSRAB (Fig. 4).

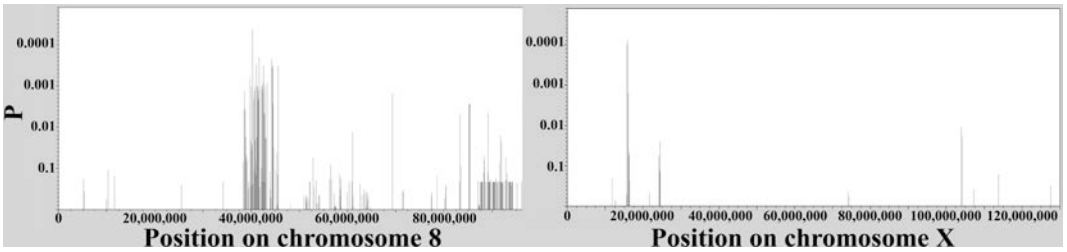


Figure 4. Chromosomal regions associated with lamb survival rate at birth in Afec-Assaf ewes. Association analysis between 10 SNP haplotypes and LSRAB was carried out using the PLINK package (Purcell *et al.* 2007). Probabilities were corrected for multiple comparisons following Bonferroni.

CONCLUSIONS

Introgression of the Booroola mutation into Awassi and Assaf breeds is a relatively fast way to increase lamb production while retaining phenotypic characteristics and important production traits, such as high milk production, large body size and adaptability to local conditions. Further research into the genetic control of LSRAB in sheep as either a maternal or fetal trait may contribute to an improvement in the economic benefits of breeding prolific sheep.

ACKNOWLEDGEMENTS

Invermay Agricultural Research Centre, New Zealand is acknowledged for supplying Booroola rams. Prof. A. Valle Zárate and Dr. A. Al Baqain, University of Hohenheim, Germany, are acknowledged for their support of the Afec-Awassi work in Bedouin flocks, and Alon Lam is acknowledged for his contribution to the genome-wide association study on lamb survival.

REFERENCES

Gootwine E. (2011) *Trop. Anim. Health Prod.* **43**: 1289.

Gootwine E. (2013) *J. Anim. Sci.* **91**: 111.

Gootwine E., Reicher S. and Rozov A. (2008) *Anim. Reprod. Sci.* **108**: 402.

Kang H.M., Sul J.H., Service S.K., Zaitlen N.A., Kong S., Freimer N.B., Sabatti C. and Eskin E. (2010) *Nat. Genet.* **42**: 348.

Moallem U., Rozov A., Gootwine E. and Honig H. (2012) *J. Anim. Sci.* **90**: 318.

Piper L.R., Bindon B.M. and Davis G.H. (1985) In 'Genetics of reproduction in sheep'. pp. 115-125, editors R.B. Land and D.W. Robinson, Butterworths, London.

Purcell S., Neale B., Todd-Brown K., Thomas L., Ferreira M.A.R., Bender D., Maller J., Sklar P., de Bakker P.I.W., Daly M.J. and Sham P.C. (2007) *Am. J. Hum. Genet.* **81**: 559.

Wilson T., Wu X.Y., Juengel J.L., Ross I.K., Lumsden J.M., Lord E.A., Dodds K.G., Walling G.A., McEwan J.C., O'Connell A.R., McNatty K.P. and Montgomery G.W. (2001) *Biol. Reprod.* **64**, 1225.

Zamir S., Rozov A. and Gootwine E. (2009) *Vet. Rec.***165**: 265.

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:82-85

YEARLING AND ADULT EXPRESSIONS OF REPRODUCTION IN MATERNAL SHEEP BREEDS ARE GENETICALLY DIFFERENT TRAITS

K.L. Bunter and D.J. Brown

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

SUMMARY

Reproductive data (N=19335 ewes recorded) from maternal breeds recorded in industry flocks were used to estimate genetic parameters for the number of lambs born (NLB) and weaned (NLW) per ewe joined, along with their component traits fertility (FERT), litter size (LSIZE) and lamb survival (LSURV). Data were analysed as different traits for ewes in different age groups (yearlings, two year olds, and 2+ year olds). Yearling performance was characterised by low FERT (54%), low LSIZE, reflecting an increased frequency of single births, and increased lamb losses relative to older ewes bred in the same flock-years. Heritability (h^2) estimates were highest for yearling FERT (h^2=0.16) and declined for this trait with ewe age group (h^2~0.07). In contrast, heritabilities and variance increased with ewe age for LSIZE (h^2: 0.05 to 0.11). Genetic correlations (rg) between yearling and later records within traits were significantly <1 (range 0.10 to 0.54). The exception was LSIZE where the genetic correlation between ewe age groups was consistently high (rg: 0.85 to 1.0). Trait values affected by fertility outcomes (FERT, NLB and NLW) had significant service sire effects, whereas service sire effects were insignificant for LSIZE and LSURV. Service sire recording was incomplete more frequently for infertile ewes. Yearling reproductive performance should be treated as genetically different to adult expressions of the same traits for genetic evaluation purposes, and the different genetic architecture of component traits towards NLB and NLW can then be appropriately accommodated.

INTRODUCTION

Reproductive performance of ewes joined to lamb as yearlings is low relative to that of maiden two-year old or multiparous mature ewes (Afolayan *et al.* 2008). Both lower fertility and prolificacy (litter size), and fewer offspring weaned per ewe lambing, are characteristic outcomes from the first joining of polyovulatory species such as sheep and pigs. This reflects variation amongst individuals in attributes like age at puberty and sufficient expression of behavioural estrus, adequate weight or body condition pre-breeding, along with differences in ovulation rate, foetal survival and pre-weaning survival. Such characteristics are all under genetic control to varying degrees. Some genetic evaluation systems treat first parity performance as a genetically different trait to performance in later parities for sheep (SIL, Walker 2008) and pigs (PIGBLUP, Crump and Henzell, 2000). Since 2012, Sheep Genetics (Brown *et al*. 2007) has also analysed yearling number of lambs born (NLB) and weaned (NLW) per ewe joined separately to the same traits recorded for older ewes. The aim of this study was to estimate genetic correlations using industry data for NLB and NLW, along with the component traits of fertility (FERT), litter size (LSIZE) and lamb survival (LSURV), when considered as different traits for ewes in different age groups. A secondary goal was to examine the importance of service sire effects for these traits.

MATERIAL AND METHODS

Reproductive data were derived from industry records submitted to Sheep Genetics. The data subset analysed included only those flocks and years where significant numbers of yearling ewes

* AGBU is a joint unit of NSW DPI and the University of New England

were recorded. The resulting data (N=19335 ewes) represented nineteen flocks and five breeds (Border Leicester, Coopworth, White Suffolk and two Maternal synthetic flocks). Most flocks had few years in which yearling ewes were recorded. Flock-years with 100% fertile ewes were excluded from analyses as these reflected incomplete recording. All yearling ewes were naturally joined whereas older ewes were either naturally joined or bred using artificial insemination (AI). A complete pedigree for the combined breeds was used for parameter estimation (N=1240091).

Traits were defined by ewe age group (yearlings, two year olds, and 2+ year olds) and included fertility of joined ewes (FERT: pregnant or not, 1/0), fecundity (LSIZE: lambs born per ewe lambing), lamb survival (LSURV: lambs weaned per ewe lambing) and number of lambs born (NLB) or weaned (NLW) per ewe joined. The fixed effect models for the yearling traits accounted for flock-year of lambing (45 levels), dam age group (6 levels: 1yo, 2yo, 3yo, 4yo, 5-8yo, 8+ yo), ewe age (in days) and month of birth fitted as linear and quadratic covariates, along with service sire age group (5 levels: 1yo, 2yo, 3-7yo, 7+ yo, and unknown). Fixed effect models for older ewes included flock-year combined with conception method (2 levels: natural or AI), along with the linear effect of ewe age and the service sire age group, as described above. Breed was confounded with flock and was not explicitly fitted in models for analysis. Ewe parity at joining was not fitted in models for two or 2+ year old ewes. Parameters were estimated for all traits under linear animal models, treating each ewe as the animal and the service sire at joining as an additional random effect (i.i.d). Ewes had only one record used per ewe age group (repeated records in the 2+ age group were not used due to low N). Estimates of heritabilities and genetic correlations were obtained under an animal model using ASREML (Gilmour et al. 2009) from a series of univariate and bivariate analyses. Where the service sire effect was only marginally significant in univariate analyses, it was removed from models for the relevant trait(s) in bivariate analyses. No covariance between service sire effects was fitted.

RESULTS AND DISCUSSION

Raw data means for each trait by ewe age group show lower fertility, prolificacy, lamb survival and therefore NLB and NLW for yearling compared to older ewes recorded in the same flock-years (Table 1). The larger trait standard deviation shows that fertility was more variable between flock-years for yearling relative to older ewes. In contrast, LSIZE was less variable for yearling ewes, resulting from a smaller range in trait values and a relatively high frequency of single births for yearling ewes (not presented). Because of low fertility, the number of yearling ewes with subsequent records for LSIZE and LSURV was low.

Table 1. Means, standard deviations (SD) and record count, by ewe age group (Y: yearling; 2yo: two-year old; 2+: older than 2 years old), for fertility (FERT), litter size (LIZE) and lamb survival (LSURV), along with lambs born (NLB) or weaned (NLW) per ewe joined

	Mean (SD)			Counts of records				
	Yearling	2yo	2+	Yearling	2yo	2+	Y/2yo	Y/2+
FERT	0.54 (0.50)	0.91 (0.28)	0.94 (0.23)	12153	9315	6657	4931	2405
LSIZE	1.40 (0.51)	1.62 (0.58)	1.74 (0.63)	6548	8487	6313	2253	1122
LSURV	1.08 (0.62)	1.39 (0.63)	1.52 (0.66)	6544	8485	6290	2220	1208
NLB	0.75 (0.79)	1.47 (0.72)	1.65 (0.73)	12153	9315	6657	4931	2405
NLW	0.58 (0.70)	1.26 (0.72)	1.44 (0.73)	12153	9315	6657	4931	2405

Estimates of heritability and service sire effects differed with ewe age class (Table 2). Fertility was most heritable (h^2=0.16) and service sire variance (s^2=0.23) was largest for yearling ewes, with both parameter estimates (h^2~0.07, s^2~0.06) and phenotypic variation decreasing in

magnitude for FERT with increasing ewe age. In contrast, heritability and variability in LSIZE increased with ewe age. Around 15% of yearling ewes lambed and lost, increasing the phenotypic variance for LSURV relative to LSIZE. However, this effect was smaller for older ewes where only 6% failed to rear any lambs. Within older ewe age groups, the heritability estimates for LSURV were typically lower than the corresponding estimates for LSIZE and residual variance was increased. Poor fertility and an increased incidence of lamb deaths for yearlings also affected parameters and variances for NLB and NLW, relative to LSIZE and LSURV, because these trait distributions become zero enriched when either fertility is low or a significant proportion of ewes rear no lambs. Genetic correlations between yearling and adult performance for FERT, LSURV, NLB and NLW were significantly less than unity, supporting the concept that genetically yearling reproductive performance differs from adult expressions for the same traits. However, genetic correlations between ewe age groups for LSIZE did not differ from unity (rg: 0.85 - 1.0).

Table 2. Estimates of heritability and the proportion of service sire effects, along with genetic correlations, by ewe age group (Y: yearling; 2yo: two-year old; 2+: older than 2 years old), for fertility (FERT), litter size (LIZE), lamb survival (LSURV), and lambs born (NLB) or weaned (NLW) per ewe joined

Parameter	Age1*	Age2*	FERT	LSIZE	LSURV	NLB	NLW
Heritability	Y		0.16±0.02	0.05±0.02	0.07±0.02	0.13±0.02	0.08±0.01
	2yo		0.07±0.02	0.10±0.02	0.07±0.02	0.10±0.02	0.08±0.02
	2+		0.07±0.02	0.11±0.02	0.06±0.02	0.09±0.02	0.06±0.02
Service sire	Y		0.23±0.02	0.02±0.01	0.02±0.01	0.11±0.01	0.14±0.02
effect	2yo		0.16±0.02	0.01±0.01	0.01±0.01	0.05±0.01	0.04±0.01
	2+		0.06±0.01	0.01±0.01	B	0.02±0.01	0.01±0.01
Phenotypic	Y		0.15	0.25	0.36	0.39	0.35
Variance	2yo		0.073	0.31	0.38	0.45	0.47
	2+		0.044	0.36	0.42	0.46	0.49
Genetic	1yo	2yo	0.44±0.12	1.0±0.21	0.38±0.20	0.33±0.11	0.40±0.14
Correlation	1yo	2+	0.10±0.15	1.0±0.18	0.54±0.24	0.33±0.15	0.42±0.20
	2yo	2+	0.28±0.21	0.85±0.12	0.65±0.22	0.91±0.12	0.95±0.21

Age*: for trait 1 or traits 1 & 2 (univariate vs bivariate analyses); B: converged to zero boundary

Few studies have reported parameter estimates for reproductive traits of ewes recorded in different age classes. Heritability estimates from combined parity data for naturally joined crossbred ewes bred in the Maternal Central Test project were 0.11±0.04 for FERT, 0.19±0.05 for LSIZE, 0.03±0.02 for LSURV, 0.17±0.04 for NLB and 0.11±0.04 for NLW (Afolayan *et al*. 2008), consistent generally with estimates from this study. Comparable estimates from combined parity Merino data tend to be lower (Safari *et al*. 2007). Newton *et al.* (2013) reported heritabilities of 0.20±0.05 and 0.16±0.05 for yearling NLB and NLW, recorded on maternal-cross ewes in the Sheep CRC INF flock. Service sire effects accounted for 21, 17 and 8% of variation in FERT, NLB and NLW in Safari *et al*. (2007); a similar pattern was observed in this study. Since service sire effects were negligible for LSIZE or LSURV, this suggests that for NLB and NLW, some service sire variation arose from an auto association between the incidence of ewe infertility and the reporting of a service sire as unknown. The percentages of records without service sires reported was 17.7% in yearling data, compared to 3.7% and 3.9% of records for older ewes. This reduced to <3% of unknown service sires for lambed ewes in any age group. However, this phenomenon would not have influenced the comparable results of Safari *et al.* (2007). In addition, since each service sire defined a joining group, other factors could also contribute to estimates of service sire variation for FERT – for example group size and paddock attributes. Approximately

50% of the variation attributed to the service sire effect for fertility can be removed by fitting additional fixed effects in models for analyses, such as joining group size (results not presented). These results suggest that the implications of service sire effects for the accurate genetic evaluation of ewe fertility should be investigated further with more complete recording of service sires for all ewes joined. The relative magnitude of service sire effects is expected to be lower for NLB and NLW because of contributions to these trait values from LSIZE and LSURV, which are unaffected by service sire effects.

Low fertility of adult ewes typically reflects a service sire failure rather than genetic inferiority of all of the ewes for fertility *per se*. Consequently, low fertility groups of ewes are typically removed from the Sheep Genetics genetic evaluation system to reduce the possibility of bias introduced by service sire failures. However, no such editing was applied to this data on the basis of yearling performance levels because in this age group low fertility of the group joined does not necessarily represent service sire failure. The extent to which genetic correlations between parities are influenced by the threshold for fertility applied to edit data needs to be examined further for yearling ewes. Service sire effects were not important for LSIZE or LSURV in these data. Genetic evaluation for these traits might be more accurate than for the compound traits of NLB or NLW when service sires are not fully reported in Industry data.

CONCLUSIONS

The relative contributions of component traits such as fertility, litter size and lamb survival to trait values for NLB and NLW varies with ewe age. This is accompanied by differences in heritabilities and phenotypic variances for fertility in particular, where performance differences between yearling and older ewes are large. Relatively low genetic correlations indicate that reproductive traits of yearling ewes should generally be treated as genetically different traits to the same traits recorded on older ewes, with the exception of LSIZE. The contribution of service sire effects to variation in reproductive performance warrants further investigation for fertility traits in particular, since this will also influence NLB and NLW. Current parameter estimates for service sire effects using Industry data may be partially influenced by incomplete recording of service sires more often when ewes are infertile, or might be eliminated by optimising management of yearling ewes at joining.

ACKNOWLEDGEMENTS

This research was funded by the Meat and Livestock Australia. The authors gratefully acknowledge the contribution of breeders to the data used in this study.

REFERENCES

Afolayan R.A., Fogarty N.M., Gilmour A.R., Ingham V.M., Gaunt G.M. and Cummins L.J. (2008). *J. Anim. Sci.* **86**: 804.

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.

Crump R. and Henzell A. (2000). In 'Pig Genetics Workshop' pp 1-4, editor Ron Crump, AGBU, UNE, Australia

Gilmour A.R., Gogel B.J., Cullis B.R. and Thompson R. (2009). ASREML User Guide Release 3.0. VSN International Ltd, Hemel Hempsted, HA1 1ES, UK.

Newton J.E., Brown D.J., Dominik S. and van der Werf J.H.J. (2013). *Proc. Assoc. Advmt. Anim. Breed. Genet.* **20**: (in press)

Safari E., Fogarty N.M., Gilmour A.R., Atkins K.D., Mortimer S.I., Swan A.A., Brien F.D., Greef J.C. and van der Werf J.H.J. (2007). *Aust. J. Agric. Res.* **58**: 177.

Walker G. (2008). SIL Technical Note. Beef + lamb New Zealand.

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:86-89

AGE AT FIRST OESTRUS. A USEFUL TRAIT FOR EARLY REPRODUCTIVE PERFORMANCE?

J.E. Newton[1,2,3], D.J. Brown[2], S. Dominik[3] and J.H.J. van der Werf[1,4]

[1]School of Environmental and Rural Science, University of New England, Armidale, NSW, 2351
[2]Animal Genetics and Breeding Unit, Armidale, NSW, 2351
[3]CSIRO Animal, Food and Health Sciences, FD McMaster Laboratory, Armidale, NSW, 2350
[4]Cooperative Research Centre for Sheep Industry Innovation, Armidale, NSW, 2351

SUMMARY

An increasing number of Australian sheep breeders are joining ewes at 6-8 months of age, which is 6-12 months earlier than ewes are traditionally first joined. When joining at a young age, additional factors such as the attainment of sexual maturity must be considered. The age of sexual maturity is a challenging trait to measure with limited data currently available in sheep. This study explored several methods of analyzing age of first oestrus (AFO) data, an indicator trait for sexual maturity, and explored the relationship between AFO and early reproductive performance. Lambing records from 2218 Maternal-cross ewes joined naturally at 7-10 months were used, a subset of 906 ewes had AFO information collected through the use of teaser wethers. Heritability estimates for AFO were low (0.03 – 0.09) whilst estimates for number of lambs born and weaned at yearling age were 0.20 and 0.16 respectively. Genetic correlation between AFO and number of lambs born and weaned at yearling age were 0.45 and 0.51, respectively, but had high standard errors. Improving reproductive performance through the use of teasers to record AFO is not recommended, thus a need exists to find reliable measures for early reproductive traits including sexual maturity.

INTRODUCTION

Under traditional sheep operations ewes are first mated at 18 months of age. However there is increased uptake from industry in joining ewes at 6-9 months of age. Early joining of ewes is currently characterized by highly variable success rates (Fogarty *et al.* 2007) and the underlying causes of poor reproductive rates in yearling ewes have not been fully quantified. Age of first oestrus (AFO) is important as an indicator for the identification of earlier maturing animals. Measuring pubertal traits is a challenge that was tackled in the Australian beef industry with a high degree of success. Moderate-highly heritable traits for pubertal development in tropical cattle such as mean age at first corpus luteum have been identified (Johnston *et al.* 2009). This study aims to identify factors influencing AFO and yearling reproductive performance and to estimate the relationship between these traits.

MATERIALS AND METHODS

Experimental Design. Information on AFO was collected on 906 Maternal-cross ewes as part of the Sheep CRC Information Nucleus Flock (van der Werf *et al.* 2010) data collection. Ewes were run with harnessed teaser wethers after weaning until joining in 2008, 2009 and 2010. The timing of exposure to teasers post-weaning and the length of time of exposure to teasers differed between years and sites and was not recorded. AFO (in days) was recorded when ewes were first marked by a raddle worn by a teaser, raddle marks were checked weekly. Live weight was also recorded when AFO was recorded. For animals missing this live weight, the live weight closest to this date (usually within 14 days) was used in the analysis. Maternal sire breeds represented were: Border Leicester, Corriedale, Bond, Coopworth, East Friesian, Booroola, PRIME SAMM and

Dohne. Lambing data was recorded for 2218 Maternal-cross ewes, naturally mated at 7-10 months from 2008-2012. All animals were assigned a contemporary group based on site and year of birth. All ewes that were alive at one year of age that were within a contemporary group that was exposed to teasers and/or joined were assumed to have also been exposed to teasers and had the opportunity to join to the ram. Ewes that met this criterion and had no lambs were assumed to have been dry for yearling number of lambs born (YNLB) and yearling number of lambs weaned (YNLW). A synthetic AFO trait was created (AFO2) for which ewes exposed to the teasers, but not marked were assigned an AFO measurement equal to the maximum for the contemporary group plus 17 days, which represented a penalty equivalent to one full oestrous cycle. In total 621 maternal crossbred ewes that had no AFO measurement were assigned a value for AFO2 according to this method.

Statistical Analysis. Traits were first analysed using animal models in univariate analyses in ASReml (Gilmour *et al. 2009*). Animal and genetic groups were fitted as random effects. Genetic groups reflected breed proportions, thus sire breed was not fitted in models. Models fitted for all traits included contemporary group (flock/drop) and age of dam (2-8 and unknown). AFO models were run with and without weight at first oestrus fitted as a covariate. Models for YNLB and YNLW were tested for the covariates age at joining and joining length. Bivariate analysis was then conducted to estimate phenotypic and genetic correlations. Due to low record numbers, a unique pedigree was generated for each analysis. The number of records retained for analysis, number of sires, number of genetic groups and number of animals in the pedigree is outlined in Table 1.

RESULTS AND DISCUSSION

Data Summaries and Fixed Effect Analyses. A summary of the traits analysed and the covariates fitted is given in Table 1. Only 40% of young ewes joined to lamb as yearlings gave birth to a lamb, whilst only 30% of these ewes weaned a lamb. Joining length ranged from 21-56 days whilst age at commencement of joining varied from 189-273 days.

Table 1. Summary of reproductive traits analysed in INF Maternal-Cross Ewes: Age at first oestrus (AFO) in days and a synthetic version of the trait (AFO2), Yearling no. lambs born (YNLB) and weaned (YNLW)

Trait	No. of Records	No. of Sires	No. in Pedigree	No. genetic groups	Mean *(st. dev)*	Range
AFO *(days)*	906	90	5007	47	202.8 *(36.5)*	130- 282
AFO2 *(days)*	1527	91	6587	51	220.3 *(38.4)*	130-299
YNLB *(no. lambs)*	2218	116	8062	54	0.40 *(0.65)*	0-4
YNLW *(no. lambs)*	2218	116	8062	54	0.31 *(0.59)*	0-4
Weight at first oestrus *(kg)*	906	-	-	-	40.8 *(8.61)*	19.4-67.0
Age at joining *(days)*	2218	-	-	-	230.6 *(19.1)*	189-273
Length of Joining *(days)*	2218	-	-	-	38.8 *(6.6)*	21-56

For all traits contemporary group was significant ($P<0.01$) whilst age of dam was not significant. Despite a difference in joining length of 35 days (Table 1.), the length of time the ewes were left with the rams was not significant for either YNLB or YNLW, most likely because this was accounted for by contemporary group. Joining length and joining age were excluded from the final models for YNLB and YNLW. Where fitted weight at AFO was highly significant ($P<0.01$).

Univariate and Bivariate Analyses. Heritability estimates for all traits were low to moderate, which was expected given the low heritability of reproductive traits in sheep reported elsewhere

(Safari and Fogarty 2003) (Table 2). Heritability estimates for both variations of AFO did not differ significantly with and without weight fitted as a covariate, 0.04 ± 0.07 and 0.03 ± 0.07 for AFO and 0.04 ± 0.07 and 0.09 ± 0.05 for AFO2. Although lower, this is similar to preliminary heritability estimates for the same trait in the New Zealand Central Progeny Test (CPT) flock of 0.09 ± 0.04 with weight and 0.10 ± 0.04 without weight fitted as a covariate (Jopson and Young pers. comm). The CPT results are from maternal sheep breeds not dissimilar from those used in the INF i.e. Corriedale, Coopworth and Romney. Higher heritability estimates for age at first corpus lutea have been reported in tropical cattle breeds, ranging from 0.52 to 0.57y (Johnston *et al.* 2011). However, age at first corpus lutea is a direct measure of sexual maturity whereas AFO in this study is an indirect measure of pubertal development. This may explain to some extent the lower heritabilities reported here.

Table 2. Phenotypic (V_p) and genetic group (V_{gg}) variances, heritabilities (on the diagonal) and genetic (above) and phenotypic (below) correlations for AFO, AFO2, YNLB, YNLW.* Standard errors in brackets.

	AFO	AFOnwt	AFO2	AFO2nwt	YNLB	YNLW
V_p	205.53 (9.97)	220.80 (10.68)	205.53 (9.97)	538.14 (20.00)	0.35 (0.01)	0.29 (0.01)
V_{gg}	30.78 (21.73)	24.08 (19.35)	30.78 (21.73)	26.16 (21.19)	0.04 (0.03)	0.02 (0.02)
AFO	**0.04** ***(0.07)***	--	--	--	0.38 *(0.26)*	0.46 *(0.27)*
AFOnwt	--	**0.04** ***(0.06)***	--	--	0.70 *(0.36)*	0.80 *(0.36)*
AFO2	--	--	**0.04** ***(0.07)***	--	0.74 *(0.42)*	0.87 *(0.45)*
AFO2nwt	--	--	--	**0.09** ***(0.05)***	0.01 *(0.20)*	0.05 *(0.21)*
YNLB	-0.02 *(0.03)*	0.02 *(0.03)*	0.06 *(0.02)*	-0.05 *(0.02)*	**0.20** ***(0.05)***	0.92 *(0.15)*
YNLW	-0.02 *(0.03)*	0.01 *(0.03)*	0.04 *(0.02)*	-0.05 *(0.02)*	0.76 *(0.02)*	**0.17** ***(0.05)***

*AFO (Age of first oestrus), AFOnwt (AFO without weight fitted as a covariate), AFO2(Synthetic AFO trait)

Heritability estimates of 0.20 ± 0.05 and 0.17 ± 0.05 for YNLB and YNLW are higher than what is commonly reported for number of lambs born and weaned from multi-parity analyses (Safari and Fogarty 2003). Bunter and Brown (2013) also reported higher heritability estimates for YNLB and YNLW, 0.13 ± 0.02 and 0.08 ± 0.01 respectively, though not as high as in this study.

Genetic and phenotypic correlations between YNLB and YNLW where high as was expected, 0.92 ± 0.15 and 0.76 ± 0.01. Phenotypic correlations between YNLB and YNLW and all AFO traits were low and generally not significantly different from zero. Genetic correlations between YNLB, YNLW and AFO traits were all positive with high standard error. Based on the assumption that animals that mature earlier are more likely to have a lamb, this is not what was anticipated as the direction of the correlation suggests that animals that mature later (marked at older ages by teaser) are more likely to rear a lamb. There are a number of possible explanations for this unexpected finding: firstly the low number of records, low number of progeny per sire and low trait heritabilities has resulted in a high standard error so we cannot be certain of this correlation. The length of teaser exposure in the collection of AFO data varied from the 4 month period from weaning to joining to a 5 week period immediately prior to joining. Due to incomplete recording of length of exposure to teasers this failure to follow the standardized procedures for collecting AFO could not be factored into the analysis.

Notter (1989) demonstrated that the continuous exposure of ewes to males delays the start of joining in comparison to ewes that are isolated from males prior to joining. NLB average for this data set was well below the 0.79 reported for yearling ewes by Bunter and Brown (2013). It is possible the use of teasers for an extended period delayed when some ewes were ready to join until

after ram removal confounding measurement of YNLB and YNLW. This study seems to suggest that the use of teasers for extended periods is not a useful indicator for successful early reproductive performance and lead to lower lambing percentages.

CONCLUSIONS

Heritability estimates for YNLW and YNLB appear to be higher than estimates from multi-parity analysis. AFO heritability estimates were lower than reported elsewhere, possibly influenced by variation in length of teasing between sites. Genetic correlations were positive though associated with high standard error. This was unexpected as it seems to indicate that earlier maturing animals were less likely to have a lamb. Whilst the use of teasers for short periods prior to joining has been shown to successfully lift reproductive rate, the use of teasers for extended periods may in fact reduce lambing percentage. This coupled with the high standard errors found in this study seem to suggest that using teasers to measure AFO to improve early reproductive performance is not desirable. Thus, a need exists to find reliable measures for early reproductive traits including sexual maturity.

ACKNOWLEDGEMENTS

The authors would like to acknowledge the Sheep CRC and the staff associated with the project for collecting and providing access to the data. The authors would also like to thank Jen Smith and Klint Gore for their invaluable advice on the data analysed and Kim Bunter for useful comments on the project and this manuscript. Thank you also to MLA for funding this PhD Research.

REFERENCES

Bunter K.L. and Brown D.J. (2013). *Proc. Assoc. Advmt. Anim. Breed. Genet.* **20**: (this publ.)
Fogarty N.M. (2007) *Aust. J. Agr. Res.* **58**: 928.
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.
Gilmour A.R., Gogel B.J., Cullis B.R. and Thmpson R. (2009) ASREML User Guide Release 3.0. VSN International Ltd, Hemel Hempsted, HA1 1ES, UK.
Johnston D. J., Barwick S.A, et al. (2009) *Anim. Prod. Sci.* **49**: 399.
Notter, D.R. (1989) *Anim. Reprod Sci.* **19**: 265.
Safari A. and Fogarty N. M. (2003) 'Genetic parameters for sheep production traits: estimates from the literature.' NSW Agriculture, Orange.
van der Werf J.H.J, Kinghorn B.P., and Banks R.G. (2010) *Anim. Prod. Sci.* **50**: 998.

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:90-93

POTENTIAL ECONOMIC RETURN FROM USE OF FIXED-TIME ARTIFICIAL INSEMINATION AS PART OF A GENETIC IMPROVEMENT PROGRAMME

S.A.A Edwards[1], B.M. Burns[2], J. Allen[3] and M.R. McGowan[1]

[1]School of Veterinary Science, The University of Queensland, Gatton, QLD 4343
[2]The University of Queensland, Centre for Animal Science, Queensland Alliance for Agriculture and Food innovation, Rockhampton, QLD 4700
[3]Agriculture Business Research Institute, University of New England, Armidale, NSW 2350

SUMMARY

To investigate the potential return on investment of implementing a genetic improvement program in a self-replacing commercial Brahman breeding herd, three different selection and breeding strategies were evaluated through modelling, 1) Natural mating with no genetic improvement (NATM-G), 2) Natural mating with genetic improvement (NATM+G), and 3) Fixed-time AI (FTAI) with genetic improvement (FTAI+G). In each scenario, the Jap Ox Index was used to quantify genetic gain and improvements were made using a Brahman sire with a top 10% Jap Ox Index ($45). A sire was selected from the progeny generated in Year 1. This sire was then used in Year 3 for natural mating in a multiplier herd. A partial budget was used to calculate the cost per calf weaned. The costs per calf weaned in Year 1 were calculated to be $46.83, $371.42 and $173.76 for NATM-G, NATM+G and FTAI+G, respectively. The Jap Ox Index for the progeny was calculated to be $20.00, $32.50 and $32.50 for NATM-G, NATM+G and FTAI+G, respectively. However, when progeny from Year 1 were used in Year 3 for breeding, the costs per calf weaned in Year 3 were calculated to be $46.83, $10.27 and $4.35 for NATM-G, NATM+G and FTAI+G, respectively. In Year 3, Total Genetic Profit was calculated to be $0, $124.38 and $1017.00 for NATM-G, NATM+G and FTAI+G, respectively. This model supports the return on investment in genetic improvement in Brahman cattle in northern Australia, and demonstrates the potential value of FTAI in both disseminating improved genetics and improving rate of genetic gain.

INTRODUCTION

A range of local and global factors are impacting on the Australian beef industry contributing to an average return on assets of only 0.3 to 2.0%. Poor reproductive performance in extensively managed tropically adapted herds is one factor contributing to this poor financial performance (McCosker *et al.* 2010). Genetic improvement to increase herd productivity with a strong emphasis on reproduction has the ability to improve the financial performance of northern breeder herds. The results from recent molecular and quantitative genetic research enable selection of superior tropical breed sires for a range of traits such as age of puberty, postpartum re-conception interval and lifetime productivity (Fortes *et al.* 2012; Johnston *et al.* 2009). The large genetic variation in reproduction traits observed in Brahman genotypes provides substantial opportunity for improvement through genetic selection (Johnston *et al.* 2009). Artificial insemination (AI) provides a practical method of increasing the dissemination of superior genetics in commercial and seed-stock bull breeding herds. The use of AI in northern Australia is currently estimated to be less than 1% of the breeder herd and traditionally considered difficult to implement in extensively managed herds. A strategy to increase the dissemination of superior genetics in northern beef herds is use of fixed-time AI (FTAI), which eliminates the need for oestrus detection. FTAI is often associated with lower labor inputs, and enables insemination of large numbers of females and production of more calves than typical oestrus detection programs (Edwards *et al.* 2012). The objective of this study was to use modelling to compare the potential return on investment of implementing three different selection and breeding strategies 1) Natural

mating with no genetic improvement (NATM-G), 2) Natural mating with genetic improvement (NATM+G), and 3) FTAI with genetic improvement (FTAI+G), in a self-replacing commercial Brahman breeding herd.

MATERIALS AND METHODS

The Brahman Jap Ox index was used to quantify genetic merit of sires (ABRI 2013) used in three different selection and breeding strategies; Strategy 1: NATM using breed average sires with no genetic improvement (NATM-G), Strategy 2: NATM with genetic improvement using a purchased top 10% Jap Ox sire (NATM+G), and Strategy 3: FTAI with genetic improvement using a top 10% Jap Ox sire (FTAI+G) and using NATM+G in Year 3 from selected progeny from Year 1. In each strategy, bulls were produced by NATM or FTAI in Year 1 from the bull breeding herd and used in Year 3 in the multiplier herd. Assumptions for purchase of sire and frozen semen, pregnancy rate to FTAI and overall weaning rate, and costs of FTAI in a 200 cow breeding herd are presented in Table 1.

The cows mated in each strategy were all assumed to have a breed average Jap Ox Index ($20). Genetic gain was calculated for each strategy using the following equations: *[(Sire Jap Ox Index) – ($20)]/2 = Calf Genetic Improvement.* In Year 3, when bulls produced from the Year 1 mating are used in the multiplier herd, the genetic gain is calculated as described above.

Table 1. List of assumptions and costs associated with NATM or FTAI

Item	Parameters and costs	Source
Breed average Brahman sire	Purchase price: $5,000	
Top 10% Jap Ox Brahman sire	Purchase price: $40,000; Semen Price: $50	
Station labour (@ $200/day)	FTAI: 5 personnel x 3 days = 15 units = $3000 NATM: 2 personnel x 1 days = 2 units = $400	
FTAI costs	Drugs to synchronise ovulation: $3524 AI technician: $1500	
Expected sire working life	4 years	(Smith *et al.* 2011)
Weaning rate (% cows joined)	71 %	(Schatz and Hearnden 2008)
Pregnancy rate to FTAI	35 %	(Edwards *et al.* 2012)
Bull:Cow ratio (NATM)	5 bulls for 200 cows (2.5%)	(Smith *et al.* 2011)

RESULTS AND DISCUSSION

The costs per calf born in Years 1 and 3 of each strategy are presented in Table 2. In the genetic improvement strategies, more genetically superior progeny were produced using FTAI than NATM (63 vs. 28, respectively). In the NATM+G scenario, as the purchase price of a natural mating sire is relatively high, only one sire was used, and thus the number of cows that could be mated to this sire was only 40 (using a 2.5% mating ratio). This strategy limits the production of genetically superior calves compared to that achieved using FTAI, where all cows in the bull breeding herd were AI once, resulting in a higher total number of genetically superior calves being produced. As a result, in both Years 1 and 3 the cost per genetically superior calf born was lower for the FTAI strategy compared to the NATM-G strategy.

Table 2. Cost per calf generated from NATM-G, NATM+G and FTAI+G strategies

Year 1	*Calculation*	*NATM-G*	*NATM+G*	*FTAI+G*
Bull breeding herd (n)	(A)	200	40[a]	200
FTAI costs[b]	(B)	-	-	$ 15,024.00
Cost per sire *(Table 1)*	(C)	$ 5,000.00	$ 40,000.00	$ 5,000.00
Sires (n) *(Table 1)*	(D)	5	1	5
Total sire expenses	C*D = (E)	$ 25,000	$ 40,000.00	$ 25,000
Labour costs	(F)	$ 400.00	$ 400.00	$ 3,400.00
Mating costs for Yr 1[c]	[B+(E/4)] + F = (G)	$ 6,650.00	$10,400.00	$ 24,674
Progeny by high genetic merit bull[d]	NATM: (A*0.71) = (H) FTAI: (A*0.35) = (H)	-	28 calves	63 calves
Progeny by average genetic merit bulls	NATM: (A*0.71) = (I) FTAI: (A*0.71)-H = (I)	142 calves	-	79 calves
Cost per calf	G/(H+I) = (K)	$ 46.83	$ 371.42	$173.76
Year 3		*Natural mating using sires generated in Yr 1*		
Bull breeding herd (n)	(L)	200	80[b]	200
Cost per sire	NATM-G: New Sires = (M) NATM+G, FTAI+G: K = (M)	$ 5,000.00	$ 371.42	$173.76
Sires (n) *(Table 1)*[e]	(N)	5	2	5
Total sire expenses	N*M= (O)	$ 25,000	$ 742.84	$ 868.80
Labour costs	(P)	$ 400.00	$ 400.00	$ 400.00
Mating costs for Yr 3	(O/4) + P = (Q)	$ 6,650.00	$ 585.71	$ 617.20
Progeny from mating	L*0.71 = (R)	142 calves	57 calves	142 calves
Total cost per calf	Q/R = (S)	**$ 46.83**	**$ 10.27**	**$ 4.35**

[a]Due to the relatively high purchase price it is assumed that only 1 purchased sire was used to breed replacement bulls.
[b] Insemination expenses include: Drugs to synchronise ovulation and, AI technician and semen costs.
[c] Mating costs include: Sire expenses and labour costs for mustering and yard handling associated with the mating strategy.
[d] Genetically improved progeny include: Number of calves born from genetic improvement mating. Weaning rate and pregnancy rates to FTAI are as per Table 1.
[e] A selection intensity of 16% was applied to sires generated from Year 1. Therefore, only 2 sires were retained to join 80 cows in the NATM+G strategy, however, 5 sires were available to join the entire bull breeding herd in the FTAI+G strategy.

The lack of adoption of artificial breeding technologies in the northern beef industry could be due to a perceived high cost per calf born. As FTAI+G can generate more high genetic merit calves than natural mating, the total costs of genetic improvement are spread across a greater number of progeny, resulting in a lower cost per calf born than NATM+G. This model assumes that the price of a natural mating sire is correlated with its genetic merit and in turn is correlated with price of semen from this sire. Some assumptions that have not been included in the model, are: 1) Genetically improved male progeny not retained for use in the herd may be sold for a higher price than average genetic merit progeny, 2) As a high selection pressure is applied to male progeny (only 16% of available progeny selected) the retained sires should have a higher actual Jap Ox index than calculated in the model, 3) Transport and other associated expenses of purchase of a high genetic merit natural mating sire have not been included, and 4) An increased proportion of females conceiving earlier in the mating period in FTAI may improve weaner values (Spitzer 1986). Total Genetic Profit was calculated to be $0, $237.25 and $1275.00 for NATM-G, NATM+G and FTAI+G, respectively (Table 3). In this comparison the FTAI+G strategy improved the genetic profit of the calves 5.4 times more than the NATM+G strategy. This is explained by the FTAI+G strategy producing 85 more calves by high genetic merit sires multiplying the effects of the genetic improvement strategy.

Table 3. Genetic profit from NATM-G, NATM+G and FTAI+G strategies.

Year 1	*Calculation*	*NATM-G*	*NATM+G*	*FTAI+G*
Bull breeding herd (n)	(A)	200	40[a]	200
Jap Ox Index of sires	(B)	$ 20	$ 45	$ 45
Average Jap Ox Index of cows	(C)	$ 20	$ 20	$ 20
Genetic gain per calf born	(B-C)/2 = (D)	$ 0	$ 12.50	$ 12.50
Calves by genetic. superior sire	(E)	0	28	63
Calves by genetic. average sire	(F)	142	-	79
Total genetic gain	E*D = (G)	$ 0.00	$ 350.00	$ 787.50
Jap Ox Index of progeny	(H)	$ 20.00	$ 32.50	$ 32.50
Year 3		*Natural mating using sires generated in Yr. 1*		
Bull breeding herd (n)	(I)	200	80	200
Jap Ox Index of sire	= (H)	$ 20.00	$ 32.50	$ 32.50
Calves from mating	I*0.71 = (J)	142	57	142
Genetic gain over average cow	(H-C)/2 = (K)	$ 0	$ 6.25	$ 6.25
Genetic gain – calves from replacement cows Yr. 1[c]	(D*0.5)*((E*0.5)*0.71)=(L)	$ 0	$ 62.13	$ 142.00
Calves from mating	(M)	140	56	140
Yr. 3 genetic gain of progeny	M*K = (N)	$ 0	$ 62.25	$ 875.00
Total Genetic Profit	L + N = (O)	**$ 0**	**$ 124.38**	**$ 1017.00**

[a] Due to the relatively high purchase price it is assumed that only 1 purchased sire will be used to breed replacement bulls.
[b] A selection intensity of 16% is applied to sires generated from Year 1. Therefore only 2 sires are retained to join 80 cows in the NATM+G Strategy, however, 5 sires are available to join to the entire bull breeding herd in the FTAI+G strategy.
[c] Assume all heifers from Year 1 are retained and bred in Year 3. Assume 50% of the calves born in Year 1 are female and the weaning percentage of these calves is 71%.

CONCLUSION

The results from this modelling support the return on investment in genetic improvement in Brahman cattle in northern Australia and demonstrate the potential value of FTAI in both disseminating improved genetics and improving rate of genetic gain.

REFERENCES

ABRI. (2013). *http://abri.une.edu.au* Accessed 8/2/2013 2013.

Edwards S.A.A., Bo G. A., Dawson K., and McGowan M.R. (2012). In: 'Proceedings 17th Interational Congress on Animal Reproduction (ICAR), Vancouver, July 2012

Fortes M. R. S. et al. (2012). *Anim. Prod. Sci.* **52**: 143.

Johnston, D. J. et al. (2009). *Anim. Prod. Sci*. **49**: 399.

McCosker T., McLean D., and Holmes P. (2010). Meat and Livestock Australia Ltd., North Sydney, NSW.

Schatz T. J., and Hearnden M.N. (2008). *Aust. J. Exp. Agr.* **48**:940.

Smith D., McCosker K., Butler S. Fordyce G., Burns B., Jephcott S. and McGowan M. (2011) *Proc. North. Beef. Res. Updat. Conf.* 122

Spitzer J. C. (1986). In: Current therapy in theriogenology: Diagnosis, treatment and prevention of reproductive diseases in small and large animals. p 320-341. editor D. A. Morrow, W.B. Saunders Company, Philadelphia.

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:94-97

BENEFITS OF MOET AND JIVET IN OPTIMISED SHEEP BREEDING PROGRAMS

T. Granleese[1,2], S.A. Clark[1], J.H.J. van der Werf[1,2]

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

SUMMARY

The additional genetic gain from implementing multiple ovulation and embryo transfer (MOET) and juvenile *in vitro* embryo production and embryo transfer (JIVET) additional to using artificial insemination (AI) and natural mating (N) in sheep breeding programs was assessed. This study was a stochastic simulation and selection based on optimum contributions for varying levels of inbreeding restriction. The genetic gain achieved after 20 years for an AI/N program was 4.89 and 5.16 units of genetic SD (h^2=0.3) when inbreeding was restricted to 1% and 2% per generation, respectively. The additional gain from MOET was 23% and 28% and the additional gain from the addition of JIVET to MOET and AI/N increased genetic gain 60% and 56% for these two levels of inbreeding when compared to AI/N. With the addition of each technology, generation interval decreased, as did the number of breeding ewes.

INTRODUCTION

Artificial insemination (AI) has been used by producers to increase selection intensity in males to increase genetic gain. Further to this, multiple ovulation and embryo transfer (MOET) and juvenile *in vitro* embryo production and embryo transfer (JIVET) are female reproductive technologies that have been employed by sheep producers to increase female selection intensity, decrease generation interval and hence increase genetic gain. There are also some limitations associated with using these technologies. One problem with using JIVET in a breeding program is that selection accuracy is often low when females are selected at a young age. Another problem is that increasing the number of progeny per breeding female can increase the rate of inbreeding significantly (Quinton and Smith 1995). Problems due to inbreeding can potentially offset any additional gains in merit that are associated with using these technologies. Optimal contribution selection principles have been developed to manage the balance between increases in genetic merit while controlling genetic diversity and inbreeding (Wray and Goddard 1994, Meuwissen 1997). Horton (1996) simulated 3% increase in genetic gain with an inbreeding rate of 8% per year in a closed Merino flock implementing AI. Brash *et al.* (1996) showed that in a closed nucleus Merino stud MOET can increase rates of genetic gain by 22% per year. However, these additional gains also resulted in a 50% increase in inbreeding rate.

This paper aims to explore the potential benefit of MOET and JIVET in sheep breeding programs while managing inbreeding. Various levels of inbreeding restrictions will be explored by invoking optimal contributions selection and applying an optimal mixture of matings using AI or natural breeding (AI/N), MOET and JIVET.

METHODS

A closed nucleus breeding program generating 250 progeny per year using a stochastic simulation program was used. Each scenario generated a base population of unrelated animals, and subsequent generations were selected on pedigree-based breeding values. Phenotypes for a single trait were simulated with a heritability of 0.3 and a phenotypic standard deviation of 10. Each year, all animals were assigned breeding values estimated using Best Linear Unbiased Prediction (BLUP).

Optimal selection was used to maximise genetic gain while maintaining genetic diversity. Using Wray and Goddard's (1994) formula, genetic merit (M) was balanced with co-ancestry (C), where, M= x'b, b is a vector of BLUP breeding values and x is a vector of genetic contributions of candidate animals with values in x summing to 0.5 for both males and females. Inbreeding rates were managed by penalizing the average co-ancestry among selected animals; $C=\lambda x'Ax$, where A is an (n x n) relationship matrix among candidates and λ is the penalty to restrict inbreeding. Price and Storn's (1997) evolutionary algorithm was used to find optimal solutions for M + C. Various values of λ were used to explore a 'frontier' of optimal selection outcomes which resulted in different levels of inbreeding and genetic gain.

In this study, three breeding programs were compared: 1) AI/N mating only, 2) AI/N + MOET and 3) AI/N + MOET + JIVET. In each breeding program AI was used and therefore, depending on the inbreeding restriction, a single male could be assigned to all dams (200+). Females however were limited to just one mating if they were assigned either an AI/N service or if they went into a MOET program. Juvenile females were assigned three matings (due to oocyte numbers recovered and individual oocyte mating ability in IVF process) if they were nominated to be used in the JIVET program only. Males were eligible to enter any breeding programs once they were over a year old. Ewes in AI/N or MOET programs were also only eligible once they were 18 months old. Ewes in the JIVET program were eligible within 3 months of age. If any individuals did not get selected in a breeding program, they were culled. However, in the JIVET program, if a ewe was not selected as a lamb it was again eligible for selection at 18 months of age. All sheep in all programs were culled once they finished five years of life. A mortality rate of 10% was applied each year. The probability of producing a certain number of offspring for AI/N, MOET (Gibbons and Marcella 2011) and JIVET (Armstrong *et al.* 1997) is summarized in Table 1. Each scenario was run for 20 years and replicated 90 times.

Table 1 Probability of producing a certain number of progeny per female per mating for the various reproductive methods.

Progeny	0	1	2	3	4	5	6	7	8	Ave
AI/N	0.1	0.7	0.2							1.1
MOET	0.1	0.05	0.05	0.15	0.25	0.15	0.13	0.07	0.05	4.02
JIVET	0.25	0.05	0.18	0.18	0.1	0.1	0.07	0.04	0.03	8.37*

*predicted average of total progeny of 3 JIVET matings

RESULTS AND DISCUSSION

When inbreeding was unrestricted, the inbreeding coefficient in the AI/N + MOET + JIVET program over 20 years was 165% higher than the maximum inbreeding in both AI/N and AI/N + MOET breeding programs. The maximum genetic gain over the 20 years was also 70% and 40% higher compared to AI/N and AI/N + MOET breeding programs, respectively (Figure 1). This genetic gain and level of inbreeding is considerably less when compared to Pryce *et al.'s* (2010) study where they reached an increase of 231% genetic gain per year and 165% inbreeding per generation in a closed Holstein nucleus. However they used genomic selection in heifers only with no mature cows mated.

There is debate over what the "ideal" amount of inbreeding per generation is and this may vary depending on the species, breed and ability to open the breeding program (Goddard 2009). Responses to the breeding programs were compared under two inbreeding rates: 1% and 2% per generation. Note the number of generations in 20 years (Table 2), and therefore the generation interval differed between the different scenarios.

At inbreeding rates of 1% per generation, JIVET programs can yield up to 60% and 30% more genetic gain than AI/N only and AI/N + MOET breeding programs, respectively (Figure 1, Table 2). At inbreeding levels of 2% per generation, these additional gains are 56% and 29%. However, for double the inbreeding, we see relatively small changes in genetic gain in all scenarios.

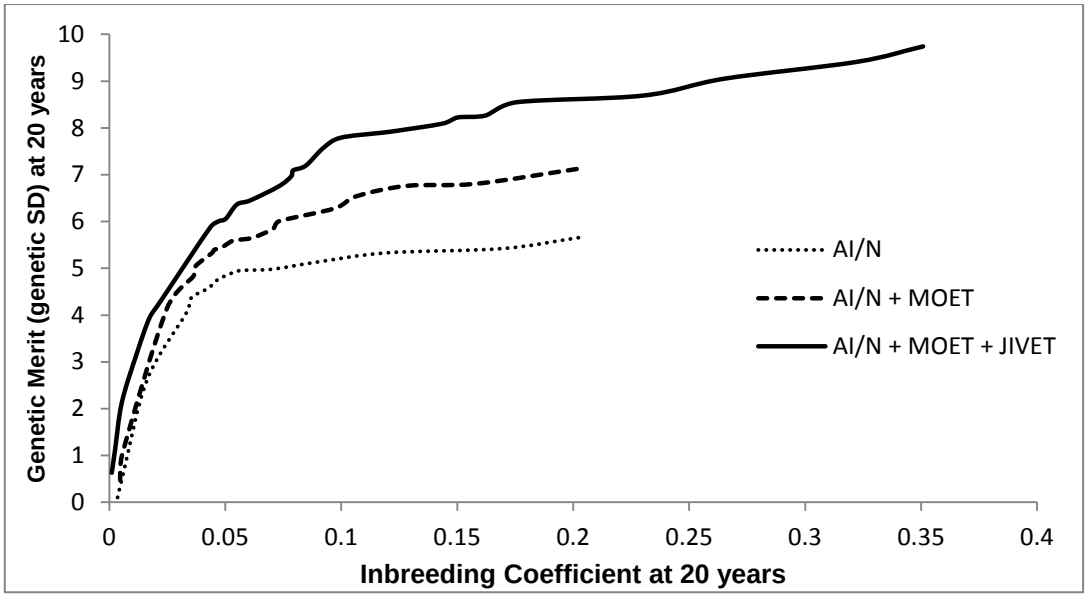


Figure 1 Level of inbreeding with level of genetic gain (in units of genetic SD) in three separate breeding programs.

The AI/N breeding program was used as a base of comparison as in either case ewes would have similar numbers of progeny. Breeding programs with small penalties often only used one or two rams each year which would be deemed not possible when servicing 200+ ewes, hence the reason to switch between natural mating and AI. As the penalty for inbreeding increased so did the number of rams used. MOET breeding programs use considerably less dams than AI/N while the inbreeding rate was not increased. Optimal selection manages to increase selection intensity in females while maintaining diversity, by selecting fewer dams per family rather than selecting fewer families.

An issue from this simulation that needs considering is shortened breeding cycle of ewes who are in JIVET programs. This study assumed all the programs' lambing occurred annually in a seasonal fashion. However further consideration needs to be taken as ewes in JIVET programs are lambing out of seasonal synchrony and generation intervals can be as short as 6 months.

If such short generation intervals were considered, further genetic gain would be expected due to the decreased generation interval. To further increase genetic gain, "age of first mating" could also be decreased for both males and females in AI/N and MOET with 18 months being conservative for a first mating age.

This study did not consider genomic selection. Genomic selection would allow earlier selection of elite juvenile animals because the accuracy of EBVs is higher and increase is relatively highest for young animals that have initially low EBV accuracy. Therefore, the next step in this study is to incorporate genomic EBVs. This would be expected to further increase the benefit of JIVET and

MOET technologies. This study has not factored in costs associated with the reproductive technologies. This will be investigated in further studies in the future.

Table 2 Number of generations (Gens) and ewes (n) used in 20 years at 1% increase in inbreeding per generation (dF), and genetic gain (SD) at 1 and 2% dF

	Gens 1%dF	Ewes	Gain at 1% dF	Gain at 2% dF
AI/N	6.04 ±0.03	235.61 ±1.77	4.89 ±0.03	5.16 ±0.03
AI/N + MOET	7.81 ±0.05	87.96 ±0.94	6.03 ±0.03	6.63 ±0.03
AI/N + MOET + JIVET	10.03 ±0.09	64.68 ±0.24	7.78 ±0.04	8.54 ±0.05

CONCLUSION

Optimal selection techniques used in breeding programs that incorporate female reproductive technologies are shown to increase genetic gain considerably while maintaining acceptable inbreeding levels. The addition of MOET to AI/N breeding programs increased genetic gain and led to shorter intervals. This trend is further increased with the introduction of JIVET. Therefore both MOET and JIVET can contribute significantly to aid in accelerating genetic gain in sheep breeding programs and this benefit is expected to be enhanced by genomic selection.

ACKNOWLEDGEMENTS

The authors would like to thank Andrew Swan for his assistance in developing computer code.

REFERENCES

Armstrong D. T., Kotaras P. J. and Earl C. R. (1997) *Repro. Fert. Devel.* **9**:333.
Brash L. D., Wray N. R. and Goddard M. E. (1996) *J. Anim. Sci.* **62**:241.
Gibbons A. and Marcela C (2011). "Embryo Transfer in Sheep and Goats: A Training Manual." National Institute for Agricultural Technology, Argentina.
Goddard M.E. (2009) In 'Adaptation and Fitness in Animal Populations', pp.41-50, editors J.H.J van der Werf, H.-U. Graser, R. Frankham, C. Gondro, Springer, United Kingdom.
Horton B. (1996) *Aust. J. Exper. Ag.* **36**:249.
Meuwissen T.H.E. (1997) *J. Anim. Sci.* **75**:934.
Nicholas F.W. (1996) *Anim. Repro. Sci.***42**:205.
Price K. and Storn R. (1997) *Dr Dobb's J.* **64**:18.
Wray N.R. and Goddard M.E. (1994) *G.S.E.* **26**:431.

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:98-101

ANALYSIS OF A SOUTH AFRICAN MERINO FLOCK DIVERGENTLY SELECTED FOR REPRODUCTIVE POTENTIAL

L. Sandenbergh[1], R. Roodt-Wilding[1], A.E. Van der Merwe[1] and S.W.P. Cloete[2,3]

[1]Department of Genetics, Stellenbosch University, Private Bag X1, Matieland, 7602, South Africa
[2]Department of Animal Sciences, Stellenbosch University, Private Bag X1, Matieland, 7602, South Africa
[3]Directorate Animal Sciences: Elsenburg, Private Bag X1, Elsenburg, 7607, South Africa

SUMMARY

A South African Merino sheep flock has been divergently selected for more than 8 generations for the ability of ewes to rear multiple offspring. Selection has resulted in a High line and a Low line that differ markedly in their reproductive output. The causative mutations and/or quantitative trait loci responsible for the difference in reproductive traits between these 2 lines have not yet been determined. Genomic regions under selection would be expected to demonstrate the highest level of genetic differentiation between these lines and would also exhibit a higher than expected degree of homozygosity within lines. Selected individuals were genotyped using the OvineSNP50 BeadChip and the genotype data were analysed to identify differences between the 2 lines. The High line and Low line ewes were shown to be phenotypically and genetically discrete; confirming the presence of 2 distinct lines. Several markers subjected to selection could be identified between the 2 lines. It can be assumed that most of these markers differ as a result of the differential selection pressure applied on reproduction. Further investigation into these loci could provide valuable information on the genes and/or quantitative trait loci involved in an improved phenotype with greater reproductive success.

INTRODUCTION

Reproduction traits are of importance in the improvement of the economic output of industry sheep flocks (Olivier 1999; Safari *et al*. 2005). Net reproduction, defined as the number of lambs (or weight of lamb) weaned per ewe mated, is a lowly heritable, gender-limited, composite trait, recorded later in life. The recording of some of the net reproduction components, such as ovulation rate, conception rate and embryo survival, is notoriously complicated, costly and labour-intensive. Conventional breeding efforts depend on recorded data linked to pedigree information to implement an efficient improvement strategy and therefore rely heavily on the recording of relevant traits (Notter 2012). This necessitates accurate performance testing of all potential breeding animals, thereby increasing costs and slowing down the rate of any potential genetic gains. Molecular markers associated with reproductive traits could accelerate genetic improvement by facilitating a more accurate estimation of reproductive potential at a much earlier age (Dodds *et al*. 2007; Hayes *et al*. 2009).

Several studies have reported mutations in a single ovine gene or closely linked group of genes that result in highly proliferative lines. Although lamb rearing ability is assumed to be a complex trait (Notter 2012), the chromosomal regions of previously identified mutations could serve as candidate regions for the current study. Mutations in 3 major genes, the bone morphogenetic protein receptor 1B, bone morphogenetic protein 15 and growth and differentiation factor 9, result in increased ovulation rate in sheep. These genes are located on chromosome 6, the X chromosome and chromosome 5, respectively (Davis 2005; McNatty *et al*. 2007).

A South African Merino sheep flock has been divergently selected for their ability to rear multiple offspring since 1986. Selection has been applied for more than 8 generations and has resulted in a High line and a Low line that differ markedly in their reproductive output (Cloete *et*

al. 2004). The causative mutations and/or quantitative trait loci responsible for the difference in reproductive traits between these 2 lines have not been determined. These lines could potentially serve as a model for identifying the genomic regions underlying reproductive traits and to determine whether mutations identified in other highly proliferative lines are also segregating in this flock.

MATERIALS AND METHODS

The divergently selected lines are maintained at the Elsenburg research farm in the Western Cape province of South Africa. The location, animal resource and the selection protocol followed is detailed in the literature (Cloete *et al*. 2004; 2009). Blood samples were obtained from 112 individuals from the lines and were genotyped using the OvineSNP50 BeadChip (Illumina) at GeneSeek Inc (Lincoln, USA). Sampled individuals were representative of the recent genetic composition of the lines (born between 2002 and 2008); with accurate estimated breeding values for number of lambs weaned per parity and total weight weaned per parity; and represent the extremes of the phenotypic distribution. Pedigree information was considered to minimise the inbreeding and relatedness in the sampling cohort to reduce potential within-line population substructure. A t-test was performed to confirm that significant differences exist between the phenotypes of the 2 lines.

Only samples with a call rate of >85% and single nucleotide polymorphism (SNP) loci with a call rate of >85%, GenCall score of >0.25, GenTrain score of >0.50 and minor allele frequency of >0.01 were included in downstream analyses. In an attempt to investigate the possibility that chromosomes cited in literate play a significantly more important role in reproduction in the current study, the genotypic data were partitioned by chromosome. A factorial component analysis was conducted in Genetix (Belkhir *et al*. 2004) to assess the multi-factorial variance between the lines and evaluate the degree of clustering on a three-dimensional scale. Markers subjected to selection were identified by an Fst outlier approach using a Bayesian method in BAYESCAN (Foll and Gaggiotti 2008) and using the FDIST2 method (Beaumont and Nichols 1996) in Lositan (Antao *et al*. 2008). Markers found to be subjected to selection based on both methods of analyses were further investigated using the *Ovis aries* genome –Annotation release 100 (www.ncbi.nlm.nih.gov/projects/mapview).

RESULTS AND DISCUSSION

Ninety-one individuals and 23 781 SNP loci met the quality control measures and were included for downstream analyses (Table 1). The relatively low SNP loci yield after quality control is most probably due to a loss of DNA quality during processing rather than a lack of polymorphic markers in the study cohort. Genotyping of other South African Merino sheep have yielded a higher number of markers after quality control (data not shown).

Table 1. A total of 91 of the original 112 samples were included in further analyses

Total number of sampled individuals		Samples included in analyses	
27 Rams	19 High line	23 Rams	16 High line
	8 Low line		7 Low line
85 Ewes	45 High line	68 Ewes	38 High line
	40 Low line		30 Low line

The t-test confirmed significant differences in the phenotypic values of the 2 lines for number and weight born per joining, consistent with the study by Cloete *et al*. (2004). The factorial component plot for each chromosome indicated 2 distinct clusters representing the divergently selected lines (Figure 1). The lines can therefore be considered to be phenotypically and genetically distinct as a result of several generations of selection.

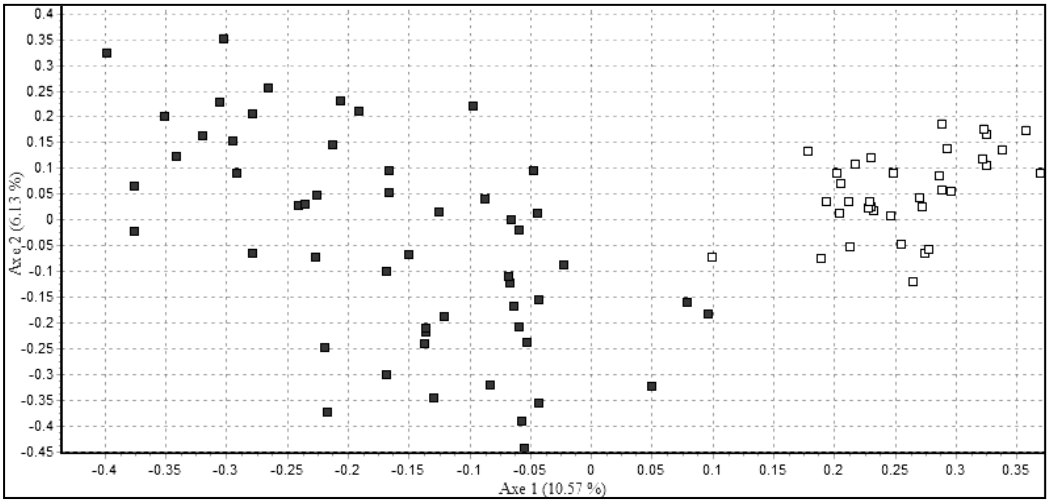


Figure 1. A factorial component plot of chromosome 5, indicating 2 distinct clusters. Dark squares represent individuals from the High line and white squares that of Low line.

The FDIST2 method indicated 1476 markers to be subjected to selection on the 27 chromosomes in the ovine genome. This number was reduced to 926 after a correction for multiple testing was implemented (Figure 2). The Bayesian-based analysis however, identified only 47 markers to be subjected to selection. The overall percentage of markers subjected to selection was 4.00% using the FDIST2 method and 0.20% using the Bayesian method.

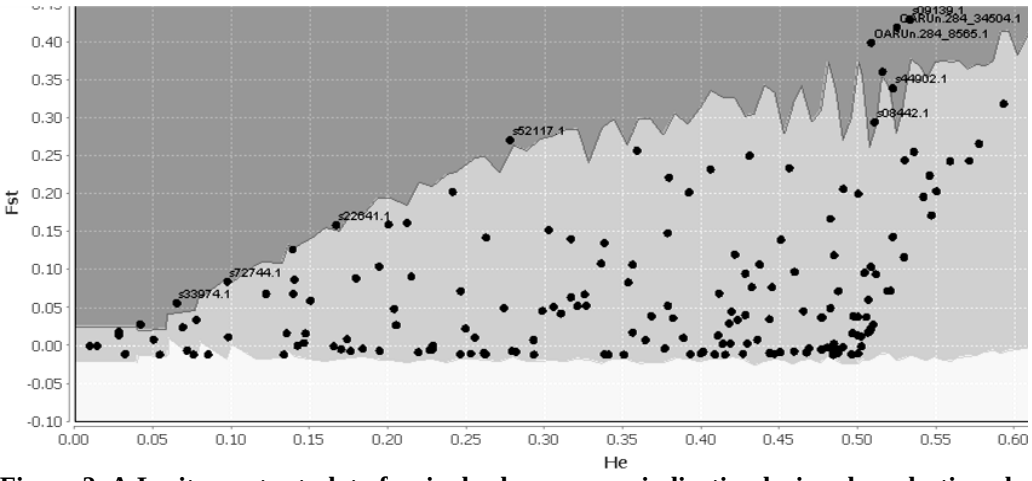


Figure 2. A Lositan output plot of a single chromosome indicating loci under selection above the 95% percentile (darker shaded area), neutral markers (lighter shaded area) and markers under balancing selection (white area).

All markers identified according to the Bayesian method were also identified by the FDIST2 method. The disparity between the numbers of markers identified by the 2 different methods has been noted in other related studies (Narum and Hess 2011). The Bayesian method implemented in BAYESCAN has been shown to be less prone to type I errors (false positives) compared to the method employed in Lositan and this could explain the larger number of markers being identified by Lositan in the current study.

Chromosome 5, 6 and the X chromosome did not exhibit a significant difference in the percentage of markers subjected to selection. No significant difference was seen between the number of markers or extent of putative markers under selection within any of the individual chromosomes.

Several markers were found to be located in or near annotated genes. One of these genes, the corticotropin releasing hormone gene on chromosome 9 is especially noteworthy as the corticoid pathway has been shown to influence reproduction in sheep (Breen *et al*. 2005).

CONCLUSIONS

The divergently selected Elsenburg flock can be considered a valuable genetic resource for studies aiming to identify genomic regions playing a role in reproductive traits. This study identified a large number of markers across the ovine genome that appear to be subjected to selection, thereby supporting the premise that reproductive traits are under the control of several loci spread throughout the genome.

Chromosome-specific partitioning of the data did not identify specific chromosomes with a greater significance pertaining to reproductive traits. However, it did facilitate the identification of loci associated with genomic areas under selection. Further investigation of these loci could provide valuable information on the genes and/or quantitative trait loci involved in an improved reproduction phenotype. A whole-genome analysis to identify signatures of selection could shed more light on the genomic regions involved in the reproduction traits of this divergently selected flock.

REFERENCES

Antao T., Lopes A., Lopes R.J., Beja-Pereira A. and Luikart G. (2008) *BMC Bioinformatic*s **9**: 323.

Beaumont M.A. and Nichols R.A. (1996) *Proc. R. Soc. Lond.* **263**: 1619.

Belkhir K., Borsa P., Goudet J., Chikhi L. and Bonhomme F. (2004) Laboratoire Génome, Populations, Interactions, CNRS UMR 5000. Université de Montpellier II; Montpellier.

Breen K.M., Billings H.J., Wagenmaker E.R, Wessinger E.W. and Karsch F.J. (2005) *Endocrinology* **146**: 210.

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.

Davis G.H. (2005) *Genet. Sel. Evol*. **37**: S11.

Dodds K.G., McEwan J.C. and Davis G.H., (2007) *Small Rumin. Res.* **70**: 32.

Foll M. and Gaggiotti O.E. (2008) *Genetics* **180**: 977.

Hayes B.J., Bowman P.J., Chamberlain A.J. and Goddard M.E. (2009) *J. Dairy Sci.* **92**: 433.

McNatty K.P., Galloway S.M., Wilson T., Smith P., Hudson N.L., O'Connell A., Bibby A.H., Heath D.A., Davis G.H., Hanrahan J.P. and Juengel J.L. (2007) *Genet. Sel. Evol.* **37**: S25.

Narum S.R. and Hess J.E. (2011) *Mol. Ecol. Resour.* **11**: 184.

Notter D.R. (2012) *Anim. Reprod. Sci.* **130**: 147.

Olivier J.J. (1999) *Proc. Assoc. Advmt. Anim. Breed. Genet.* **13**: 119.

Safari E., Fogarty N.M. and Gilmour A.R. (2005) *Livest. Prod. Sci.* **92**: 271.

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:102-105

CURRENT FLOCK EFFECTS ON LIFETIME REPRODUCTIVE PERFORMANCE OF SIMULATED SELECTION AT HOGGET AGE IN MERINO SHEEP, FOR FLEECE WEIGHT, FIBRE DIAMETER, BODY WEIGHT AND RELEVANT SELECTION INDEXES. III. HIGH RAINFALL REGION RESULTS.

L.R. Piper[1], A.A. Swan[2] and H.G. Brewer[1]

[1]CSIRO Animal, Food and Health Sciences, F D McMaster Laboratory, Locked Bag 1, Armidale, NSW 2350, [2]Animal Genetics and Breeding Unit, University of New England, Armidale, 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, litter size, lamb survival and reproduction rate have been examined for a medium-wool random breeding flock of Merino sheep grazing in the New England tablelands of NSW. Simulated single trait selection for increased fleece weight, fibre diameter, and body weight had significant positive effects on lifetime fertility ($P<0.05$; $P<0.01$; $P<0.01$ respectively), litter size ($P<0.001$; $P<0.001$; $P<0.001$ respectively) and reproduction rate ($P<0.001$; $P<0.001$; $P<0.001$ respectively)). There were no significant effects on lifetime reproduction rate or on any of the component traits, of simulated selection for either of the two selection indexes. Despite the substantial range in yearly mean reproduction rate (0.55 to 1.09), the selection group x lambing year effect was not significant for any combination of the reproduction and production traits and there was no tendency for the selection group differences in reproduction rate to increase or decrease over the observed range in mean reproduction rate. These data reinforce the findings from earlier papers in this series and do not support the view that selection for increased fleece weight may adversely affect lifetime reproduction rate.

INTRODUCTION

Adams *et al.* (2006) have suggested that reproductive performance may potentially be compromised in animals with increased capacity for wool production especially if feed supplies are reduced. To examine the genetic consequences of selection for increased fleece weight on reproduction rate, in an environment where feed supplies are regularly compromised, Piper *et al.* (2007) analysed data from long term fleece weight selection and control flocks of medium-wool Merino sheep grazing at Cunnamulla, south west Queensland. Rainfall at Cunnamulla averages 375 mm *per annum* but there is considerable variation and rainfall unreliability is the main factor limiting feed supply from pasture. In this environment, 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 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 the breeding objective, can be implemented without necessarily reducing reproduction rate."

To examine the effects on ewe lifetime reproduction rate of simulated phenotypic selection for wool and body traits, Piper *et al.* (2009) analysed data from a random mating flock grazing at Cunnamulla, Queensland. In this relatively harsh, semi-arid environment, there were no significant effects on lifetime reproduction rate or on any of the component traits (fertility, litter size, lamb survival), of simulated selection for fleece weight, fibre diameter or either of the selection indexes. Simulated 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 ($P=0.059$). In a follow up study, Piper *et al.* (2011) demonstrated that the effect of simulated selection for production traits on lifetime reproductive performance was not significantly influenced by variability in the available feed resources as assessed by the year to year variation in mean

reproduction rate. In this paper, we have applied the same analysis to a much larger data set collected on a medium Peppin Merino flock grazing in a high rainfall region at Armidale, NSW.

MATERIALS AND METHODS

Sheep. The reproductive performance of 2248 medium-wool, mixed Peppin origin, Merino ewes, first mated at around 18 months of age (mo) between 1970 and 1986, was analysed. There were 8293 records for fertility (ewes lambing/ewe joined, EL/EJ), 7135 for litter size (lambs born/ewe lambing, LB/EL), 7126 for lamb survival (lambs weaned/lamb born, LW/LB) and 8284 for reproduction rate (lambs weaned/ewe joined, LW/EJ). Details relating to the origin of the flock, to the environment and management of the flock at CSIRO's Longford Field Station, Armidale, NSW, have been given by Turner and Jackson (1978). The mating design for the flock consisted of 20 sire groups each of 20 mixed age ewes. Replacement sires were mated in the clockwise adjacent mating group to that in which they were born and ewe replacements were distributed at random to the mating groups with the restriction that mating of close relatives was avoided.

Observations and data analysis. Ten wool and body characteristics were measured on all animals using the techniques described by Turner *et al.* (1953). For these analyses, the data comprises measurements of greasy fleece weight (GFW), fibre diameter (FD), and body weight (BWT) taken from 12-14 mo ewes and the reproduction records (fertility, litter size, lamb survival, and reproduction rate) of the same ewes at their first five lambings (aged 2-6 years).

Allocation of ewes to High (H) and Low (L) selection groups for the production traits. As described by Piper *et al.* (2009), linear models adjusting for significant fixed effects were fitted using the statistical software R (R Development Core Team, 2011). 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 (H-L) for each trait is shown in Table 1.

Table 1. Predicted means for, and differences between the High and Low groups for GFW (kg), CFW (kg), FD (micron), BWT (kg), and Merino 7% and 14% indexes (M7 and M14)

	High (se)	Low (se)	H-L(se)	(H-L)/L*100
GFW	3.56 (0.01)	2.87 (0.01)	0.69 (0.01)	24.0
FD	22.01 (0.04)	19.60 (0.03)	2.41 (0.04)	12.3
BWT	35.39 (0.08)	30.17 (0.08)	5.21 (0.09)	17.3
M7	104.69 (0.14)	95.59 (0.14)	9.10 (0.15)	9.5
M14	105.97 (0.18)	94.41 (0.18)	11.56 (0.19)	12.2

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 lambing year, management group, lambing year x management group, birth type, age of dam (years), own age (years), selection group (High or Low) and lambing year x selection group all fitted as factors with ewe fitted as a random effect. Lambing year ($P<0.001$), management group ($P<0.001$ to $P<0.053$), lambing year x management group ($P<0.001$ to $P<0.057$) and own age ($P<0.001$ to $P< 0.002$) were significant or nearly so for all combinations of reproduction and production traits. Age of dam was significant ($P<0.05$) for each of the reproduction rate (LW/EJ), production trait combinations but not for any other combination of reproduction and production traits. Birth type was not significant for any combination of the reproduction and production traits.

RESULTS AND DISCUSSION

The predicted mean values for the High and Low groups for each production trait by reproduction trait combination are shown in Table 2. For GFW the difference between the high and low groups (H-L) was positive and significant for fertility (2.2%; $P<0.05$), litter size (4.4%; $P<0.001$) and reproduction rate (6.6%; $P<0.001$). For FD and BWT the difference between the high and low groups (H-L) was positive and significant for fertility (2.6%; 2.6%; $P<0.01$), litter size (3.2%; 7.0%; $P<0.001$) and reproduction rate (6.2%; 10.2%; $P<0.001$). There were no significant differences between the H and L groups for any of the production traits for survival and no differences between the H and L, M7 or M14 index groups for any of the reproduction traits. As found in Piper *et al.* (2009), the simulated selection for increased body weight produced a significant increase ($P<0.001$) in litter size but in the current study, it also produced a significant increase in fertility ($P<0.01$) and reproduction rate ($P<0.001$). Also in contrast to the results of Piper *et al.* (2009), there were positive and significant differences between the high and low GFW and FD groups for fertility, litter size and reproduction rate. These differences between the high and low groups for GFW and FD tend to balance each other in their contributions to the M7 and M14 indexes. This balancing effect probably accounts for the lack of difference between the high and low M7 and M14 groups for any of the reproduction traits.

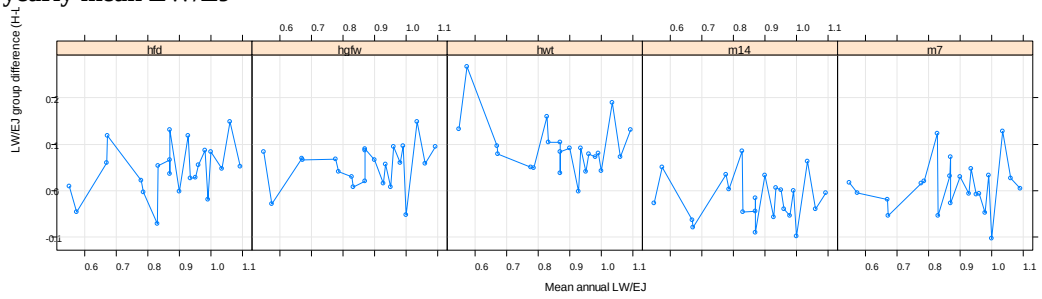
Table 2. **Predicted mean values (se) for the high (H) and low (L) groups for each production trait by reproduction trait combination**

	Fertility (EL/EJ)	Litter Size (LB/EL)	Survival (LW/LB)	Rep.Rate (LW/EJ)
GFW - H	0.883 (0.006) *	1.278 (0.009) ***	0.841 (0.006)	0.930 (0.011) ***
GFW - L	0.864 (0.006) *	1.225 (0.009) ***	0.847 (0.006)	0.872 (0.011) ***
FD - H	0.885 (0.006) **	1.272 (0.009) ***	0.846 (0.006)	0.928 (0.011) ***
FD - L	0.862 (0.006) **	1.232 (0.009) ***	0.842 (0.006)	0.874 (0.011) ***
BWT - H	0.885 (0.006) **	1.294 (0.009) ***	0.846 (0.006)	0.944 (0.011) ***
BWT - L	0.862 (0.006) **	1.209 (0.009) ***	0.842 (0.006)	0.857 (0.011) ***
M7 - H	0.880 (0.006)	1.258 (0.009)	0.838 (0.006)	0.904 (0.011)
M7 - L	0.868 (0.006)	1.247 (0.009)	0.851 (0.006)	0.898 (0.011)
M14 - H	0.870 (0.006)	1.245 (0.009)	0.840 (0.006)	0.889 (0.011)
M14 - L	0.877 (0.006)	1.259 (0.009)	0.848 (0.006)	0.914(0.011)

Significance of difference between high and low groups; * $P<0.05$; ** $P<0.01$: *** $P<0.001$; remainder, ns

The yearly mean LW/EJ ranged from 0.55 in 1976 to 1.09 in 1987. The differences in LW/EJ between the High and Low selection groups for each production trait in each year are shown in Figure 1 plotted against the yearly mean LW/EJ. There is clearly no tendency for the production trait differences in LW/EJ to increase or decrease as the mean LW/EJ moves from 0.55 to 1.09 and, despite the substantial range in mean LW/EJ, the lambing year x selection group effect was not significant for any combination of the reproduction and production traits.

Figure 1. Yearly production trait group differences (H-L) in LW/EJ plotted against the yearly mean LW/EJ



CONCLUSIONS

The results from Piper *et al.* (2009, 2011) on the phenotypic consequences of simulated selection for production traits on reproductive performance did 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. The results from this study, on a much larger flock grazing in a very different production environment, reinforce that conclusion. These current findings are again broadly 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.

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.

Piper L.R., Swan A.A and Brewer H.G. (2007) *Proc. Assoc. Advmt. Anim. Breed. Genet.* **17:** 103.

Piper L.R., Swan A.A and Brewer H.G. (2009) *Proc. Assoc. Advmt. Anim. Breed. Genet.* **18:** 374.

Piper L.R., Swan A.A and Brewer H.G. (2011) *Proc. Assoc. Advmt. Anim. Breed. Genet.* **19:** 335.

R Development Core Team (2011) 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.

Turner H.N. and Jackson N. (1978) *Aust. J. agric. Res.* **29:** 615.

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.

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:106-109

GENETIC ESTIMATES FOR ALONG AND ACROSS FIBRE DIAMETER VARIATION AND ITS USE TO IMPROVE STAPLE STRENGTH IN MERINO SHEEP

J.W.V Preston[1,2] and S. Hatcher[1,2]

[1]CRC for Sheep Industry Innovation, Homestead Building, UNE, Armidale, NSW, 2350
[2]NSW DPI, Orange Agricultural Institute, Forest Road, Orange, NSW, 2800

SUMMARY

Genetic parameters for a range of along and across fibre diameter traits measured from the Optical Fibre Diameter Analyser 2000 (OFDA2000) were estimated and correlated to staple strength and percentage of midbreaks. Fibre diameter, overall fibre diameter standard deviation, overall fibre diameter coefficient of variation, across fibre coefficient of variation, minimum fibre diameter along the fibre and maximum fibre diameter along the fibre were all highly heritable (h^2>0.30); however along fibre coefficient of variation was lowly heritable (h^2=0.07). Fibre diameter had a strong positive genetic correlation with maximum fibre diameter along the fibre and minimum fibre diameter along the fibre (≥0.86). Along fibre coefficient of variation had a low genetic correlation with across fibre coefficient of variation (0.22), which suggests that selecting for along fibre diameter distribution components will result in small changes in across fibre diameter distribution overtime. Staple strength had moderate negative genetic correlations with overall fibre diameter coefficient of variation (-0.59) and across fibre coefficient of variation (-0.51), while all other genetic correlations were low to negligible (<0.4). However genetic correlations between the percentage of midbreaks and across fibre coefficient of variation (0.45) and along fibre coefficient of variation (-0.40) had opposing direction of effect.

Therefore overall fibre diameter coefficient of variation had a greater heritability albeit lower observed variation and stronger correlation with staple strength than either the along or across fibre components. This would result in greater genetic gain when used in a breeding program to improve staple strength. However across fibre coefficient of variation proved to be useful in indirectly reducing the percentage of midbreaks and maybe valuable as a secondary trait to improve staple strength properties.

INTRODUCTION

OFDA2000 measures fibre diameter and splits the profile into along and across fibre diameter attributes. OFDA2000 has been utilised to monitor fibre diameter changes in response to environment (Gloag *et al.* 2004), aid clip preparation (Brien *et al.* 2001; Ferguson *et al.* 2002; Hansford *et al.* 2002) and to indirectly select for staple strength (Greeff 2002; Yamin *et al.* 1999). The importance of fibre diameter distribution to wool processing is considered significant as it affects the average fibre length in top (Lamb 2000), yarn evenness (Lamb 1992), fabric bending/rigidity (Degroot 1992) and yarn tenacity (Lamb 1992). Overall fibre diameter coefficient of variation has been used as a indirect selection criteria for staple strength and correlates to processing performance (Rottenbury *et al* 1983). The position of break in tender wool is important as fibre breakage in the middle of the staple has greater implications on wool processing. Therefore, overall fibre diameter coefficient of variation is included in breeding objectives (Piper and Lax 1992) and is used in current MERINOSELECT^{TM} breeding indexes. However there are few genetic estimates for the along and across fibre components of fibre diameter coefficient of variation measured by the OFDA2000. Providing an accurate estimation of the genetic parameters will allow the likely response to selection on staple strength and the percentage of midbreaks to be predicted. Other studies have reported high heritability estimates for along and across fibre components but with relatively high standard errors (Greeff 2002; Yamin *et al.* 1999). This paper

reports genetic parameters for along and across fibre traits and provides comment on the use of these traits to influence staple strength properties.

MATERIALS AND METHODS

Data was obtained from the Sheep CRC Information Nucleus Flock (INF) (Fogarty *et al.* 2007; van der Werf *et al.* 2010). The INF is comprised of eight diverse flocks located around Australia in geographically different locations. All sites were linked by common sires via the use of an extensive artificial insemination program. This paper presents results from 4,958 Merino progeny that were born between 2007-2010, drawn from 143 sires from a range of Merino wool types. Midside samples were collected at yearling age (10-13 months) and measured at a commercial fleece measurement laboratory (AWTA Limited Melbourne). OFDA2000 traits were measured by choosing one staple at random from the midside sample and cleaving it into a number of smaller micro staples. The micro staples where then placed on the OFDA2000 fibreglass xy slide; measured for fibre diameter and divided into it's along and across fibre components. The traits measured included minimum fibre diameter along the fibre (AMIN), maximum diameter along the fibre (AMAX), fibre diameter (FD), overall fibre diameter standard deviation (FDSD), overall fibre diameter coefficient of variation (FDCV), across fibre diameter coefficient of variation (ACCV) and along fibre diameter coefficient of variation (ALCV). For staple strength (SS) and percentage of mid breaks (MID), ten staples were chosen from the midside sample and measured using the automatic tester for length and strength (ATLAS) in accordance to IWTO 30 (2009).

ASReml 3.0 (Gilmour *et al.* 2009) was used to estimate the genetic parameters using general linear mixed and residual maximum likelihood methods. Initially, a univariate analysis of all traits included the following fixed effects: flock (8 sites), year of birth (2007, 2008, 2009, 2010), sex (male, female), dam age (2, 3, 4, 5, 6), birth type rearing type (born single raised single, born twin raised single, born and raised as a multiple) and siregroup (ultra/superfine, fine/fine-medium & medium/strong), with significant two way interactions. A sequence of models was fitted to each trait including varying combinations of random effects (i.e. effect of animal, sire.flock and overall maternal effect) and an effect to account for genetic groups (fitted as random or fixed). Genetic groups were allocated according to the back pedigree obtained from the data set. Ancestors with only 1 progeny were removed and groups with insufficient data merged. The genetic grouping accounted for the differing ewe foundation flocks at each of the sites and strain differences within the INF. The most appropriate model for each trait was determined by log likelihood ratio tests. Phenotypic and genetic correlations for each combination of traits with standard errors were estimated from bivariate analyses.

RESULTS AND DISCUSSION

Based on the log likelihood ratio test; fitting genetic groups as fixed was the most appropriate method for FD, FDSD, AMAX, AMIN, ALCV, ACCV and MID. FDCV and SS were best modelled with genetic groups as random. There was a significant sire.flock interaction (Table 1) for all traits. ALCV and SS were influenced by a maternal effect. Heritability estimates for FD, FDSD, FDCV, AMAX, AMIN, ACCV and SS were all high (>0.30), while ALCV and MID were lowly heritable (≤0.13). FD, FDCV and AMIN heritability estimates were consistent with previous reports; while AMAX was considerably higher (Greeff 2002; Yamin *et al.* 1999). ACCV and ALCV were lower (Greeff 2002; Yamin *et al.* 1999). The data recorded (amount and structure) and genetic diversity of the INF compared to the flocks studied by Greeff (2002) and Yamin *et al.* (1999) may have contributed to the differences in the estimates. Greeff (2002) recorded progeny bred from sires of Collinsville, Peppin, and Bungaree Merino families (12 studs, 100 sires); while Yamin *et al.* (1999) recorded progeny bred from Collinsville and Bungaree Merino families (4 studs, 47 sires). The phenotypic variation of fibre diameter distribution traits estimated in our

study was similar to that of Yamin *et al.* (1999); though lower than estimated by Greeff (2002).

Table 1. Variance components, coefficient of variation (CV %) and heritability (h^2) for yearling OFDA2000 and staple strength properties

Trait	Mean	Variance components					CV	Heritability
		Phenotypic	Residual	Additive	Sire.flock	Maternal	(%)	h^2
FD	17.15	1.54	0.39	1.12	0.03	-	8.98	0.73±0.05
FDSD	3.16	0.14	0.06	0.07	0.00	-	4.43	0.52±0.05
FDCV	18.46	3.09	1.46	1.53	0.10	-	16.73	0.49±0.05
AMAX	18.78	2.11	0.62	1.43	0.06	-	11.24	0.68±0.05
AMIN	15.80	1.38	0.51	0.85	0.02	-	8.73	0.62±0.05
ACCV	20.42	5.85	3.80	1.85	0.20	-	28.65	0.32±0.04
ALCV	5.69	2.58	2.14	0.19	0.11	0.13	45.34	0.07±0.03
SS	31.84	85.93	50.72	28.72	2.43	4.32	269.59	0.33±0.06
MID	62.59	771.24	641.54	103.64	26.06	-	1232.21	0.13±0.03

There were strong positive genetic correlations between FD and FDSD, AMIN and AMAX (≥0.64) (Table 2) which suggests that as fibre diameter becomes finer, the maximum and minimum diameter along the fibre will also become finer and result in less fibre diameter variation. This is in agreement with Greeff (2002) and Yamin *et al.* (1999). ACCV and ALCV both had moderate to strong positive genetic correlations with FDSD and FDCV (≥0.42). ALCV and ACCV had low positive genetic correlation between each other and indicates that they are controlled by different genes. The unfavourable negative genetic correlation between ACCV with both AMIN (-0.37) and AMAX (-0.25) indicates that as fibre diameter distribution across the fibre is less variable; the maximum and minimum diameter along the fibre will be broader.

Table 2. Phenotypic correlations (below diagonal) and genetic correlations (above diagonal)

	FD	FDSD	FDCV	AMAX	AMIN	ACCV	ALCV
FD		0.64±0.04	-0.09±0.06	0.99±0.00	0.86±0.02	-0.33±0.07	0.30±0.09
FDSD	0.58±0.01		0.70±0.03	0.70±0.03	0.56±0.05	0.42±0.06	0.57±0.08
FDCV	-0.05±0.02	0.77±0.01		0.00±0.06	-0.18±0.07	0.42±0.06	0.40±0.09
AMAX	0.93±0.00	0.65±0.01	0.09±0.02		0.95±0.01	-0.25±0.07	0.41±0.09
AMIN	0.88±0.00	0.43±0.01	-0.17±0.02	0.78±0.01		-0.37±0.07	0.15±0.10
ACCV	-0.21±0.02	0.41±0.01	0.67±0.01	-0.09±0.02	-0.18±0.02		0.22±0.12
ALCV	0.03±0.02	0.35±0.01	0.41±0.01	0.32±0.01	-0.30±0.01	0.13±0.02	

Phenotypic correlations for SS and MID with along and across fibre components were shown to have a similar direction of effect as genetic correlations but in smaller magnitude (Table 3). Negative genetic correlations were estimated between both SS and FDCV (-0.59) and SS and ACCV (-0.51). All other correlations were low to negligible (<0.40). MID had a moderate positive genetic correlation with ACCV (0.45) and an antagonistic moderate negative correlation with ALCV (-0.40). Therefore selecting for low ACCV or low FDCV would result in improvements in both SS and the percentage of midbreaks, while selection for low ALCV would result in a decrease in SS and a higher percentage of midbreaks.

Table 3. Phenotypic (r_p) and genetic (r_g) correlations between staple strength (SS), percentage of midbreaks (MID) and along and across fibre diameter distribution traits

Trait	SS		MID	
	r_p±SE	r_g±SE	r_p±SE	r_g±SE
FD	0.16±0.02	0.19±0.07	-0.14±0.02	-0.23±0.10
FDSD	-0.25±0.02	-0.32±0.07	-0.04±0.02	0.04±0.11
FDCV	-0.43±0.01	-0.59±0.06	0.06±0.02	0.33±0.11
AMAX	0.08±0.02	0.14±0.07	-0.17±0.02	-0.25±0.10
AMIN	0.23±0.02	0.26±0.07	-0.07±0.02	-0.16±0.10
ACCV	-0.29±0.02	-0.51±0.07	0.11±0.02	0.45±0.11
ALCV	-0.20±0.02	-0.21±0.12	-0.16±0.02	-0.40±0.13

The high heritability of FDCV (0.49) and moderate favourable correlations with SS confirms Greeff's (2002) conclusion that it is unnecessary to further divide this trait into its along and across fibre components when using FDCV as an alternative selection criteria for staple strength. However selection using ACCV reduces the percentage of midbreaks and therefore it would be beneficial to include ACCV in a breeding program perhaps as a secondary trait. Further work is required to evaluate the usefulness of the current staple strength measurement in a breeding program. Measurement of SS via the ATLAS was designed to provide a prediction of fibre breakage during processing and was not intended for use to compare the of rank animals to achieve on-farm genetic improvement (Semmel 2003). Other measurements that describe the shape of strength vs extension curve such as "staple specific work to rupture" produced by the ATLAS may provide a more efficient method for selection for SS.

ACKNOWLEDGEMENTS

We would like to thank all staff involved at the INF sites for the work involved in the collection of the data used in this paper.

REFERENCES

Brien F.D., Ferguson M.B. and Konstantinov K. (2001) *Proc. Assoc. Advmt. Anim. Breed. Genet.* **14**: 277.

Degroot G. (1992) *Wool Tech. Sheep Breed.* **40:** 60.

Ferguson M.B., Gloag C.M., Behrendt R. and Brien F.D. (2002) *Wool Tech. Sheep Breed.* **50**: 787.

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

Gilmour A., Gogel B., Cullis B. and Thompson R. (2009) ASReml User Guide Release 3.0.

Gloag C., Kearney G. and Behrendt R. (2004) *Aust. Soc. Anim. Prod.* **25:** 249

Greeff J.C. (2002) *Wool Tech. Sheep Breed.* **50**: 11.

Hansford K.A., Marler J.W. and McLachlan I.M. (2002) *Wool Tech. Sheep Breed.* **50**: 140.

IWTO 30. (2009) *IWTO* 1-30.

Lamb P.R. (2000) *IWTO* Christchurch meeting (Report CTF 02),

Lamb P.R. (1992) *Wool Tech. Sheep Breed.* **40**: 65.

Piper L. and Lax. J. (1992) *Wool Tech. Sheep Breed.* **40**: 10.

Rottenbury R.A., Andrews M.W. and Brown G.H. (1983) *Text. Res. J.* **53**: 29.

Semmel P.J. (2003) *IWTO* Instanbul meeting (Report RWG 04).

van der Werf J., Kinghorn B. and Banks R. (2010) *Anim. Prod. Sci.* **50**: 998.

Yamin M., Hynd P.I., Ponzoni R.W., Hill J.A., Pitchford W.S. and Hansford K.A. (1999) *Wool Tech. Sheep Breed.* **47**: 151.

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:110-103

GENETIC CORRELATIONS ACROSS AGES FOR GREASY FLEECE WEIGHT AND FIBRE DIAMETER IN MERINO SHEEP

D.J. Brown[1], A.A. Swan[1] and J.S. Gill[2]

[1]Animal Genetics and Breeding Unit[*], University of New England, Armidale, NSW 2351
[2]Meat and Livestock Australia, University of New England, Armidale, NSW 2351

SUMMARY

While adult wool production is a key determinant of profitability of Merino flocks, much of the on-farm recording through MERINOSELCT is focused on assessments at 12 or 18 months of age. While there have been numerous studies in research flocks examining these relationships these results have not been validated in industry flocks. The aim of this paper was to investigate the genetic correlations between multiple age expressions of fibre diameter and greasy fleece weight using the MERINOSELECT database.

The results support earlier research in that, while the correlations across ages are generally very high they also suggest that at least one adult assessment of fleece weight and fibre diameter would be beneficial for the breeding program. As a result of this work the genetic evaluation for Sheep Genetics will be modified to include annual expressions of adult fleece traits.

INTRODUCTION

The correlations between early age performance (12 and 18 months of age) and adult performance have been examined in numerous research studies (Atkins 1990; Atkins and Mortimer 1987; Coelli *et al.* 1998; Hickson *et al.* 1994; Fozi *et al.* 2012). These studies all suggest that the correlations between early age measurements and adult expression are moderate to high (r_g=0.60 to 0.90). Furthermore correlations between adult expressions are very high and can be treated as repeated expressions of the same trait. This has been the approach currently adopted by Sheep Genetics in the MERINOSELECT analysis (Brown *et al.* 2007), although most industry breeders choose not to record animals at an adult age.

The aim of this study was to evaluate the value of recording fleece weight and fibre diameter throughout the animal's lifetime by estimating genetic correlations between multiple age expressions of these traits from the MERINOSELECT database.

MATERIALS AND METHODS

Data. Pedigree and performance data were extracted from the Sheep Genetics MERINOSELECT database (Brown *et al.* 2007). This database consists of pedigree and performance records submitted by Australian and New Zealand Merino ram breeders which are used for genetic evaluation purposes.

As the complete database was too large for parameter estimation analyses, a subset of 78 flocks were used based on their recording of adult wool traits. These flocks are a mix of industry ram breeders, research and sire evaluation flocks. Within these flocks all animals with at least sire pedigree and born from 1990 and later were included. Data were extracted for all greasy fleece weight (GFW) and fibre diameter (FD) records from these flocks. Records were classified to one of 7 age based traits, yearling (Y, 12 months), hogget (H, 18 months) and 2 year old adult (A2) through to 6 year old adult records (A6). All contemporary groups were transformed to a common mean within each group for both greasy fleece weight and fibre diameter, as is done routinely for

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

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 all ancestral information available. This resulted in pedigree files comprising 243,996 and 240,989 animals and data files comprising 337,846 and 338,909 records from 204,540 and 199,663 animals recorded at least once for greasy fleece weight and fibre diameter respectively.

Table 1. Summary of the data used in this study

	Fibre diameter (micron)					Greasy fleece weight (kg)				
Age	**Count**	**Mean**	**SD**	**Min**	**Max**	**Count**	**Mean**	**SD**	**Min**	**Max**
Y	125,810	17.5	1.2	12.6	25.5	127,291	3.5	0.60	0.7	7.0
H	84,610	19.0	1.4	13.1	27.0	71,524	4.5	0.67	1.4	8.0
A2	92,230	19.5	1.4	13.8	29.3	94,988	5.5	0.75	1.9	11.6
A3	17,498	19.5	1.5	14.7	27.5	21,309	5.5	0.82	1.9	10.2
A4	11,312	19.5	1.6	14.5	27.3	13,802	5.5	0.83	2.6	9.4
A5	6,756	19.5	1.6	14.2	27.2	8,680	5.5	0.83	2.4	9.4
A6	693	19.5	1.2	16.5	23.0	252	5.5	0.82	3.5	7.5

In this dataset approximately 47% and 53% of the animals studied had multiple records for fleece weight and fibre diameter respectively (Table 2). On average animals had 1.7 age expressions recorded for each trait. Furthermore 91% and 94% of the sires had progeny recorded across multiple age expressions for fleece weight and fibre diameter respectively.

Table 2. Number (%) of the animals by number of records observed per animal and number (%) of sires by the number of ages at which they have progeny were recorded

		Number of records per animal or per sire						
		1	**2**	**3**	**4**	**5**	**6**	**7**
GFW	**animals**	108,785 (53)	73,650 (36)	11,225 (5)	6,376 (3)	4,442 (2)	62 (<1)	0 (0)
	sires	414 (9)	1,730 (39)	888 (20)	339 (8)	623 (14)	461 (10)	9 (<1)
FD	**animals**	94,022 (47)	86,352 (43)	9,659 (5)	5,467 (3)	3,744 (2)	315 (<1)	104 (<1)
	sires	274 (6)	1,721 (41)	945 (22)	269 (6)	433 (10)	559 (13)	43 (1)

Models of analysis. Parameters were estimated in multivariate animal model analyses including all 7 age expressions for each trait in ASReml (Gilmour *et al.* 2006). For both traits the fixed effects of contemporary group, birth type, rearing type, age of dam, and animal's age at measurement were fitted. Contemporary group was defined as flock, year of birth, sex, date of measurement, management group subclass. A single random term for the direct genetic effects was modelled.

RESULTS AND DISCUSSION

The phenotypic variance for greasy fleece weight increases with age (Table 3) while the heritability increased from 0.37 at yearling to 0.51 at 2 year old adult age and then plateaued thereafter. The genetic correlations between yearling and adult performance were moderate to high ranging from 0.81 at 2 years to 0.62 at 6 years of age. Hogget traits were more highly correlated with adult expressions ranging from 0.88 to 0.73. The results confirm that assessments for greasy fleece weight made on young animals are good predictors of adult performance genetically (0.62

to 0.81) even if the phenotypic correlations are lower (0.44 to 0.61). Adult expressions were all very highly correlated and can be treated as the same trait genetically. The phenotypic correlations were generally moderate to high and also slightly lower than the genetic correlations.

Table 3. Phenotypic variance (σ^2_p), direct (h^2) heritability, genetic (below) and phenotypic (above) correlations for greasy fleece weight (s.e. in parentheses)

	Y	H	A2	A3	A4	A5	A6
σ^2_p	0.32 (0.00)	0.39 (0.00)	0.56 (0.00)	0.63 (0.01)	0.63 (0.01)	0.65 (0.01)	0.72 (0.05)
h^2	0.37 (0.01)	0.48 (0.01)	0.51 (0.01)	0.48 (0.01)	0.50 (0.02)	0.49 (0.02)	0.44 (0.10)
Y	.	0.59 (0.00)	0.61 (0.00)	0.47 (0.01)	0.46 (0.01)	0.44 (0.01)	0.47 (0.05)
H	0.84 (0.01)	.	0.65 (0.00)	0.63 (0.01)	0.63 (0.01)	0.58 (0.01)	0.63 (0.09)
A2	0.81 (0.01)	0.88 (0.01)	.	0.67 (0.00)	0.69 (0.00)	0.66 (0.01)	0.75 (0.03)
A3	0.71 (0.02)	0.82 (0.01)	0.93 (0.01)	.	0.70 (0.00)	0.69 (0.01)	0.66 (0.04)
A4	0.67 (0.02)	0.79 (0.01)	0.92 (0.01)	0.95 (0.01)	.	0.72 (0.01)	0.72 (0.04)
A5	0.66 (0.02)	0.77 (0.02)	0.90 (0.01)	0.94 (0.01)	0.97 (0.01)	.	0.77 (0.03)
A6	0.62 (0.12)	0.73 (0.15)	0.82 (0.09)	0.90 (0.09)	0.96 (0.09)	0.90 (0.08)	.

The phenotypic variance for fibre diameter also increased with age (Table 4) while the heritability increased from 0.60 at yearling to 0.68 at 3 year old adult age and then plateaued thereafter. The genetic correlations between yearling and adult performance were moderate to high ranging from 0.92 to 0.74 at 6 years of age. Hogget traits were more highly correlated with adult traits ranging from 0.91 to 0.79. The results again confirm that assessments of key fleece traits made on young animals are good genetic predictors of adult performance. The adult expressions were all very highly correlated and can be treated as the same trait genetically. The phenotypic correlations were generally moderate to high and slightly lower than the genetic correlations.

Table 4. Phenotypic variance (σ^2_p), direct (h^2) heritability, genetic (below) and phenotypic (above) correlations for fibre diameter (s.e. in parentheses)

	Y	H	A2	A3	A4	A5	A6
σ^2_p	1.34 (0.01)	1.70 (0.01)	1.74 (0.01)	1.78 (0.02)	1.98 (0.02)	2.12 (0.03)	2.56 (0.11)
h^2	0.60 (0.01)	0.61 (0.01)	0.65 (0.01)	0.68 (0.02)	0.66 (0.02)	0.66 (0.03)	0.67 (0.08)
Y		0.71 (0.00)	0.67 (0.00)	0.60 (0.01)	0.59 (0.01)	0.56 (0.01)	0.58 (0.02)
H	0.92 (0.01)		0.71 (0.00)	0.67 (0.00)	0.66 (0.01)	0.61 (0.01)	0.59 (0.02)
A2	0.85 (0.01)	0.91 (0.01)		0.76 (0.00)	0.75 (0.00)	0.71 (0.01)	0.71 (0.02)
A3	0.79 (0.01)	0.87 (0.01)	0.96 (0.01)		0.77 (0.00)	0.75 (0.01)	0.74 (0.02)
A4	0.78 (0.01)	0.86 (0.01)	0.94 (0.01)	0.96 (0.01)		0.79 (0.00)	0.80 (0.01)
A5	0.75 (0.01)	0.81 (0.01)	0.91 (0.01)	0.94 (0.01)	0.98 (0.01)	.	0.84 (0.01)
A6	0.74 (0.03)	0.79 (0.03)	0.87 (0.03)	0.92 (0.02)	0.98 (0.02)	0.98 (0.02)	.

These results suggest that measurement of at least one adult expression would improve the accuracy of selection for lifetime wool production and value. A companion paper in these proceedings (Swan and Brown 2013) used the genetic parameters presented here and estimated the trait and economic gains that can be achieved for various combinations of measurements across ages and as well as incorporating genomic selection. These results confirm that recording at least one adult assessment produced significantly greater trait and economic gain for both traits. Furthermore genomic selection also increased the progress in both fleece weight and fibre diameter.

Additional analysis of the breeding values of sires from these analyses shows that despite the

very high correlation between traits some sires have breeding values that either increase or decrease over time. Thus breeders with concerns about changes in fleece value across age should be encouraged to record annual fleece value traits. The genetic evaluation for Sheep Genetics will be modified to include annual expressions of adult fleece traits.

These were preliminary analyses, so they ignored maternal effects (genetic and environmental), genetic group effects and the effects of previous and current physiological state which are known to affect wool production (Hinch *et al.* 1996; Huisman and Brown 2009). It is likely that accounting for these effects would further improve the correlations between traits recorded across different ages.

There is also a general lack of recording of liveweight at adult ages in the MERINOSELECT database. With the increasing focus on mature weight of sheep due to maintenance costs, welfare and occupation health and safety concerns this appears to be an opportunity for breeders to record this trait and increase the focus on this trait in the breeding program.

It is significant to note that this industry data set was/is large enough, and contains sufficient recording of fixed effects, to support very accurate genetic parameter estimation.

CONCLUSIONS

These preliminary estimates from industry data are consistent with those in the literature and reconfirm that assessments made on young animals are good genetic predictors of adult performance. However, genetic correlations do support the need for breeders to collect at least one adult assessment. The adult expressions were all very highly correlated and can be treated as the same trait genetically and support the model currently employed by Sheep Genetics. Additional data and analyses are required to investigate other lifetime traits such as live weight, fertility and wool quality traits.

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. The authors acknowledge the contributions of the Sheep CRC Information Nucleus, the Australia Merino Sire Evaluation Association and industry-funded research flocks.

REFERENCES

Asadi Fozi, M., van der Werf J.H.J. and Swan A.A. (2012) *Anim. Prod. Sci.* **51**: 1019.

Atkins K.D. and Mortimer S.I. (1987) *Proc. Ass. Adv. Anim. Breed. Genet.* **6**: 79.

Atkins K.D. (1990) *Proc. 4th World Cong. on Genet. App. to Liv. Prod.* **15**: 17.

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.

Coelli K.A., Gilmour A.R. and Atkins K.D. (1998) *Proc. of 6th World Cong. on Genet. App. to Liv. Prod.* **24**: 31.

Gilmour A.R., Gogel B.J., Cullis B.R., and Thompson R. (2009) ASReml User Guide Release 3.0 VSN International Ltd, Hemel Hempstead, HP1 1ES, UK.

Hickson J., Kinghorn B.P., Swan A.A., and Piper L.R. (1994) *Proc. 5th World Cong. on Genet. App. to Liv. Prod.* **18**: 139.

Hinch G.N., Lynch J.J., Nolan J.V., Leng R.A., Bindon B.M. and Piper LR (1996) *Aust. J. Exp. Agric.* **36:** 129.

Huisman A.E. and Brown D.J. (2009) *Anim. Prod. Sci.* **49**: 289.

Swan A.A. and Brown D.J. (2013) *Proc. Assoc. Advmt. Anim. Breed. Genet.* **20:** 232

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:114-117

GENETIC AND NON-GENETIC EFFECTS ON FLIGHT SPEED AND AGITATION IN WEANED LAMBS

C.L. Dodd[1,2], J.E. Hocking Edwards[1,3], S.J. Hazel[2] and W.S. Pitchford[2]

[1]Cooperative Research Centre for Sheep Industry Innovation
[2]School of Animal and Veterinary Sciences, The University of Adelaide, Roseworthy SA 5371, Australia
[3]South Australian Research Development Institute, Struan Research Centre, Naracoorte SA 5271

SUMMARY

By identifying sheep with a genetic or environmentally-induced propensity for stress, it is possible to manage or select against those sheep to minimise stress and improve the welfare and ease of handling of the entire flock. In this study we used established behavioural measures, including flight speed and agitation scores. However, behavioural measurements can be difficult and time consuming. Therefore we also assessed the possibility of using facial wool cover, a subjective score already used by industry, as an indicator for behavioural reactivity in sheep.

This study investigated the 2008-2010 cohorts of the Information Nucleus. Eight flocks totaling 11,047 lambs were tested. Flight speed and agitation were measured at 2-6 weeks post-weaning. Lambs were assigned face cover scores at 8 months of age.

Low to moderate heritability estimates of flight speed (0.11±0.02) and agitation (0.20±0.02) indicate that while there is an inherent component to behaviour as measured in these tests, that component is small. A moderate genetic correlation was found between flight speed and agitation (0.19±0.10), though the phenotypic correlation was low. Heavier and female lambs were more reactive than lighter and male lambs in both behavioural tests. Terminal sire x Merino cross lambs were faster in the flight speed test than other types (pure Merino or second cross). In one flock, younger lambs were more reactive in the agitation test. The two behavioural traits varied independently such that flocks with high average flight speeds did not necessarily have high average agitation scores.

Face cover score was highly heritable (0.39±0.03), similar to earlier work. Phenotypic and genetic correlations between behaviours and face cover were low, indicating face cover will not be a useful indicator for these behavioural tests.

INTRODUCTION

Behavioural reactivity is an animal's behavioural response to stress. This is underlain by a pattern of neuro-endocrine system responses which may be controlled by genetics, and permanent and temporary environmental effects. By identifying sheep with a genetic or environmentally-induced propensity to react negatively to stress, it is possible to manage those sheep to minimise stress, or to select against them during breeding to improve the welfare and ease of handling of the entire flock (Burrow 1997).

Successful divergent breeding of sheep based on apparent differences in behaviour has been achieved, indicating that behaviour is heritable (Beausoleil *et al.* 2012). Identification of behavioural measures that are correlated with breeding objectives and that are heritable could result in the development of genetic improvement programs to reduce sheep reactivity to handling, enhancing flock welfare and ease of handling (Ponzoni and Newman 1989).

It is likely that non-genetic factors also affect the expression of behavioural reactivity. Agitation and flight speed are objective behavioural tests which have been used in Australian sheep research. Previous estimates of heritability for these traits were low to moderate (Blache and

Ferguson 2005; Hocking Edwards *et al.* 2011; Plush *et al.* 2011), supporting the hypothesis that flight speed and agitation are also influenced by non-genetic factors, such as sex, age and breed.

Measuring behaviour is difficult and time consuming. An indicator trait which is easy to assess, correlated to the behavioral trait of interest and heritable would make selection much easier. Facial hair patterning may be a potential candidate. There is evidence in cattle that facial hair patterning is related to behaviour, with associations found between the position of the facial hair whorl and agitation during restraint and handling (Olmos and Turner 2008). Hair whorl position has also been related to behaviour in other species (Tomkins *et al.* 2012). Sheep do not exhibit a facial hair whorl. However, a similar trait, face cover, is measured routinely in sheep.

This study aimed to estimate the heritability of agitation and flight speed, assess the impacts of non-genetic factors on these behaviours and evaluate the usefulness of face cover as an indicator for behaviour in lambs

MATERIALS AND METHODS

This study investigated the 2008-2010 cohorts of the Cooperative Research Centre for Sheep Industry Innovation's "Information Nucleus" lambs. The full structure of this flock has been described by Fogarty *et al.* (2007).

Flight speed and agitation were measured on lambs at two to six weeks after weaning. Flight speed was the average speed at which a lamb crosses a 1.7m distance (Burrow 1997), measured using infra-red start and stop beams attached to a timer. Flight speed was measured as the lamb exited the weigh crate. Agitation was measured using an isolation test. The lamb was restrained within a fully enclosed box with dimensions 1.5m x 0.7m x 1.5m. Vibrations caused by movement and vocalisation of the lamb over 30 seconds were measured (Plush *et al.* 2011).

Face cover is a subjective score of the amount of wool on the face (*Visual Sheep Scores* 2007). Scores range from 1: open face with no wool in front of the ears and top knot or on the jowls, to 5: heavy wool growth over the entire face with wool from the top and side of the muzzle joining. Animals are scored at four months of age or older.

Data. In 2008 a total of 3992 lambs were measured, with 3841 lambs in 2009, and 3214 lambs in 2010. Records for each of these animals included behavioural measures, weights and demographics such as age, sire and dam breeds, management group, flock and year of birth. Across the three years 5599 males and 5440 females were measured. A pedigree was available with up to three generations of data plus source genetic groups for most animals.

Analysis. A linear mixed animal model was fitted to the behaviours using ASReml (Gilmour *et al.* 2009). The model contained fixed effects of flock, sex, lamb age (nested within flock), birth-rearing type (11, 21, 22, 31, 32, 33), lamb breed (Merino, Maternal x Merino, Terminal x Maternal, Terminal x Merino), weaning weight and faecal worm egg count. No interactions were significant. Animal (pedigree), management group (within flock and drop) and sire (within flock) were fitted as random terms. This model was used to calculate variances and heritabilities for each trait, in addition to probabilities for the fixed effects. The model was then fitted as a bivariate to estimate phenotypic and genetic correlations between flight speed, agitation and face cover.

RESULTS AND DISCUSSION

The low to moderate heritability estimates of flight speed and agitation (Table 1) indicate that while there is an inherent component to behaviour as measured in these tests, that component is small. The values found here are in agreement with previous estimates (Hocking Edwards *et al.* 2011; Plush *et al.* 2011), but are lower than that suggested by Blache and Ferguson (2005).

Table 1. Variance (on the transformed scale) explained by each of the random effects as proportion of total variance, and heritability estimates for flight speed, agitation and face cover.

	Flight speed	Agitation	Face cover
Management group	0.035 (22%)	1.11 (20%)	0.081 (19%)
Genetic (animal)	0.014 (9%)	0.86 (16%)	0.139 (32%)
Genetic x Environmental (sire x site)	0.005 (3%)	0.11 (2%)	0.030 (7%)
Residual	0.107 (66%)	3.42 (62%)	0.186 (43%)
Heritability estimate	0.11±0.02	0.20±0.02	0.39±0.03

Flight speed and agitation were poorly correlated with each other phenotypically, though moderately genetically correlated (Table 2). This is supported by previous analyses of these tests in related sheep (Hocking Edwards *et al.* 2011; Plush *et al.* 2011), and suggests either that these are measuring different aspects of behavioural reactivity, or that one or both of these are poor measures of behavioural reactivity.

Table 2. Correlations between flight speed, agitation and face cover.

	Flight speed x Agitation	Agitation x Face cover	Flight speed x Face cover
Management group	0.23 ± 0.13*	0.09 ± 0.15	0.25 ± 0.13*
Genetic	0.19 ± 0.10*	-0.06 ± 0.07	-0.10 ± 0.08*
Genetic x Environment	0.15 ± 0.15	0.03 ± 0.13	-0.45 ± 0.10*
Residual	0.03 ± 0.02*	0.03 ± 0.03	-0.01 ± 0.03
Phenotypic	0.06 ± 0.01*	0.0004 ± 0.01	-0.05 ± 0.01*

* = correlation significantly different from zero

Heavier lambs were more reactive in both behavioral tests. Previous studies had mixed results (Amdi *et al.* 2010; Horton and Miller 2011). Hyper-responsiveness of the hypothalamic-pituitary adrenal (HPA) axis has been demonstrated in individuals with increased adipose tissue in sheep, rats and humans, suggesting a true physiological link between adiposity and stress response (Tilbrook and Clarke 2006).

Female lambs were more reactive in the agitation test than were males. This may indicate increased fear, or higher social motivation of females (Boissy *et al.* 2005). Several studies have found evidence that there are functional differences between sexes in the HPA axis in sheep (Hernandez *et al.* 2010).

Terminal x Merino first-cross lambs had faster flight speeds than the three other pure, first-cross and second-cross lamb types. Breed differences in behaviour have been found in several studies (Boissy *et al.* 2005) with some breeds displaying an active coping mechanism (high levels of locomotion, high-pitched bleats, escape attempts) and others a passive mechanism (immobilisation, quiet bleating, retreat from stimuli).

Younger lambs were more active in the agitation test than older lambs in a single flock (Katanning), similar to Viérin and Bouissou's (2003) work in which 3-4 month old lambs were more fearful than 5-6 month olds. Age and experience are difficult to separate, and although lambs in this study were tested at a young age (2-5 months), habituation to handling by humans may have contributed to the lack of significance across flocks for this effect.

There were significant but ambiguous effects of flock on both of the behavioural tests, with the two behaviours apparently varying independently of one another. Given that the eight flocks were chosen specifically to represent the diversity of sheep production across Australia, this effect may

be due to a variety of factors, including weather, day length, facilities and handling style. Faecal worm egg count and birth-rearing type were not associated with either behavioural measure.

Face cover score was highly heritable, yielding similar values to previous studies (Mortimer *et al.* 2009). However, lack of correlation with either behavioural score indicates that face cover will not be a useful indicator for behavioural reactivity. The present study was opportunistic in utilising a face cover score designed to assess the risk of a sheep for wool blindness, rather than a measure designed to describe patterning. It is possible a relationship exists between facial hair patterning and behaviour in sheep, as demonstrated in cattle (Olmos and Turner 2008). More descriptive measures of facial wool patterning may be useful as indicators of inherent sheep behaviour.

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 by NSW DPI, the University of New England, Vic DPI, SARDI and DAFWA. The authors gratefully acknowledge the contributions of the many staff involved with the Information Nucleus program. This program has been a very large collaborative effort involving teams of scientists and technical officers from 7 different research agencies working at 9 Information Nucleus flock sites, 7 abattoirs, and 7 laboratories across Australia. Individual staff are not listed simply because of the number involved, but their contributions are duly acknowledged. Participation by Sheep Genetics and data from the former Sheep Genomics Program is integral. Development and maintenance of the IN database is supported by AGBU.

REFERENCES

Amdi C., Williams A., Maloney S., Tauson A., Knott S. and Blache D. (2010) *Anim. Prod. Sci.* **50**: 683

Beausoleil N., Blache D., Stafford K., Mellor D. and Noble A. (2012) *Appl. Anim. Behav. Sci.* **139**: 74

Blache D. and Ferguson D. (2005) Meat & Livestock Australia.

Boissy A., Bouix J., Orgeur P., Poindron P., Bibé B. and Le Neindre P. (2005) *Genet. Sel. Evol.* **37**: 381

Burrow H. (1997). *Animal Breeding Abstracts* **65**: 477

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

Gilmour A., Gogel B., Cullis B. and Thompson R. (2009). VSN International Ltd, Hemel Hemptead, UK.

Hernandez C., Matthews L., Oliver M., Bloomfield F. and Harding J. (2010). *Physiol. Behav.* **101**: 588

Hocking Edwards J., Brien F., Hebart M., Hinch G., Hoad J., Hart K., Gaunt G., Robertson M., Refshauge G. and Bird-Gardiner T. (2011). *Proc. Assoc. Advmt. Anim. Breed. Genet.* **19**: 163

Horton B. and Miller D. (2011) *Appl. Anim. Behav. Sci.* **129**: 74

Mortimer S., Robinson D., Atkins K., Brien F., Swan A., Taylor P. and Fogarty N. (2009) *Anim. Prod. Sci.* **49**: 32

Olmos G. and Turner S. (2008) *Appl. Anim. Behav. Sci.* **115**: 25

Plush K., Hebart M., Brien F. and Hynd P. (2011) *Appl. Anim. Behav. Sci.* **134**: 130

Ponzoni R. and Newman S. (1989). *Anim. Prod.* **49**: 35

Tilbrook A. and Clarke I. (2006) *Front. Neuroendocrin.* **27**: 285

Tomkins L., Thomson P. and McGreevy P. (2012) *Vet. J.* **192**: 359

Viérin M. and Bouissou M. (2003) *Dev. Psychobiol.* **42**: 131

Visual Sheep Scores (2007) Australian Wool Innovation Limited, Meat and Livestock Australia

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:118-121

GENETIC ESTIMATES FOR ALONG AND ACROSS FIBRE DIAMETER VARIATION AND CORRELATIONS WITH SUBJECTIVE WOOL QUALITY TRAITS IN MERINO SHEEP

J.W.V Preston[1,2], S. Hatcher[1,2]

[1]CRC for Sheep Industry Innovation, Homestead Building, UNE, Armidale, NSW, 2350
[2]NSW DPI, Orange Agricultural Institute, Forest Road, Orange, NSW, 2800

SUMMARY

Genetic and phenotypic correlations were estimated for along and across fibre diameter components measured by the Optical Fibre Diameter Analyser 2000 (OFDA2000) and a range of subjectively assessed wool quality scores. Results demonstrated that greasy wool handle, wool staple structure and wool character were strongly genetically correlated with fibre diameter, overall fibre diameter standard deviation, minimum fibre diameter along the fibre and maximum fibre diameter along the fibre (0.51-0.68). Genetic correlations for across fibre co-efficient of variation with subjective wool quality scores were all low to negligible (<0.40). However, along fibre co-efficient of variation had a moderate favourable genetic correlation with staple structure (0.47±0.11) and greasy wool handle (0.38±0.13). All other correlations were low to negligible. Correlations for fibre diameter, overall fibre diameter standard deviation and overall fibre diameter co-efficient of variation with subjective wool quality scores were generally higher in magnitude than the along or across fibre diameter distribution traits. They were also estimated in a favourable direction with most subjective wool quality scores. Therefore selecting for sheep with less variable fibre diameter will result in correlated improvements in subjective wool quality scores.

INTRODUCTION

The OFDA2000 is utilised to measure fibre diameter and is capable of splitting fibre diameter into along and across fibre components. This provides the opportunity to investigate the genetic relationships between fibre diameter and its components with the suite of visual wool quality scores as well as greasy wool handle. Subjectively recorded traits can be difficult to assess accurately and precisely and determining the estimated correlated response with along and across fibre components may reveal a method of achieving greater genetic gain. Overall fibre diameter co-efficient of variation has been shown to be highly heritable (0.48) (Mortimer *et al.* 2009) with a moderate to strong relationship with greasy wool handle (Roberts 1956; Stevens 1994), wool character (Mortimer *et al.* 2009; Robinson *et al.* 2007), staple weathering (Mortimer *et al.* 1990) and fleece rot (Evans and McGuirk 1983; Watts *et al.* 1981). Estimating the phenotypic and genetic correlations between the visual wool quality wool scores, greasy wool handle and along/across fibre components will further determine if there is any benefit in terms of genetic gain in dividing fibre diameter distribution into its along or across fibre components compared to utilising overall mean fibre diameter coefficient of variation. A previous paper by Preston and Hatcher (2013) provided heritability estimates and variance components for the along and across fibre diameter traits. This paper will report the correlated response of along and across fibre attributes with visual wool quality scores and greasy wool handle.

MATERIALS AND METHODS

Data was obtained from the Information Nucleus Flock (INF) conducted by the Sheep Cooperative Research Centre (Fogarty *et al.* 2007; van der Werf *et al.* 2010). The data used in this study has been previously described by Preston and Hatcher (2013). In addition all progeny were Merino's assessed at yearling age (10-13 months) for a range of subjectively assessed visual wool

quality scores (AWI & MLA 2007). Traits assessed in this study included greasy wool colour (GCOL), wool character (CHAR), dust penetration (DUST), staple weathering (WEATH), fleece rot (FLROT) and staple structure (STRUCT). All traits were scored on a 1-5 scoring system with a higher score representing the less desirable expression. GCOL, STRUCT and CHAR were all assessed on the midside of the sheep; while DUST, WEATH and FLROT were assessed along the top-line of the sheep where the expression of the traits was likely to be most pronounced. Textural greasy wool handle (HAND) was assessed according to Casey and Cousins (2010). This involves the textural components of the wool and requires the assessor to rub their finger along the fibre in a base to tip direction and allocated a 1-5 score with a higher score represented a harsher and more abrasive surface. Midside samples were collected and measured at a commercial fleece measurement laboratory (AWTA Limited Melbourne). For OFDA2000 traits, one staple was chosen at random from the midside sample and then cleaved into the formation of smaller micro staples. The micro staples were then placed on the OFDA2000 fibreglass xy slide and measured for fibre diameter. OFDA2000 output included fibre diameter (FD), overall fibre diameter standard deviation (FDSD), overall fibre diameter co-efficient of variation (FDCV), maximum fibre diameter along the fibre (AMAX), minimum fibre diameter along the fibre (AMIN), across fibre diameter co-efficient of variation (ACCV) and along fibre diameter co-efficient of variation (ALCV).

ASReml 3.0 (Gilmour *et al.* 2009) was used to estimate the genetic parameters using general linear mixed model and residual maximum likelihood methods. As described in Preston and Hatcher (2013), a univariate analysis of all traits including the addition of the following fixed effects: flock (8 sites), year of birth (2007, 2008, 2009, 2010), sex (male, female), dam age (2, 3, 4, 5, 6), birth type rearing type (born single raised single, born twin raised single and born and raised as a multiple) and siregroup (ultra/superfine, fine/fine medium & medium/strong) with appropriate two way interactions. A number of models were fitted to each trait which varied in the combination of random effects (i.e. sire.flock and maternal effect) and a means to account for genetic groups (fitted as random or fixed). Genetic groups were allocated according to the obtained pedigree. Progeny from ancestors with a low number of offspring were removed and then merged into groups with insufficient data. The genetic grouping accounted for the differing ewe foundation flocks at each of the Information Nucleus flocks. The most appropriate model for each trait was determined by log likelihood ratio tests. Phenotypic and genetic correlations for each trait were estimated from the appropriate co-variances in ASReml.

RESULTS AND DISCUSSION

The phenotypic correlations between the along and across fibre traits and subjective wool quality traits were all low to negligible (<0.4) (Table 1). This indicates that within flock there is little to no relationships between these traits. ACCV and ALCV had a lower phenotypic relationship with subjective wool traits than FDCV or FDSD.

Genetic correlations were all stronger in magnitude than the corresponding phenotypic correlations and had higher standard errors which may influence the estimates (Table 2). STRUCT and HAND both had strong positive genetic correlations with FD, FDSD, AMAX and AMIN (>0.60). Therefore selection for either finer fibre diameter, less variable fibre diameter distribution, lower maximum or minimum diameter along the fibre will result in wool with smaller fibre bundles that are texturally softer. STRUCT, HAND and CHAR also had low to moderate positive genetic correlations with ALCV (0.47±0.11, 0.38±0.13 and 0.25±0.13 respectively). Genetic correlations with ACCV were much lower in magnitude (≤0.05) inferring that STRUCT, HAND and CHAR are more associated with along fibre diameter components rather than across fibre attributes. FLROT had a low positive genetic correlation with FDCV (0.30±0.08) and ACCV (0.39±0.09), which supports previous reports that fleecerot is linked to fibre diameter distribution

(Evans and McGuirk 1983; Watts *et al.* 1981). All other genetic correlations with FLROT were negligible (≤0.13) including ALCV (0.00±0.13); inferring that greater response to selection will be achieved if ACCV was used. In agreement with Hatcher *et al* (2004), DUST had a negligible genetic correlation with ALCV or ACCV (≤0.02). DUST and GCOL both had favourable low to moderate positive genetic correlation with FD, FDSD, AMAX and AMIN ($0.44 \leq r_g \geq 0.25$). Therefore selection for finer fibre diameter, reduced overall fibre diameter distribution and lower minimum and maximum diameter along the fibre will generate correlated improvements in greasy wool colour (i.e. whiter wool) and reduced dust penetration along the wool staple.

Table 1. Phenotypic correlations (r_p) between OFDA traits and subjectively assessed traits

	GCOL	CHAR	DUST	WEATH	FLROT	STRUCT	HAND
FD	0.15±0.02	0.20±0.02	0.05±0.02	0.04±0.02	-0.08±0.02	0.30±0.01	0.28±0.02
FDSD	0.16±0.02	0.26±0.02	0.05±0.02	0.04±0.02	0.03±0.02	0.29±0.01	0.25±0.02
FDCV	0.08±0.02	0.17±0.02	0.02±0.02	0.02±0.02	0.10±0.02	0.12±0.02	0.09±0.02
AMAX	0.14±0.02	0.19±0.02	0.05±0.02	0.05±0.02	-0.07±0.02	0.29±0.01	0.27±0.02
AMIN	0.16±0.02	0.20±0.02	0.05±0.02	0.03±0.02	-0.06±0.02	0.30±0.01	0.28±0.02
ACCV	0.09±0.02	-0.01±0.02	-0.04±0.02	0.02±0.02	0.13±0.02	-0.03±0.02	-0.04±0.02
ALCV	0.00±0.02	0.07±0.02	0.01±0.01	0.01±0.02	0.00±0.02	0.05±0.02	0.06±0.02

Table 2. Genetic correlations (r_g) between OFDA traits and subjectively assessed traits

	GCOL	CHAR	DUST	WEATH	FLROT	STRUCT	HAND
FD	0.33±0.07	0.51±0.07	0.25±0.10	0.11±0.09	-0.13±0.08	0.68±0.06	0.63±0.07
FDSD	0.44±0.07	0.68±0.06	0.27±0.10	0.22±0.10	0.11±0.09	0.70±0.06	0.61±0.08
FDCV	0.24±0.08	0.39±0.08	0.08±0.11	0.21±0.10	0.30±0.08	0.25±0.09	0.16±0.10
AMAX	0.37±0.07	0.52±0.07	0.27±0.10	0.19±0.09	-0.11±0.08	0.70±0.07	0.68±0.07
AMIN	0.37±0.07	0.53±0.07	0.28±0.10	0.11±0.09	-0.10±0.08	0.67±0.06	0.65±0.07
ACCV	0.24±0.09	0.03±0.10	0.02±0.12	0.26±0.11	0.39±0.09	-0.01±0.10	-0.05±0.12
ALCV	0.13±0.13	0.25±0.13	0.02±0.15	0.22±0.14	0.00±0.13	0.47±0.11	0.38±0.13

Merino breeding programs with a focus on selecting for finer less variable wool will generate correlated improvements in STRUCT, HAND, CHAR, GCOL and DUST. All results had favourable correlations; which suggests that there would be minimal negative effect on visual wool quality scores when producers select for wool with less variable fibre diameter distribution. These results indicate most visual wool quality scores would have a greater response to selection when utilising a overall fibre diameter distribution trait rather then splitting it into its along and across fibre components. Therefore there would be little value in including the latter in a Merino breeding program.

ACKNOWLEDGEMENTS

We would like to thank all staff involved at the INF sites for the work involved in the collection of the data used in this paper.

REFERENCES

AWI & MLA (2007) 'Visual Sheep Scores Research Edition.' Research Edition. (AWI & MLA: Sydney, New South Wales)

Brien F.D., Ferguson M.B. and Konstantinov K. (2001) *Proc. Assoc. Advmt. Anim. Breed. Genet.* **14**: 277.
Casey A. and Cousins P. (2010) Sheep Scores - Wool Handle – Texture (NSW Department of Primary Industries: Orange Agricultural Institute)
Evans R. and McGuirk B. (1983) *Aust. J. Agric. Res* **34**: 47.
Ferguson M.B., Gloag C.M., Behrendt R. and Brien F.D. (2002) *Wool Tech. Sheep Breed.* **50**: 787.
Fogarty N., Banks R., van der Werf J., Ball A. and Gibson J. (2007) *Proc. Assoc. Advmt. Anim. Breed. Genet* **17**: 29.
Gilmour A., Gogel B., Cullis B. and Thompson R. (2009) ASReml User Guide Release 3.0.
Gloag C., Kearney G. and Behrendt R. (2004) *Aust. Soc. Anim. Prod.* **25:** 249
Greeff J.C. (2002) *Wool Tech. Sheep Breed.* **50**: 11.
Hansford K.A., Marler J.W. and McLachlan I.M. (2002) *Wool Tech. Sheep Breed.* **50**: 812.
Hatcher S., Atkins K. and Thornberry K. (2004) *Aust. Soc. Anim. Prod* **25**: 252.
Mortimer R., Taylor P.J., Jackson N., Lax J. and Maddocks I.G. (1990) *Proceedings of the 20th Sheep & Wool seminar and refresher course'* ***11***
Mortimer S., Robinson D., Atkins K., Brien F., Swan A., Taylor P. and Fogarty N. (2009). *Anim. Prod. Sci* **49**: 32.
Preston J.W.V. and Hatcher S. (2013) *Proc. Assoc. Advmt. Anim. Breed. Genet* **20**
Roberts N. (1956) *Text. Res. J* **26**: 687.
Robinson D., Mortimer S., Swan A. and Purvis I. (2007). *Proc. Assoc. Advmt. Anim. Breed. Genet* **17**: 336.
Stevens D. (1994) *Wool spec 94* H1-H10
van der Werf J., Kinghorn B. and Banks R. (2010) *Anim. Prod. Sci* **50**: 998.
Watts J., Merritt G., Lunney H., Bennett N. and Dennis J. (1981). *Aust Vet Journal* **57**: 372.
Yamin M., Hynd P.I., Ponzoni R.W., Hill J.A., Pitchford W.S. and Hansford K.A. (1999) *Wool Tech. Sheep Breed.* **47**: 151.

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:122-125

GENETIC ARCHITECTURE AND EVOLUTION OF QUANTITATIVE TRAITS

M.E. Goddard[1,2]

[1]Department of Environment & Primary Industries, AgriBio, Bundoora, VIC 3083, Australia
[2]Melbourne School of Land & Environment, University of Melbourne, Parkville, VIC 3010, Australia

SUMMARY

Genome wide association studies in livestock and in humans typically find many SNPs of small effect and intermediate allele frequency associated with each quantitative trait. This is in contrast to theories that predict most variance to be due to rare alleles of large effect. This paper reports a computer simulation of the evolution of a quantitative trait under the effects of mutation, selection and genetic drift. The simulation can approximate the experimental findings but only by assuming that there are >1,000,000 sites at which mutation can affect a typical trait, mutation at these sites is much more likely to cause an allele of small effect than of large effect and selection against the mutant allele increases more than linearly with the size of the mutation's effect on the quantitative trait. Thus the experimental results are consistent with the theory for the control of genetic variation in quantitative traits at least under these assumptions.

INTRODUCTION

Quantitative or complex traits are important in agriculture, medicine and evolution but we have no good understanding of the forces that control genetic variation in these traits. Most theories explaining quantitative genetic variation assume that it is controlled by a balance between mutation, which creates new variants, and selection which eliminates these mutant alleles. Consequently, most versions of this theory predict that the genetic variance will be mainly due to rare mutations of large effect (eg Eyre-Walker 2010).

Until recently we had little knowledge of the genes that cause variation in quantitative traits and so this prediction was difficult to test. However, in the last 6 years assays for thousands of genetic markers or single nucleotide polymorphisms (SNPs) have become available for livestock and humans and this has allowed a new type of experiment known as a genome wide association study (GWAS). In a GWAS, individuals are measured for a trait and genotyped for thousands of SNPs. Then the SNPs are searched for those that are significantly associated with the trait. A SNP that is significantly associated with the trait is assumed to 'tag' a nearby mutation that causes variation in the trait because it is in linkage disequilibrium (LD) with this causal mutation (or quantitative trait locus, QTL). GWASs typically find many SNPs with small effects and intermediate allele frequencies but very few with large effects. This seems to contradict the theory that predicts that most of the variance will be due to rare mutations of large effect. In this paper I consider how the theory and the experimental results can be reconciled by using a computer model which simulates the evolution of a quantitative trait.

The simulation requires inputs concerning the number of mutations that can affect a typical quantitative trait, the size of their effects and the selection to which they are subject. Prior to the era of GWASs, some relevant information about the genetic architecture of quantitative traits was available. The variance added each generation by mutation is in the range 0.001 to 0.01 times the environmental variance (V_e) for most traits studied. This variance could be due to many mutations of small effect or few mutations of large effect or a mixture of both. In mice, many experiments have been reported in which a gene is 'knocked out' or replaced by an inactive form. As many as 1 in 3 of these knock outs affect body size (Reed *et al.* 2008). If this applies to most quantitative traits, it implies that over 5000 genes can affect each trait. Some mutations do have a large effect on a quantitative trait. For instance, mutations in the gene FBN1 in humans can cause Marfan's

syndrome which includes a large increase in height. Over 500 different mutations in FBN1 cause Marfan's syndrome (Kemper *et al.* 2012). FBN1 may be unusual, but if even 200 mutations can cause a large change in height, it is likely that even more mutations can cause a small effect on height. Therefore, it seems reasonable to expect that at least 5000 x 200 = 1,000,000 different mutations might affect a trait such as height.

It is usually assumed that natural selection favours an intermediate value for many quantitative traits. That is, individuals with extreme phenotypes are less fit than individuals with intermediate phenotypes. It is also likely that some mutations, which affect a quantitative trait, are detrimental regardless of their effect on phenotype for a particular quantitative trait. Zhang and Hill (2002) call this model of selection a joint effect model because it combines stabilising selection directly on the trait with selection directly against a deleterious mutation, and it is a joint effect model of selection that I have used in the simulation.

In this paper I compare GWAS results for weight in cattle with a simulation of the evolution of a quantitative trait under the influence of mutation, selection and genetic drift. The aim is to find a theory for the control of genetic variation in quantitative traits that is consistent with experimental GWAS results.

MATERIAL AND METHODS

GWAS. Bolormaa *et al.* (2013) analysed data from the Beef CRC on 6000 animals that were measured for post weaning weight and had genotypes for 700,000 SNPs. The method of analysis was called "BayesR' by Erbe et al (2012) and it fits all the SNPs simultaneously. The effects of the SNPs are assumed to be random variables drawn from a mixture of 4 normal distributions. The 4 distributions are such that SNPs in each of the distributions explain on average either zero, 0.0001, 0.001 or 0.01 of the genetic variance of the trait. The analysis used a Gibbs sampling chain and in each cycle the number of SNPs in each of the 4 distributions was counted. In this way the distribution of SNP effects and the variance they explain was estimated. The distribution of variances explained was compared to the simulation results.

Simulation. The computer simulation assumed a constant population size of 10,000. Each generation gametes are formed by recombination between the paternal and maternal gametes of the parent. Mutations occur in these gametes at a rate of 10^{-8} per site and there are 10^6 sites in the genome where mutation affects the trait. The effect of each mutation is drawn from a gamma distribution with a shape parameter of 0.1 and a scale parameter such that the variance added by mutation each generation is 0.001 times the environmental variance. The effect of the mutation is negative in a random 50% of cases. The parents mate at random and the offspring are subject to selection. The fitness of each offspring is obtained by multiplying together a fitness due to stabilising selection and a fitness which is constant for the mutation. The fitness from stabilising selection is $\exp(-0.5\ y^2/V_s)$ where y is the phenotype in environmental standard deviations and V_s is 200Ve. The constant fitness component of a mutation is 1-s where $s = 0.8\ x^2$ where x is the effect of the mutation on the trait in units of environmental standard deviations. After the simulation reaches an equilibrium state it is run for 1000 generations and the number of mutations segregating, their effects sizes and allele frequencies recorded each generation.

RESULTS

In the simulation, the effect of a mutation is drawn from a gamma distribution with shape parameter of 0.1. This is a distribution with many very small effects but a long tail of larger effects (Fig 1). Although mutations of large effect (eg. > 1 standard deviation) are rare they explain most of the mutation variance (Fig 1). However, mutations of large effect are strongly selected against and so they are kept rare and eventually eliminated by natural selection. Eventually when an equilibrium is reached, as much variance is lost each generation by genetic drift and selection as is

added by mutation. At this point the heritability of the trait was 0.33. Table 1 gives the distribution of mutation effects when an equilibrium state has been reached. Table 1 shows that mutations explaining less than 0.0001 of the total variance are common and mutations explaining more than 0.1 of the variance are rare.

The simulation includes the mutations that cause variation in the trait and these have been counted in Table 1 regardless of their allele frequency. However, a GWAS is based on SNPs that are not the causal variants but are hopefully in LD with them. Most of the SNPs on commercial SNP 'chips' such as used in our cattle GWAS, have a minor allele frequency (MAF) in the range 0.1 to 0.5. Therefore a causal mutation with MAF <0.1 cannot be in complete LD with a SNP with MAF > 0.1 and so the SNPs will underestimate the true effect of the causal mutation. The most optimistic assumption would be for causal mutations with MAF > 0.1 to be in complete LD with one of the SNPs and for causal mutations with MAF = q <0.1 to be in LD with a SNP that explains a fraction q/0.1 of the variance explained by the causal mutation. When this assumption is used to calculate the number of SNPs in each variance class (Table 1) the number of SNPs explaining >0.01 of the variance is much less than the number of causal mutations because many of these mutations are rare and hence incompletely detected by the SNPs. Consequently, the number of SNPs explaining <0.0001 of the variance is more than the number of causal mutations because it includes some causal mutations that explain a greater variance but are incompletely 'tagged' by the SNPs.

The BayesR analysis of weight in cattle provides a distribution of the effects of SNPs on weight. The distribution has been summarised (Table 1) by calculating the number of SNPs that fall into each proportion of variance class. The results are broadly similar to those predicted by the simulation model but the real data has even more SNPs explaining <0.0001 of the variance than predicted by the simulation.

Table 1. Number of segregating sites in the computer simulation and number of SNPs in the Bayes R analysis of cattle weight classified by the proportion of genetic variance that they explain.

Proportion of variance explained	Number of causal sites in simulation	Number of simulated SNPs	Number of SNPs from BayesR
$< 10^{-4}$	1466	1562	3166
10^{-4} to 10^{-3}	190	161	1492
10^{-3} to 10^{-2}	145	91	52
10^{-2} to 10^{-1}	21	8	5
$> 10^{-1}$	0.3	0	0.05

DISCUSSION

The simulation parameters might be regarded as extreme in certain respects. For instance, I assumed 1,000,000 sites in the genome affect a typical trait when mutated, that the distribution of their effects is very leptokurtotic (ie has a long tail) and that selection against mutations rises with the effect of the mutation squared. These assumptions all act to increase the importance of SNPs with small effects and decrease the number of SNPs with large effects when an equilibrium is reached. Despite this, the simulated data has fewer SNPs of small effect than the real data on cattle weight. To mimic the real GWAS results more closely the simulation would need to assume that >1,000,000 sites in the genome affect a typical trait when mutated. Thus the true parameters may be even more extreme than those assumed in the simulation.

If the theory of Eyre-Walker (2010) is applied to the parameters assumed in the simulation it also predicts that most of the variance is due to mutations of small effect. However, most authors have ignored this conclusion perhaps because they regarded the input parameters to be too extreme to be realistic. The conclusion of this paper is that they are not extreme enough.

Qualitatively the simulation matches an important feature of real data on quantitative traits. Mutations of large effect (> 2 standard deviations) occur for many traits. For instance, mutations causing dwarfism are known in many species but they are usually kept rare by natural selection so that they explain little of the total genetic variance.

The simulation results make a prediction with important practical consequences. The simulation predicts that there are a number of QTL segregating that explain > 1% of the variance but which go undetected by GWAS because their MAF is too low. This could explain the 'missing heritability' discussed in human genetics (Yang *et al.* 2010) and in cattle (Haile-Mariam *et al.* 2013). If this is indeed the case, the use of genome sequence data instead of SNP genotypes or haplotypes of SNPs should lead to the discovery of more QTL of medium size effects and their exploitation in genomic selection.

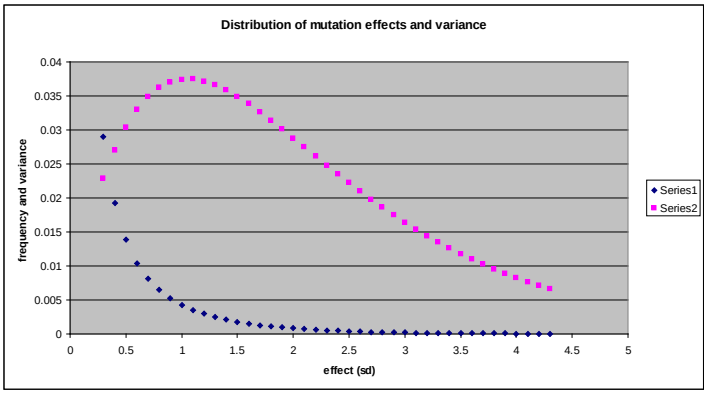


Figure 1. Distribution of the effects of mutations (series 1) and the mutation variance explained (series 2) by mutations of different size measured in units of environmental standard deviations.

REFERENCES

Bolormaa S., Pryce J.E., Kemper K., Savin K., Hayes B.J., Barendse W., Zhang Y., Reich C.M., Mason B.A., Bunch R.J., Harrison B.E., Reverter A., Herd R.M., Tier B., Graser H-U and Goddard M.E. (in press) *J. Anim. Sci.*

Erbe M., Hayes B.J., Matukumalli L.K., Goswami S., Bowman P.J., Reich, C.M., Mason B.A., and Goddard M.E. (2012) *J.Dairy Sci.* **95**: 4114.

Eyre-Walker A (2010) *Proc, Nat. Acad. Sci*. **107**: 1752.

Haile-Mariam M., Nieuwhof G.J., Beard K.T., Konstatinov K.V. and Hayes B.J. (2013) *J. Anim Breed Genet.* doi:10.1111/j.1439-0388.2012.01001.x

Kemper K.E., Visscher P.M. and Goddard M.E. (2012). *Genome Biology*. **13**: 244.

Reed D, Lawler M, Tordoff M (2008) *BMC Genetics* **9**: 4.

Yang J., Benyamin B., McEvoy B.P., Gordon S., Henders A. Nyholt DR, Madden PA, Heath AC, Martin NG, Montgomery GW, *et al* (2010) *Nature Genet.* **42**: 565.

Zhang X-S. and Hill W.G. (2002). *Genetics* **167**: 459.

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:129-129

COMBINING MULTIPLE TEST-STATISTICS INCREASES THE POWER OF SELECTIVE SWEEP ANALYSES IN CATTLE

I.A.S. Randhawa[1], P.C. Thomson[1], M.S. Khatkar[1] and H.W. Raadsma[1]

[1]ReproGen - Animal Bioscience Group, Faculty of Veterinary Science, University of Sydney, Camden, NSW 2570, Australia

SUMMARY

Technological advances in targeted DNA sequencing, SNP genotyping and biometrical tools, allow for accurate localization of selection signatures. We present a simple method of combining ranks (mean fractional ranks, MFR) of multiple test-statistics as evidence of selection from single (F_{ST}, ΔDAF) and multiple (XP-EHH) marker based tests. *P*-values and FDR (*q*-values) to assess significance of an association can be determined from MFR: this cannot be done for its constituent tests. MFR is validated in two datasets (grouped for the presence or absence of either polledness or double muscling) from 375 animals of 21 cattle breeds with genotypes on 38,610 SNP assays from an Illumina BovineSNP50 chip. Candidate regions under selection (CRS) on chromosomes 1 and 2 were localized to regions of 610 and 680 kb near the functional mutations causing polledness and double muscling in cattle, respectively. The existence of strong signals of low FDR (i.e., > 85% of SNPs in CRS have $q < 0.05$) close to the candidate genes confirms the robustness of MFR.

INTRODUCTION

Trait-specific signals of selection are very challenging to identify. Multiple methods have been developed for the detection of selection signatures from genome-wide single nucleotide polymorphism (SNP) data. These have been extensively implemented in population studies for many species. The specificity of each selection test is limited to certain aspects of selective forces operating under various models of selection. Hence, many tests being used to link genotypes with phenotypes often provide differing results for the same genomic data (Lin *et al.* 2010).

Non-neutral patterns of local genomic variation may reflect historical selective sweeps resulting in signatures of selection. A population undergoing positive selection for specific traits can exhibit signals of selection at the underlying genomic regions when measured by various selection tests of allele frequency spectrum and haplotype structures (Qanbari *et al.* 2011). Therefore, a combination of multiple strategies would appear to be a robust approach in localizing candidate regions under selection (CRS) and correlate them with phenotypic variation. Recently, several approaches have been developed (Grossman *et al.* 2010; Lin *et al.* 2010; Pavlidis *et al.* 2010) which combine multiple summary statistics in order to improve the power of detecting selection signatures. However, the complexity of methods, extensive range of computational resources and prior knowledge required to implement available combining approaches leaves researchers with limited resources at a disadvantage. To improve trait-specific genome-wide selection scans, we present a simple method of combining evidence from the ranks of several selection tests requiring no prior information and it is potentially ideal for outbred populations.

MATERIALS AND METHODS

This study was conducted on two well characterized traits under selection in cattle to validate the MFR method. We investigated multi-breed panels from 212 (dataset I: polled versus horned breeds) and 357 (dataset II: double muscle versus normal muscle breeds) cattle samples genotyped with the Illumina BovineSNP50 chip assays, available from Gautier *et al.* (2010). We used 38,600 SNPs that were mapped on the UMD3.1 bovine assembly. Imputation of missing genotypes and haplotyping were performed with BEAGLE 3.3 (Browning and Browning 2007). Ancestral and

derived allelic polarity was acquired from Decker *et al.* (2009) and Matukumalli *et al.* (2009).

Mean Fractional Ranks (MFR). We combined three popular constituent tests to capture evidence for selection across multiple populations from genetic polymorphism data namely change in allele diversity by F_{ST} (Weir and Cockerham 1984), across population extended haplotype homozygosity (XP-EHH) test (Sabeti *et al.* 2007) and change in derived allele frequencies (ΔDAF) (Grossman *et al.* 2010). We derived composite test statistics (i.e., MFR) by combining 3 tests statistics at the same SNP, as well as determine *P*-values for these composite tests, to test the presence of a common signal as follows:

Let T_{ij} be the test statistic using method $i, (i = 1, \dots, m)$ calculated at SNP $j, (j = 1, \dots, n)$. Then for each test statistic type i obtain the rank of each observed test statistic across all n SNPs, say $R_{ij} = \text{rank}\,(T_{ij})$, which take values $1, \dots, n$ (using R program's `rank` function with default options so that it averages the sequential ranks for equal scores on multiple SNPs of a test). Next, these ranks are converted to fractional ranks by re-scaling them to lie between 0 and 1, i.e. $R'_{ij} = R_{ij}/(n+1)$, giving values from $1/(n+1)$ through $n/(n+1)$. Next, the MFR of the test statistics at each SNP is calculated, averaging over all the test statistic methods, $\bar{R}'_j, j = 1, \dots, n$. If there is a common signal across the multiple test statistics, this will show up as an excess in the $\bar{R}'_j$ value at that point, otherwise, $\bar{R}'_j$ may be dampened down, i.e. regressed to the average. Under the null hypothesis of no common signal, we can regard the values of R'_{ij} as m independent observations from a uniform $U(0,1)$ distribution, and using the results of Sadooghi-Alvandi *et al.* (2009) for the sum of m $U(0,1)$ random variables, we can derive the distribution of the mean $\bar{R}'_j$ as follows.

The probability density function (PDF) of $\bar{R}'_j$ is obtained as

$$f(r) = \frac{1}{(n-1)!}\sum_{k=0}^{n}(-1)^k \binom{n}{k}[(rn-k)_+]^{n-1}, 0 \le r \le 1$$

where $x_+ = x$ if $x > 0$, or 0 otherwise. By integration, the cumulative distribution function (CDF) is obtained as

$$F(r) = \frac{1}{n!}\sum_{k=0}^{n}(-1)^k \binom{n}{k}[(rn-k)_+]^{n}, 0 \le r \le 1$$

So for a mean scaled rank of $\bar{R}'_j$, the p-value for a test of no common signal would be calculated as $p = 1 - F(\bar{R}'_j)$.

The top 0.1% of $-\log_{10}$ of the empirical p-values were used to declare a SNP to be significant relative to the rest of the genome. The effectiveness of multiple tests was also compared gradually at various thresholds. Further, empirical *p*-values were calibrated using the ConReg-R method (Li *et al.* 2011) and the tail area based false discovery rate (FDR) i.e., q-values were estimated.

RESULTS AND DISCUSSION

Genome-wide distribution of empirical scores (non-smoothed) indicates that the highest $-\log_{10}(p)$ of MFR values above various thresholds were in the candidate regions in both datasets (Figure 1). The three component tests (FST, ΔDAF and XP-EHH) were found significant in the candidate gene regions but with fewer and lower ranked SNPs as compared to the MFR test (results not shown). To reduce spurious signals, the test statistics were smoothed by averaging statistics over SNPs within 1 Mb sliding windows centered at each SNP (Figure 2). Putative regions under selection (PRS) were defined from windows containing at least 3 significant SNPs and first to last SNP (top 0.1 %) positions as its boundaries. In total, 9 and 12 PRSs detected by at

least one of the constituent selection tests were substantially reduced at common signals to 3 and 4 PRSs by the MFR method in datasets I and II, respectively (Table 1). Genes located within the PRS ± 0.5 Mb positions were investigated for previously reported candidates of selection to localize CRS. MFR shows clusters of significant SNPs as peaks of selection signatures in CRSs on bovine autosome (BTA) 1 and 2 (Figure 2). The presence of non-candidate selection signals was much lower in MFR as compared to constituent tests (results not shown). The strategy of combining multiple test statistics has neutralized the unique patterns of each constituent selection test. In the empirical MFR distribution, the significant scores have an FDR < 0.0001, and after smoothing > 85% of SNPs in CRSs have $q < 0.05$. Additional peaks at PRSs by MFR also indicate the presence of genes under selection, for example; in the dataset I, a strong phenotypic diversity also exists for stature on BTA13 and 14, see Randhawa *et al.* (2013).

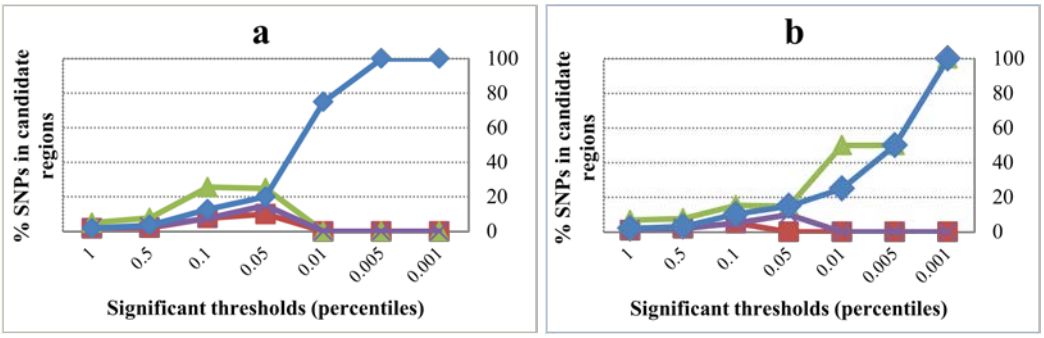


Figure 1: Percentage of significant SNPs present within the candidate gene regions (*y*-axis) identified by MFR (◆), XP-EHH (▲), F_{ST} (■) and ΔDAF (×) in a) polled and b) double muscle cattle at various thresholds (*x*-axis).

Polledness in cattle. In dataset I, out of 39 SNPs above the top 0.1% MFR scores, 10 SNPs within 610 kb span were found in the CRS harbouring *POLL* locus on BTA1 (Figure 2a). The *POLL* locus contains candidate mutations at the proximal end of BTA1 (1.65–2.05 Mb) where dominant alleles cause the polledness in cattle (Matukumalli *et al.* 2009; Allais-Bonnet *et al.* 2013).

Double muscling in cattle. In dataset II, among 39 significant MFR scores, a cluster of 10 SNPs was localized at CRS of 680 kb flanking myostatin (MSTN) gene at BTA2 (Figure 2b). Bovine MSTN gene (6.21–6.22 Mb) harbours various loss-of-function mutations or an 11 bp deletion in its third exon that underlie the muscular hypertrophy in some beef cattle (Piedmontese, Belgian Blue, South Devon and Asturiana de los Valles) breeds (Georges 2010).

Table 1: Regions under selection (putative = PRS, candidate = CRS) and significant SNPs in constituent and composite tests, and FDR of MFR in both datasets of cattle

Dataset	Total number of			Number of PRS and (SNPs* in CRS) in				% FDR¶ of MFR in Genome and (CRS)
	PRS	SNPs* in PRS	SNPs† in CRS	Constituent tests			Composite MFR	
				XPEHH	F_{ST}	ΔDAF		
I	9	105	14	3 (9)	5 (1)	5 (0)	3 (10)‡	9.8 (86.0)
II	12	129	10	5 (10)‡	4 (3)	5 (0)	4 (10)‡	6.2 (90.0)

* Significant SNPs † Total genomic SNPs ‡ Extreme scoring SNPs ¶ $q<0.05$

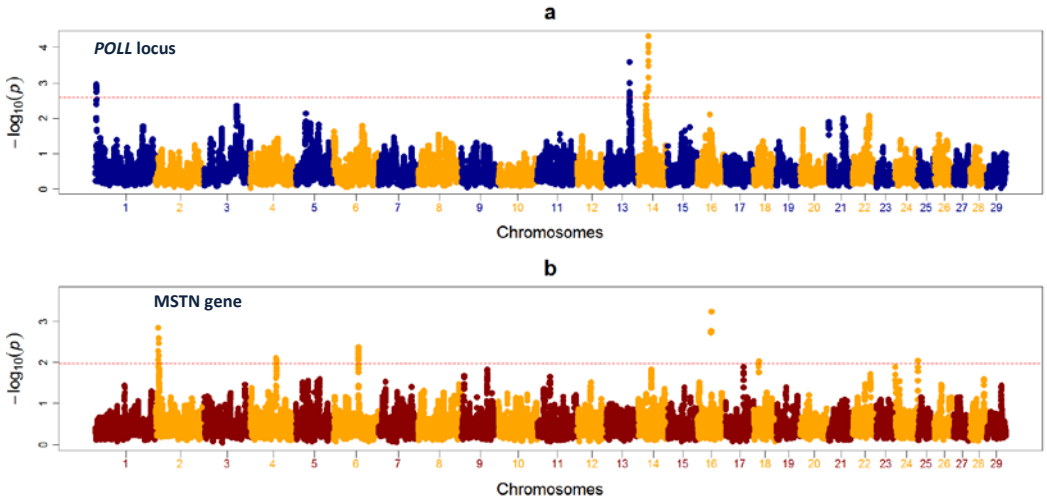


Figure 2: Manhattan plots of smoothed $-log_{10}(p)$ of MFR for a) polled and b) double muscle cattle. Dashed lines indicate genome-wide top 0.1% thresholds in both datasets.

Overall, MFR demonstrates its robustness even in the absence of any casual SNP in the genotype data. It provides an improvement for the predictions of positive selection as compared to its constituent tests of selection. MFR can be further improved by incorporating a strategy so that it can use the magnitudes of the actual test statistics. Moreover, MFR can easily accommodate additional selection tests given their sufficient power to distinguish selected and neutral loci in the genetic polymorphism data. This method can be used to identify the CRSs harbouring functional SNPs in genes for simple and potentially also for complex traits in domestic species.

REFERENCES

Allais-Bonnet A., Grohs C., Medugorac I., Krebs S., Djari A. *et al.* (2013) *PLoS ONE.* **8**:e63512.
Browning S.R. and Browning B.L. (2007) *Am. J. Hum. Genet.* **81**:1084.
Decker J.E., Pires J.C., Conant G.C., McKay S.D. *et al.* (2009) *Proc. Natl. Acad. Sci.* **106**:18644.
Gautier M., Laloë D. and Moazami-Goudarzi K. (2010) *PLoS ONE.* **5**:e13038.
Georges M. (2010) *Immun., Endoc. & Metab. Agents in Med. Chem.* **10**:240.
Grossman S.R., Shylakhter I., Karlsson E.K., Byrne E.H. *et al.* (2010) *Science.* **327**:883.
Li J., Paramita P., Choi K.P. and Karuturi R.K.M. (2011) *Biol. Direct.* **6**:27.
Lin K., Li H., Schlotterer C. and Futschik A. (2010) *Genetics.* **187**:229.
Matukumalli L.K., Lawley C.T., Schnabel R.D., Taylor J.F. *et al.* (2009) *PLoS ONE.* **4**:e5350.
Pavlidis P., Jensen J.D. and Stephan W. (2010) *Genetics.* **185**:907.
Qanbari S., Gianola D., Hayes B., Schenkel F., Miller S., Moore S., Thaller G. and Simianer H. (2011) *BMC Genomics.* **12**:318.
Randhawa I.A.S., Khatkar M.S., Thomson P.C. and Raadsma H.W. (2013) *Proc. Assoc. Advmt. Anim. Breed. Genet.* **20**:(in press).
Sabeti P.C., Varilly P., Fry B., Lohmueller J., Hostetter E. *et al.* (2007) *Nature.* **449**:913.
Sadooghi-Alvandi S., Nematollahi A. and Habibi R. (2009) *Statistical Papers.* **50**:171.
Weir B.S. and Cockerham C.C. (1984) *Evolution.* **38**:1358.

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:130-133

BOVINE FAT DEPOTS DISCRIMINATE BY GENE EXPRESSION PATTERNS

N.J. Hudson[1], A. Reverter[1], D.W. Pethick[2], J.P. Siddell[3], P.L Greenwood[4] and Brian P. Dalrymple[1]

[1]CSIRO Animal, Food and Health Sciences, 306 Carmody Road, St. Lucia, Brisbane, QLD4067
[2] Murdoch University, Perth, WA 6150
[3] NSW Department of Primary Industries, Glen Innes Agricultural and Research Advisory Station, Glen Innes NSW2370
[4]NSW Department of Primary Industries Beef Industry Centre, University of New England, Armidale, NSW2351

SUMMARY

To improve our molecular understanding of bovine fat metabolism, global patterns of gene expression were explored in 5 fat depots: subcutaneous rump (SC), intermuscular (Inter), intramuscular (IMF), omental (Omen) and kidney (Kid). All depots share conserved co-expression gene sets relating to fundamental adipocyte cytoskeletal architecture, metabolism and inflammation. However, the various depots could clearly be discriminated from each other by gene expression. Differences in fatty acid saturation between SC and the other depots are reflected by differential expression of the *SCD* gene that encodes the Δ 9 desaturase enzyme. The fundamental lipogenic machinery such as the *ACACA* gene encoding the rate limiting synthetic enzyme acetyl coA carboxylase is expressed at lower levels in IMF. We also detected differences in expression consistent with divergent lipogenic fuel preferences. Across depots, the most differentially expressed (DE) genes align with those published in the literature for non-ruminants, illustrated by SC rump's highly divergent expression of *HOXA10* and *DLK1*. These genes are likely markers for populations of pre-adipocytes whose properties vary between depots.

INTRODUCTION

Deposition of marbling fat has a positive impact on sensory meat quality through enhanced flavour, juiciness and tenderness. Development of the non-edible fat depots, particularly subcutaneous fat (SC), is considered energetically and commercially wasteful. Therefore, a better understanding of fat depot biology contributes to the challenge of efficiently maintaining product quality in a resource-constrained world. Genetics and nutrition can alter percent intra muscular fat (IMF%) and fat depot distribution. However, IMF development remains an enigmatic trait. In cattle, there are few, if any, known causal mutations, phenotypic variation in IMF% explained by single nucleotide polymorphisms (SNP) is modest (Barendse *et al.* 2010) and the key precursor cell populations have not been unequivocally identified (Harper and Pethick 2004). Physiological differences between depots have been postulated. For example, IMF adipocytes are thought to have a lipogenic preference for glucose and lactate carbon while SC adipocytes prefer acetate (Smith *et al.* 2009). The expression research described here underpins gene and pathway discovery in *bovine* fat metabolism.

MATERIALS AND METHODS

In brief, 15 individual 250 day grain fed Angus, Hereford and Wagyu × Angus steers (n = 5 per breed) were slaughtered at ~26 months of age as part of a larger experiment detailed by (Greenwood *et al.* 2011). Fat depot samples were dissected from each carcass as soon as possible after slaughter from the *m. longissimus dorsi* (IMF), Inter, Omen, Kid and SC depots and snap frozen in liquid nitrogen. The *longissimus dorsi* muscle with IMF intact (LD) was also sampled. Total RNA was phenol chloroform extracted using Trizol (Invitrogen) following the

manufacturer's instructions. RNA yield and purity were determined using microphotometry and RNA integrity by agarose gel electrophoresis. RNA was submitted to the Ramaciotti Institute (Randwick, NSW, Australia) for hybridisation to the 4×44K one colour Agilent bovine array. Data was normalised using a previously described mixed-model approach (Reverter *et al.* 2005) and expressed as log2 values. The expression measurements represent mRNA abundance on a per unit total RNA basis.

Data-driven clustering. The expression profiles of 10,000 genes chosen at random were imported into Permut Matrix software (Caraux and Pinloche 2005). Global relationships between depots based on the molecular data were determined using unsupervised hierarchical clustering performed on columns for all tissues. LD muscle was included in this analysis for comparison.

Co-expression network. To gain insight into molecular relationships within and between fat depots, we used a co-expression approach to build the first exclusively fat depot-based *bovine* network. LD muscle samples were not included here. We filtered the normalised data to leave a manageable subset of genes that satisfied at least one of the following criteria: top 10% in terms of variability of expression; top 10% most abundant expression; annotated as either a transcription factor, cofactor, or chromatin remodeller by (Zhang *et al.* 2012). Significance of differences between genes across depots was determined by PCIT (Reverter and Chan 2008) followed by a hard threshold of 0.975.

RESULTS AND DISCUSSION

Data-driven clustering. Permut matrix software produced a dendrogram of relationships, interpretable from the top down. The text labels include breed and tissue. The first major split shows the LD was discriminated from the fat samples. The next split shows each fat depot could be clearly resolved (Figure 1A). IMF was awarded a unique branch within the fat tree. Inter and Omen were most closely related, followed by Kid then SC.

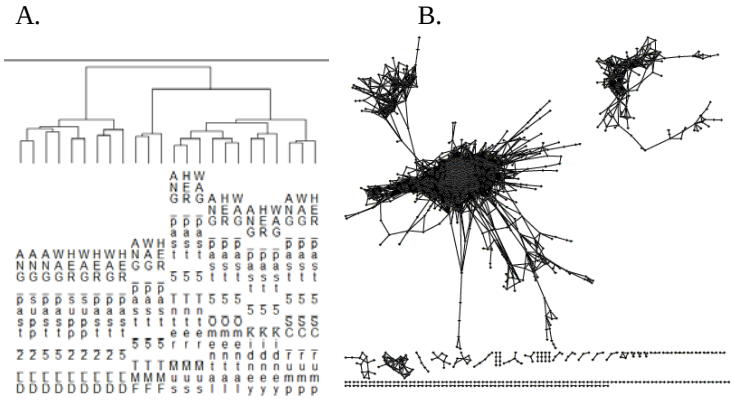


Figure 1. A. Hierarchical clustering of gene expression separated depots. B. Bovine fat depot co-expression network resolves into 3 major clusters reflecting cellular diversity.

Co-expression network. The co-expression network (Figure 1B) was visualised in Cytoscape (Shannon *et al.* 2003). We clustered using a Cytoscape algorithm called 'organic.' This shortens the path length between highly inter-connected genes producing visually coherent representations of gene interactions. Overall, it resolved into three major sub networks (Figure 1B) functionally enriched for 1) cytoskeletal architecture and metabolism (largest cluster; hypergeometric *P*-value = 7.57e-17) 2) inflammation (top left; *P* = 3.63e-18) and 3) peptidyl serine phosphorylation (top

right; *P* = 1.34e-4). These major clusters reflect the diversity of cell types which include macrophages and other immune cells in addition to adipocytes, pre-adipocytes and endothelial cells (Lee *et al.* 2013). Given the emerging link between inflammation and adiposity (Smorlesi *et al.* 2012), the inflammatory cluster is noteworthy. Molecules present in these modules display highly coordinated changes in expression across depots. The network contained several representatives from the three major gene sets previously identified (De Jager *et al.* 2013) representing triacylglyceride (TAG) synthesis (e.g. fatty acid binding protein 4, *FABP4*), fatty acid synthesis (e.g. fatty acid elongase 6, *ELOVL6*) and *PPARG* (e.g. acetyl coenzyme A synthetase, *ACSS2*).

In comparing SC (the most divergent depot) against the other depots we detected extreme DE of delta-like 1 homolog (*DLK1*) and homebox A10 (*HOXA10*) among others (Figure 2). In humans, it has previously been noted that genes regulating early development, including members of this family of phylogenetically ancient homeotic (*HOX*) genes, differ among undifferentiated pre-adipocytes between depots (Tchkonia *et al.* 2007). Similarly, *DLK1* has also been described as a marker for adipocyte progenitors (Shan *et al.* 2013). Gene expression clearly detects the presence of RNA diagnostic of skeletal muscle in our IMF sample. It is unclear to what extent the muscle RNA complicates the marbling adipocyte interpretation.

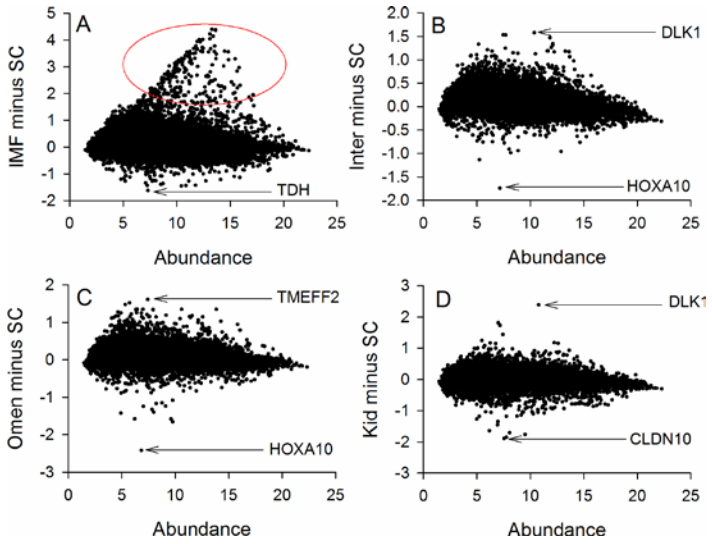


Figure 2. SC rump versus other fat depots. Oval highlights muscle derived transcripts

A targeted examination of enzymatic expression profiles across depots informed by known differences in tissue phenotypes relating to saturation (*SCD*), elongation (*ELOVL6*), TAG synthesis (*DGAT2*), synthetic capacity (*ACACA*), and acetate (*ACSS2*) and glucose (*MDH2*) fuel usage highlighted the following possible molecular drivers (Table 1). IMF displays lower expression of key lipogenic enzymes in line with, but to a lesser extent than, previous biochemical measurements made in cattle and pigs (Bonnet *et al.* 2007, Gardan *et al.* 2006).

Table 1. Log2 expression of genes encoding rate limiting enzymes of fatty acid composition.

Gene	Enzyme (EC#)	Probe	IMF	Inter	Kid	Omen	SC
SCD	Δ 9 desaturase (1.14.19.1)	A_73_P101286	10.49	11.15	11.03	11.01	11.76
ELOVL6	fatty acid elongase 6 (2.3.1.n8)	A_73_119372	10.88	11.44	11.43	11.42	11.54
ACACA	acetyl coA carboxylase (6.4.1.2)	A_73_P038926	6.90	7.45	7.41	7.27	7.74
DGAT2	diacylglycerol O-acyltransferase (2.3.1.20)	A_73_118582	15.86	16.40	16.42	16.34	16.72
ACSS2	acetyl coenzyme A synthetase (6.2.1.1)	A_73_P037091	13.10	13.77	13.71	13.62	14.01
MDH2	malate dehydrogenase 2 (1.1.1.37)	A_73_P422416	18.42	18.03	18.09	18.04	18.11

CONCLUSIONS

Biological similarity between fat depots is reflected by shared clusters of some highly co-expressed genes. Having said this, the 5 bovine depots can be clearly separated by global gene expression patterns, in a manner similar to other species. These depot-specific differences reflect, in part, the proportion and behaviour of populations of pre-adipocytes coupled with metabolic differences such as saturation and lipogenic fuel preference.

REFERENCES

Barendse W., Bunch R. J. and Harrison B. E. (2010) *J Anim Sci*. **88**: 47.

Bonnet M., Faulconnier Y., Leroux C., Jurie C., Cassar-Malek I., Bauchart D., Boulesteix P., Pethick D., Hocquette J. F. and Chilliard Y. (2007) *J Anim Sci*. **85**: 2882.

Caraux G. and Pinloche S. (2005) *Bioinformatics*. **21**: 1280.

De Jager N., Hudson N. J., Reverter A., Barnard R., Cafe L. M., Greenwood P. L. and Dalrymple B. P. (2013) *J Anim Sci*.

Gardan D., Gondret F. and Louveau I. (2006) *Am J Physiol Endocrinol Metab*. **291**: E372-80.

Greenwood P. L., Siddell J., McPhee M., Walmsley B., Geesink G. and Pethick D. W. (2011) *Adv Anim Biosci*. **2**: 408.

Harper G. and Pethick D. (2004) *Australian Journal of Experimental Agriculture*. **44**: 653.

Lee M. J., Wu Y. and Fried S. K. (2013) *Mol Aspects Med*. **34**: 1.

Reverter A., Barris W., McWilliam S., Byrne K. A., Wang Y. H., Tan S. H., Hudson N. and Dalrymple B. P. (2005) *Bioinformatics*. **21**: 1112.

Reverter A. and Chan E. K. (2008) *Bioinformatics*. **24**: 2491.

Shan T., Liu W. and Kuang S. (2013) *FASEB J*. **27**: 277.

Shannon P., Markiel A., Ozier O., Baliga N. S., Wang J. T., Ramage D., Amin N., Schwikowski B. and Ideker T. (2003) *Genome Res*. **13**: 2498.

Smith S. B., Kawachi H., Choi C. B., Choi C. W., Wu G. and Sawyer J. E. (2009) *J Anim Sci*. **87**: E72-82.

Smorlesi A., Frontini A., Giordano A. and Cinti S. (2012) *Obes Rev*. **13 Suppl 2**: 83.

Tchkonia T., Lenburg M., Thomou T., Giorgadze N., Frampton G., Pirtskhalava T., Cartwright A., Cartwright M., Flanagan J., Karagiannides I., Gerry N., Forse R. A., Tchoukalova Y., Jensen M. D., Pothoulakis C. and Kirkland J. L. (2007) *Am J Physiol Endocrinol Metab*. **292**: E298-307.

Zhang H. M., Chen H., Liu W., Liu H., Gong J., Wang H. and Guo A. Y. (2012) *Nucleic Acids Res*. **40**: D144.

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:134-137

USING TWO DIFFERENT APPROACHES TO INFER THE GENETIC STRUCTURE OF POPULATIONS WITH COMPLEX RELATIONSHIPS: THE CASE OF THE AVILEÑA-NEGRA IBÉRICA

D. Martin-Collado[1,2], K.J. Abraham[3], S.T. Rodriguez-Ramilo[1], M.A. Toro[4], M.J. Carabaño[1] and C. Diaz[1]

[1]Dpto. Mejora Genética Animal. INIA. Ctra. de la Coruña km.7.5 28040, Madrid, Spain
[2]AbacusBio Limited, PO Box 5585, Dunedin, New Zealand
[3]Dpto. de Biologia Celular e Molecular, Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo Ribeirão Preto SP, Brazil
[4]Dpto. Producción Animal, E.T.S.I. Agrónomos, Universidad Politécnica de Madrid. Ciudad Universitaria, 28040, Madrid, Spain

SUMMARY

The inference of the genetic structure of domestic animal populations has important implications in the design of breeding programs. In this paper, we assessed the utility of a graphical clustering algorithm (GCA) to identify the genetic structures of real livestock populations with complex relationships comparing it to a Bayesian clustering algorithm (STRUCTURE). The genetic structure of the Spanish cattle breed Avileña-Negra Ibérica was inferred by the analysis of 13,343 animals from 70 herds genotyped for 17 microsatellites. We compared the results of GCA and STRUCTURE regarding the ability to restore Hardy-Weinberg equilibrium in each subpopulation and the average coancestry within and between subpopulations. Both approaches described a similar structure for the ANI breed, which was found to have three genetic subpopulations and a pool of individuals that cannot be assigned without ambiguity to any of the subpopulations. This structure is coherent with the history of the breed. The GCA showed to be a much faster method to infer genetic structure with high ability to determine the core hidden structure of populations with complex relationships.

INTRODUCTION

The demographic and correlated genetic structure of livestock populations has important implications for the design of breeding and conservation programs. In addition, there is a renovated interest in studying the genetic structure of livestock populations since population stratification could bias the prediction of genomic breeding values as well as the results of GWAS (Janss *et al.* 2013). Genetic structures can be analysed using molecular information and several methodologies have been developed for this aim. The bayesian methods and in particular the STRUCTURE software (Pritchard *et al.* 2000) has become very popular. However, STRUCTURE might show difficulties in assesing the genetic structure of populations with a complex pedigree structure. In this paper, we introduce the use of graphical clustering algorithms (GCA) for the inference of the genetic structure of livestock populations. We used a new GCA (Abraham *et al.* forthcoming), that make use of the population molecular coancestry matrix to determine its genetic structure. The Avileña-Negra Ibérica (ANI) breed is an example of a population with complex relationships, where herds are subpopulations with different degrees and patterns of connection among them. Vasallo and Díaz (1986) determined that ANI breed had a pyramidal structure with the majority of herds recurrently buying bulls from the same leading herds. Therefore, the genetic status of the breed was highly dependent on the genetic management of those few leading herds. Since then the ANI population might have evolved. The aim of this paper is twofold: first, to assess the usefulness of GCA to identify the genetic structure of real livestock populations by comparing its performance to Bayesian clustering algorithms (STRUCTURE); and second, to

study the current genetic structure of the ANI beef cattle breed using microsatellites (MS) genotypes to assess the genetic relationships among individuals.

MATERIAL AND METHODS

Material. The ANI breed is a Spanish beef cattle breed reared under extensive conditions. We analysed a data set of 13,343 individuals from 70 herds genotyped for 17 MS.

Methods. We compared the performance of a GCA with the model based algorithm implemented by STRUCTURE.

STRUCTURE algorithm. The Bayesian clustering algorithm implemented in STRUCTURE can assign either the individuals or a fraction of their genome (a proportion of inferred ancestry) to a number of clusters (K) based on multilocus genotypes (Pritchard *et al.* 2000). To determine K we used Evanno's *et al*. (2005) criterion that was found to perform better than the one initially proposed by Pritchard *et al.* (2000) to detect the more likely number of subpopulation (K) when the pattern of dispersal of individuals was not homogeneous.

Graphical clustering algorithm. The GCA works on a symmetric matrix whose off-diagonal elements are the values of the correlation between the corresponding elements to be clustered. In our case, the matrix contained the molecular coancestry values among the 13,343 ANI individuals analysed, which were obtained from the information on the frequencies of the markers following Caballero and Toro (2002). The matrix was calculated by Metapop software (Pérez-Figueroa *et al.*, 2009). The GCA used comprises two algorithms that are run one after another. The first one identifies all possible independent (or less related) individuals using a modification of a method shown in Abraham and Fernando (2012) and the second builds the clusters around these independent animals, as described in detail in Williams *et al*. (2011). Two thresholds (molecular coancestry values) have to be set to determine which individuals are considered as independent and which are defined as closely related. The thresholds were established according to the percentiles of the distribution of the molecular coancestry values. The percentile chosen depends on the expected genetic differentiation of the subpopulations. We did not expected ANI population to have a simple genetic structure; therefore, the threshold that defined independent animals was set to be very conservative. Several thresholds were used to define independent animal however the number of independent animals was similar. The case we present correspond with the percentile 1.25 of the molecular coancestry matrix. The second threshold (closely related individuals) corresponded with the percentile 75.

Genetic contribution of herds. We analysed 209,694 animals of 732 herds included in ANI breed Herdbook to complement the result of the genetic analysis. We determined the contribution of the different herds to the genetic composition of the ANI breed with ENDOG software (Gutiérrez and Goyache, 2005) by calculating the probability of gene origin of the ancestors and then summing the contribution values of the ancestors belonging to each herd.

RESULTS

None of the genotyped MS was in Hardy-Weinberg equilibrium (HWE) when considering the population as a whole. The lack of HWE might be an indicator of the presence of a stratified genetic structure. The F_{ST} differentiation index among herds was on average 0.074. The average molecular coancestry within (f_{ii}) and among (f_{ij}) herds was 0.329 and 0.278, respectively. We estimated the different statistics with Metapop software (Perez-Figueroa *et al*. 2009) except for the HWE which was calculated with Genepop (Rousset 2008).

STRUCTURE algorithm. STRUCTURE inferred the existence of three genetic clusters and assigned a proportion of ancestry coming from each cluster for all participating individuals. Animals were assigned to a certain cluster when at least 90% of their genome (the proportion of inferred ancestry given by STRUCTURE) was coming from that cluster. According to this

definition of clusters, there were 1134, 1054 and 1015 animals in the first, second and third clusters, respectively. Those animals that were not included in the clusters were grouped together in a pool (ST-Pool). The number of loci in HWE within the clusters increased with respect to the whole population; 16, 9 and 11 MS were found to be in HWE for clusters 1, 2 and 3, respectively. Only 2 MS were in HWE in the ST-Pool. Clusters f_{ii} were higher than f_{ij}, as expected, while the f_{ii} of the ST-Pool was the same as the f_{ij} of herds (Table 1).

We analysed the distribution of herds across the clusters and found two types of herds, those whose individuals were mostly associated to one specific cluster and those whose individuals were distributed among different clusters. It should be noted that the majority of individuals in most herds were assigned to the ST-Pool. However, there were nine herds with a majority of individuals not assigned to the pool but to a specific cluster.

Graphical clustering algorithm. The GCA also identified three clusters. Cluster sizes were 257, 534 and 458 for clusters 1, 2 and 3 respectively. Those animals not assigned to any cluster were also grouped in pool (GCA-Pool). In this case, 15, 16 and 15 MS were in HWE in clusters 1, 2 and 3, respectively. Only 1 microsatellite was in HWE in the GCA-Pool. As expected, f_{ii} was higher than f_{ij} (Table 1). Both the f_{ii} and the f_{ij} of the GCA clusters were larger than the ones of the clusters inferred by STRUCTURE.

Table 1. Molecular coancestry within (f_{ii} in diagonal) and across (f_{ij} off-diagonal) the genetic clusters of the Avileña-Negra Ibérica population inferred by STRUCTURE software (values on the left of the slash) and the graphical clustering algorithms -GCA- (on the right of the slash)

Cluster	1	2	3	ST-Pool/CGA-Pool
1	0.343/0.427	0.244/0.327	0.241/0.306	0.271/0.280
2	0.244/0.327	0.311/0.409	0.239/0.366	0.276/0.316
3	0.241/0.306	0.239/0.366	0.343/0.417	0.271/0.301
ST-Pool/CGA-Pool	0.271/0.280	0.276/0.316	0.271/0.301	0.278/0.269

In line with STRUCTURE, GCA described the same two types of herds; those whose animals are mostly associated to one cluster and those associated to different ones. Furthermore, the distribution of herds across clusters was very similar. However, in this case a higher percentage of individuals of a herd were assigned to the GCA-Pool. In general, many herds had individuals assigned to clusters 2 and 3 as expected from the f_{ii} and f_{ij} values (Table 1). This connection between clusters 2 and 3 is also observed in the analysis of herds of STRUCTURE.

Genetic contribution of herds. Three herds were the origin of the majority (56.5%) of the genes in the population in 2012. These three herds are among those nine herds whose animals are mostly associated to one specific cluster. Furthermore, each of them appeared assigned to a different cluster both in STRUCTURE and GCA solutions.

DISCUSSION

STRUCTURE and the GCA inferred similar genetic structures suggesting that the results are robust. However, there were important differences in terms of the computational time. When using the K determination criterion suggested by Evanno *et al.* (2005), the STRUCTURE algorithm needs to be run several times per K to calculate a standard deviation of the replicates. In our case, we tested the algorithm from K= 1 to 50 which took several weeks to run. GCA took less than an hour to obtain the whole set of solutions. According to the results of both analyses, the ANI population can be divided in three genetic clusters and one pool of animals that could not be

assigned to any of the clusters and that grouped the majority of animals. We set very strict criteria in both approaches for animals to be assigned to a cluster, aiming to get the core ANI population structure, given its expected complexity. Thus, we expect that the size of the pool would be reduced if the criteria were looser or, as observed in human population, once a first level of stratification is identified then new stratification levels may appear. The average HWE across MS increased within the clusters, validating the clustering, while the pool was in HW disequilibrium indicating that there may be secondary genetic structures within it. The distribution of individuals in different herds among the three clusters was quite consistent when comparing GCA and STRUCTURE inferences. Both approaches gave the same solution regarding the herds that are mostly associated to one cluster; each of the three herds that, according to the pedigree analysis, have had a major genetic contribution to the breed was found in each of the three clusters. However, there were some differences among those herds that cannot be clearly associated to one cluster. The results of the analysis indicate that currently ANI is stratified in three linages with a number of herds mainly related to one of the lineages and then, another group of herds whose individuals are distributed across lineages. Due to the large number of herds and individuals included in this last group and the limited number of herds with a major contribution to the breed, it will be of great importance to understand the composition of the pool and see how it is related with the genetic variability of the breed.

ACKNOWLEDGEMENTS

We thank ANI breeders Association for providing the data and helpful comments for the results interpretation. DMC was funded by the Collaboration Agreement (CC09-009) between INIA and FEAGAS supported by Spanish Ministry of Agriculture, Food and Environment. KJA wished to acknowledge the support of Capes-Brazil (Programa PVE) .

REFERENCES

Abraham K.J. and Fernando R. (2012) In 'New Frontiers in Graph Theory', pp. 307-322, editor Y. Zhang, InTech publishers, New York.

Abraham K.J., Martin-Collado D., Toro M.A., Carabaño M.J, Rodriguez-Ramilo S. and Diaz C. (forthcoming) A graph theoretical approach to estimating the number of constituent populations in an admixed population.

Caballero A. and Toro M.A. (2002). *Conserv. Genet.* **3**: 289.

Evanno G., Regnaut S. and Goudet J. (2005) *Mol. Ecol.* **14**: 2611.

Gutiérrez J.P. and Goyache F. (2005) *J. Anim. Breed. Genet.* **122**: 172.

Janss L., de los Campos G., Sheehan N. and Sorensen D. (2012) Genetics, **192**: 693.

Pérez-Figueroa A., Saura M., Fernández J., Toro M.A. and Caballero A. (2009). *Conserv. Genet.* **10**: 1097.

Pritchard J.K., Stephens M. and Donelly P. (2000) *Genet.* **155**: 945.

Rousset F. (2008) *Mol. Ecol. Resour.* **8**: 103.

Vasallo J.M. and Díaz C. (1986) *Livest. Prod. Sci.* **15**: 285.

Williams T.D., Turan N., Diab A.M., Wu H., Mackenzie C., Bartie K., Hrydziuszko, O., Lyons B.P., Stentiford G.D., Herbert J.M., Abraham J.K., Katsiadaki I., Leaver M.J., Taggart J.B., George S., Viant M.R., Chipman K.J. and Falciani F. (2011) *Plos Comput. Biology* **7**: e1002126.

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:138-141

ACROSS-BREED GENOMIC EVALUATION BASED ON BOVINE HIGH DENSITY GENOTYPES, AND PHENOTYPES OF BULLS AND COWS

C. Schrooten[1], G.C.B. Schopen[1], A. Parker[2], A. Medley[2] and P. Beatson[2]

[1]CRV BV, P.O. Box 454, 6800 AL Arnhem, the Netherlands
[2]CRV AmBreed, NZ PO Box 176, Hamilton 3240, New Zealand

SUMMARY

Most applications of genomic selection are based on a reference population of bulls only, genotyped with 50k SNP-chips. In some populations, the size of the reference population is limited, resulting in relatively low reliabilities of genomic breeding values. In this study we looked at the possibility of expanding the reference population by combining several breeds in one genomic evaluation, and making use of reference cows in addition to reference bulls. Because such an evaluation needs genotypes at higher density than 50k, high density (777k) SNP-chip genotypes were used. Presentation of results was limited to 7 traits. On average, reliabilities were 1-4% higher than reliabilities from a single breed evaluation using a bull reference population with 50k SNP-chip genotypes, and 0-2% higher than reliabilities from an across-breed evaluation based only on reference bulls and high density genotypes.

INTRODUCTION

Genomic evaluation at CRV Ambreed has been based on 50k SNP-chip genotypes and single-breed reference populations of bulls. The individual reference populations for Friesians and Jerseys consist of approximately 2,200 and 1,200 reference bulls, respectively. These reference populations are relatively small, compared to the reference populations in North America (VanRaden 2010) and Europe (Lund *et al*. 2011). Therefore, in these small populations, the use of genomic information is predicted to result in only a moderate increase in reliability of breeding values of animals without phenotype.

Reliabilities may increase further if reference populations are combined in a multi-breed genomic evaluation. To make use of genomic information across breeds, markers must be in Linkage Disequilibrium (LD) with the mutations affecting the trait of interest, and the linkage phase must be the same in the individual breeds. De Roos *et al.* (2008) looked at LD and phase persistency in Holstein Friesian, Jersey, and Angus populations in Australia, New Zealand and the Netherlands. They concluded that strong enough and persistent LD could be obtained when genotyping with at least 300k SNP. Therefore, to combine reference populations for CRV Ambreed, a higher density is needed than obtained with the currently used 50k SNP-chip.

The reference population can also be expanded by adding cows with phenotypic information to the reference population. Because reliability of phenotypic information is lower for cows than for bulls, the benefit of adding a certain number of cows to the reference population is lower than the benefit of adding the same number of bulls. Nevertheless, when no additional bulls are available and genotyping cost are sufficiently low, cows offer a good opportunity to expand the reference population.

The objective of this study was to estimate the effect on reliability of genomic breeding values, when single-breed reference populations are combined, and the reference population is augmented with high density genotypes and cow genotypes and phenotypes.

MATERIALS AND METHODS

Genotypes. Genotypes of 465 Friesians, 227 Jerseys and 57 crossbreds were obtained using the Illumina BovineHD BeadChip, containing 777k SNP-markers:

(http://www.illumina.com/Documents/products/datasheets/datasheet_bovineHD.pdf).

Genotypes of approximately 9,000 animals, obtained with 50k chips, were imputed to BovineHD with Beagle version 3.0 (Browning and Browning, 2007), using the 749 HD-genotyped animals as reference set for imputation. After data edits, 9,486 animals were available for evaluation. Ancestral haplotype scores were obtained for 622k loci on the 29 autosomes. To reduce computer requirements for genomic evaluation and because the full SNP-set contains redundant information due to complete or nearly complete LD between neighboring SNP, the number of HD loci based on hidden states (i.e. ancestral haplotypes) was reduced by considering only the first locus out of every 10 consecutive loci in genomic evaluation. After this reduction, 62,302 loci remained for evaluation.

Genomic evaluation. Genomic evaluation was performed for 26 traits, but presentation of results will be restricted to a subset of 7 traits with moderate to high heritability that are part of the New Zealand Merit Index (NZMI, http://www.crv4all.co.nz/Library/NZMI.html). Depending on the trait, de-regressed proofs (DRP) of 3,200-3,700 bulls and 1,300 – 2,600 cows were available. Effective daughter contributions (EDCs) were used as weight for the phenotype. Phenotypes of the youngest cohorts of bulls (born from 1-1-2005 onwards), and their daughters were not used to estimate effects. About 150-200 cow phenotypes for each breed were not included in the genomic evaluation because these cows were sired by a validation bull. The bulls from these cohorts were considered a validation bull if they had a genotyped sire with a phenotype, and no genotyped sons with phenotype. Furthermore, animals (cows) with phenotype reliability below a trait-dependent threshold (ranging from 0.25 to 0.50) were not used to estimate effects, because initial analyses indicated a negative effect on reliability when including low reliability phenotypes. The minimum reliability per trait was chosen on the one hand to discard records with a very low reliability and on the other hand to keep enough records to be able to estimate the impact of cow data on the reliability of genomic proofs.

The following model was used for genomic evaluation to obtain genomic breeding values (GBV):

$$y_i = \mu + u_i + \sum_j^n (q_{ij1} + q_{ij2})v_j + e_i$$

where y_i is the deregressed proof of bull i, μ is the overall mean, u_i is the random polygenic effect of bull i, n is the total number of loci, v_j is the direction vector of the effects of the haplotypes at locus j, $q_{ij1\,(2)}$ is the size of the effect for the first (second) haplotype ID of animal i at locus j and e_i is the residual for bull i. A Markov Chain Monte Carlo method using Gibbs sampling was used to obtain posterior estimates for all effects in the model. The Gibbs sampler was run for 10,000 iterations with a 2,000 burn-in. Four replicates per trait were run. More details on the method can be found in Meuwissen and Goddard (2004) and Calus et al. (2008).

Pedigree based breeding values (PBV) were estimated with the same data using a comparable model without genomic information:

$$y_i = \mu + u_i + e_i$$

Validation. The genomic prediction of the validation bulls was compared to their daughter-based phenotype, as an approximation of increased reliabilities due to genomic information. Squared correlations (R^2) between DRP and both GBV and PBV were computed and compared to each other to obtain ΔR^2 using the following formula:

$$\Delta R^2 = \frac{R^2_{GBV,DRP} - R^2_{PBV,DRP}}{REL_{DRP}}$$

where $R^2_{GBV,DRP}$ is the squared correlation between GBV and DRP, $R^2_{PBV,DRP}$ is the squared correlation between PBV and DRP, and REL_{DRP} is the reliability of the DRP.

The increase in GBV reliability (measured as ΔR^2) in the reduced HD loci subset was compared to the increase in GBV reliability when using only HD genotyped bulls, and compared to the increase in GBV reliability of the routine genomic evaluation (50k bull genotypes) obtained in earlier validations.

RESULTS AND DISCUSSION

In this paper, presentation of results was limited to a subset of 11 traits out of 26 traits, all with heritability >= 0.15. Nine of the 26 analyzed traits had a heritability below 0.15. Only one (Jerseys) or three traits (Friesians) out of these nine traits benefited from genomics when a single-breed reference population consisting of bulls was used. Therefore, the comparison with results from a multi-breed evaluation with cows included was not made for traits with heritability below 0.15.

The number of reference bulls ranged from 3,276 (udder, Table 1) to 3,548 (protein and milk). The number of reference cows showed more variation, from 1,584 (rump angle) to 2,640 (milk), mainly due to the applied threshold for minimum reliability of the phenotype. The number of bulls used for validation was 345 or 346 for Friesians, 163 for Jerseys and 56-67 for crossbreds. Average increase in R^2 due to genomic information was 9.4% and 10.9% for Friesians and Jerseys, respectively, when an across-breed evaluation was performed using high density genotype data and reference bulls as well as reference cows. With only bulls genotyped with 50k as reference animals in a single breed analysis, increase in R^2 was 8.6% (Friesians) and 7.0% (Jerseys). This indicates that Jerseys, which initially had the smallest reference population, gained most from expanding the reference population. Without cows in the reference population (results not shown), average increase in R^2 due to genomic information was 9.3% (Friesians), 8.5% (Jerseys), and 7.6% (crossbreds). For Jerseys, both reference cows and animals from the other breed seem to contribute to higher reliabilities of genomic breeding values, although not all traits show this result. For crossbreds, increase of R^2 was 1.2% higher when cows were added to the reference population.

For Friesians and Jerseys, the across-breed evaluation with cows added to the reference population gave higher reliabilities for 10 out of 14 analyzed trait-breed combinations, compared to reliabilities when only 50k-genotyped reference bulls were used. Exceptions were protein and milk in Friesians, and milk and udder overall in Jerseys.

The size of the bull reference population ranged from 3,276 to 3,548, whereas the number of cows added to the reference population was lower, ranging from 1,584 to 2,640. Because reliability of cow phenotypes is lower, the amount of information added to the genomic evaluation is relatively low when converted to bull equivalents. Nevertheless, there was benefit from adding a relatively low number of cows to the reference population for most trait-breed combinations presented here. This offers opportunities to further increase reliabilities of genomic breeding values by adding more cows to the reference population.

Table 1. For each trait: number of reference bulls (Nrefb) and cows (Nrefc), and for each breed: increase in R^2 due to genomics for the across-breed evaluation based on BovineHD genotypes and a reference population of bulls and cows (ΔR^2), increase in R^2 in a single-breed evaluation based on 50k genotypes and a reference population of bulls only (ΔR^2s). For crossbreds, number of validation bulls (Nval) is indicated

			Friesian[1)]		Jersey[1)]		Crossbred[2)]	
Trait	Nrefb	Nrefc	ΔR^2 (%)	ΔR^2s (%)	ΔR^2 (%)	ΔR^2s (%)	Nval[1)]	ΔR^2 (%)
Protein	3548	1985	5.4	11.7	7.0	4.7	67	12.2
Milk	3548	2640	11.6	15.4	7.0	20.4	67	9.8
Liveweight	3343	2313	7.8	3.5	16.4	4.9	59	7.7
Somatic Cells	3493	2275	8.8	6.9	18.5	5.1	66	12.5
Capacity	3470	2072	10.3	9.5	8.7	4.0	59	-0.2
Rump angle	3281	1584	12.9	11.2	11.2	1.3	57	11.1
Udder overall	3276	2481	9.3	2.1	7.2	8.4	56	8.5
Average			9.4	8.6	10.9	7.0		8.8

1) Number of validation bulls was 345-346 for Friesians, and 163 for Jerseys
2) No results from single breed evaluation based on 50k genotypes available for Crossbreds

CONCLUSIONS

Averaged across 7 analyzed traits, across-breed genomic evaluation resulted in 0.7% (Friesian) and 1.0% (Jersey) higher reliabilities than the single-breed genomic evaluation based on 50k genotypes. Adding cows to the reference population was beneficial in Jerseys and crossbreds, where reliabilities increased with 2.2% (Jersey) and 1.2% (crossbreds). Considering that a relatively low number of cows was added, higher reliabilities can be obtained by adding more cows.

REFERENCES

Browning S.R. and Browning B.L. (2007) *Am. J. Hum. Genet.* **81**: 1084.
Calus M.P.L., Meuwissen T.H.E., de Roos A.P.W., and Veerkamp R.F. (2008) *Genetics* **178**: 553.
De Roos A.P.W, Hayes B.J, Spelman R.J. and Goddard M.E. (2008) *Genetics* **179**: 1503.
Lund M.S., de Roos A.P.W., de Vries A.G, Druet T., Ducrocq V., Fritz S., Guillaume F., Guldbrandtsen B., Liu Z., Reents R., Schrooten C., Seefried F. and Su G. (2011) *Genet. Sel. Evol.* **43**: 43.
Meuwissen T.H.E. and Goddard M.E. (2004) *Genet. Sel. Evol.* **36**: 261.
VanRaden P. (2010). *Proc. 9th WCGALP* **27**: 8.

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:142-146

LACTATIONAL PERFORMANCE OF STRAIGHTBRED ANGUS COWS AND THREE ANGUS-DAIRY-CROSS GENOTYPES

F.J. Roca Fraga, R.E. Hickson, N. Lopez-Villalobos, P.R. Kenyon and S.T. Morris

Massey University, Private Bag 11222, Palmerston North 4442, New Zealand.

SUMMARY

The objective of this study was to determine the milk yield and lactation curves of 117 three-year-old Angus (AA), Angus×Friesian (AF), Angus×Jersey (AJ) and Angus×Kiwi-Cross (AK) cows, rearing singleton calves sired by either Angus or Simmental bulls. Milk yield was estimated using the weigh-suckle-weigh technique (WSW) recorded at day 32, 49, 80, 120 and 160 days postpartum (dpp). A third-order Legendre polynomial was fitted to lactation data using random regression. Cows from AF, AJ and AK groups reached peak lactation at a similar ($P>0.05$) stage of lactation (71±7, 74±6 and 82±9 days, respectively), but later ($P<0.05$) than AA cows (46±6 days). Peak milk yield was the greatest ($P<0.05$) for AF cows (12.1±0.4 kg) followed by AK and AJ cows (11.4±0.5 kg 10.9±0.3 kg, respectively). Cows from AJ and AK groups produced more milk at peak lactation ($P<0.05$) than AA cows (9.8±0.31 kg). Overall, AF cows produced more ($P<0.05$) milk from day 32 to day 160 (1337.3±22.3 kg) than AJ and AA cows (1244.5±19.9 kg and 1017.5±19.99 kg, respectively), although their milk production did not differ ($P>0.05$) from AK cows (1307.8±31.6 kg). Angus×Friesian, AJ and AK cows produced more milk throughout lactation than AA cows.

INTRODUCTION

Maternal milk production is one of the most important factors affecting the weaning weight of calves and the production costs associated with the maternal metabolisable energy requirements (Montano Bermudez *et al*. 1990). Therefore, the profitability of the cow-calf producer is directly affected by changes in the lactational performance of beef cows. Introduction of dairy genetics into a beef cattle herd can result in higher milk yields in cows (Deutscher andWhiteman 1971), higher calf growth rates and better biological and economic efficiencies regarding beef production (Morris 2008). Morris (2008) suggested that both Friesian- and Jersey-cross cows are highly adapted to New Zealand's pastoral conditions and consequently have high potential for use as suckler cows. The objective of this study was to identify how the inclusion of dairy genetics into a straightbred Angus herd may have affected the characteristics of lactations curves of the beef × dairy cows compared to the straightbred Angus cows.

MATERIAL AND METHODS

One hundred and seventeen 3-year-old Angus (AA; n=43), Angus×Friesian (AF; n=32), Angus×Jersey (AJ; n=40) and Angus×Kiwi-Cross (AK; n=21) cows (Hickson *et al*. 2012) rearing calves sired by Angus (n=4) or Simmental (n=4) bulls were used in this study. Cows with single calves were allocated into one of three groups based on their calving date, i.e. early (E), mid- (M) and late (L) calving cows, respectively. Cows were grazed under pastoral conditions to an average post-grazing cover of 1534 kg DM/ha at Massey University's Tuapaka Farm. Calves remained with their dams until weaning at an average 148±19 days of age. The milk production of cows was estimated using the weigh-suckle-weigh technique (WSW) on 5 occasions for groups E and M at an average 32, 49, 80, 120 and 148 days postpartum (dpp) and on 4 occasions for group L at an average 49, 80, 120 and 148 dpp.

Statistical analyses were carried out using the Statistical Analysis System (SAS version 9.2, SAS Institute Inc., Cary, NC, USA, 2009). Third-order Legendre polynomials (standardized for units of time) were fitted to lactation data using a random regression to obtain an average lactation curve for the population and for each cow using the following model:

$$y_{tm} = \sum_{i=0}^{3} b_i P(x)_t^i + \sum_{i=1}^{3} \alpha_{im} P(x)_{tm}^i + e_{tm}$$

where y_{tm} is the observation at time t in cow m for daily milk yield, b_i are fixed regression coefficients of days in milk on variable *y* (b_0 = intercept, b_1 = linear effect, b_2 = quadratic effect and b_3 = cubic effect); α_{im} are random regression coefficients of days in milk on variable *y* in cow *m* (α_{0m} = intercept, α_{1m}=linear effect, α_{2m}= quadratic effect and α_{3m} = cubic effect), x_{tm}^i is the observation of standardized days in milk at time *t* in cow *m* at the power 0, 1, 2, and 3; e_{tm} is the residual error associated with observation y_{tm}.

Random coefficients for each cow were obtained using the MIXED procedure assuming an unstructured covariance structure between the variance and covariances of the random regression coefficients of the model. Using the estimated random regression coefficients for each cow, parameters of the lactation curve for each cow were estimated. Analysis of variance for each of these parameters was performed with the MIXED procedure with a linear model that included the fixed effects of breed of the dam, mob, sex of calf, and the interaction between sex of calf and breed of the dam.

RESULTS

Predicted lactation curves from 32 to 160 dpp based on test-day records varied across genotypes and the shapes can be classified into two types according to the pattern of milk production (Figure 1). Type 1 (AA group): from Figure 1 it is evident that the highest milk production from AA cows was likely to have occurred somewhere during the first month of lactation or before 32 dpp, with a slow decrease until 83 dpp, thereafter it remained fairly constant until approximately 115 dpp after which it decreased rapidly towards weaning; Type 2 (AF, AJ and AK groups): milk production increased continuously from the beginning of the lactation period until it reached a peak around 80 dpp and then decreased until the end of the lactation. The effect of birth weight and differing breed proportions in the calf on lactation curve shape were investigated, however, no effect of bull breed ($P>0.05$), proportion of maternal breed in the calf ($P>0.05$) or birth weight of the calves ($P>0.05$) was observed. This indicates that it was primarily the genotype of the cow that determined milk production in this study.

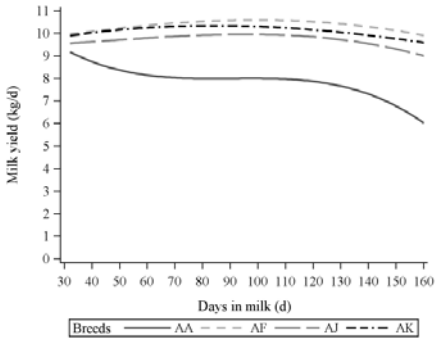


Figure 1. Milk yield from day 32 to day 160 of lactation, of straightbred Angus (AA), Angus×Friesian (AF), Angus×Jersey (AJ) and Angus×Kiwi-Cross (AK) cows.

The corrected least square means and SE for total milk yield from 32 to 160 dpp, milk yield at peak lactation, days in milk at which peak lactation was reached and milk yield at weaning are shown in Table 1. Absolute days at which peak milk yield was reached differed between AF, AJ and AK cows, although differences were not significant (P>0.05). Pure Angus cows differed (P<0.05) from the crossbred groups and reached peak lactation at 46 dpp. Angus×Friesian cows produced more milk (P<0.05) during peak lactation than AJ and AA cows but did not differ (P>0.05) from AK cows. Angus×Jersey and AK cows produced more milk at peak (P<0.05) than AA cows. On average, beef-cross-dairy crossbreds produced approximately 2 kg more milk during peak lactation than AA cows. The sex of calf affected milk production at peak lactation, such that dams nursing female calves (P=0.05) produced less milk (approximately 0.8 kg) than those nursing male calves.

Angus×Friesian cows produced more milk at weaning (P<0.05) than AJ and AA cows but not AK cows. The AA cows had the lowest milk yield at weaning with an average difference compared to the other genotypes of 3.4 kg. The AF, AJ and AK cows produced more (P<0.05) milk from 32 to 160 dpp than the AA cows. The AF cows produced more (P<0.05) milk from 32 to 160 dpp than AJ cows, with AK cows being intermediate and not differing (P>0.05) from either AF or AJ cows. In the present study, as the proportion of Friesian or Jersey in the crossbreds increased from 0 to 50%, an extra 325 kg and 240 kg of milk, respectively, was expected compared with the AA cows. Given that a Kiwi-Cross is a Friesian-Jersey hybrid, AK cows' production was intermediate between AF and AJ cows, and they produced an extra 282.5 kg of milk compared with the AA cows.

Table 1. Least square means and standard error for the lactation curve parameters of Angus (AA), Angus×Friesian (AF), Angus×Jersey (AJ) and Angus×Kiwi Cross (AK) cows.

Genotypes	Milk yield (kg/d)			Day of peak lactation (d)
	Peak	Weaning 160dpp	Estimated Total†	
AA	9.8±0.3^{c}	6.1±0.3^{c}	1017.5±19.9^{c}	46±6^{b}
AF	12.1±0.4^{a}	9.8±0.3^{a}	1337.3±22.3^{a}	71±7^{a}
AJ	10.9±0.3^{b}	8.9±0.3^{b}	1244.5±19.9^{b}	74±6^{a}
AK	11.4±0.5^{a,b}	9.5±0.4^{ab}	1307.8±31.6^{ab}	82±9^{a}

† Total milk yield from day 32 to day160 postpartum.
^{abc} Values within columns with different superscripts differ at the P<0.05 level.

DISCUSSION

The lactation curves for beef × dairy cows are of similar shape and resemble a typical lactation curve seen in dairy cows. The findings in the present study are similar to those of Walker and Pos (1963) in New Zealand, where AF and AJ cows reached peak lactation at an average 74 dpp; and to those reported by Chennete and Frahm (1981) whereby peak lactation in AJ cows was detected at approximately 70 dpp, followed by a steady decrease as lactation progressed. Post peak lactation, milk production levels were maintained until approximately 120 dpp when a decrease in milk production occurred for all three crossbred genotypes. Gaskins and Anderson (1980) reported peak lactation in AJ cows during the first month of lactation, which is earlier than reported in the present study, however, milk yield remained constant until 84 dpp in 2-year-old cows and until 112 dpp in 3-year-old cows. These results suggest lactation curve persistency in beef cows can be greatly improved by the introduction of genes from dairy breeds.

The days in milk at peak lactation reported here for AA cows is earlier compared to the findings of Jenkins and Ferrell (1984) and Hohenboken *et al*. (1992) using the "Jenkins" equation, however, the latter equation has been criticised since it repeatedly produced curves that peaked approximately around 60 and 70 dpp and also underestimated milk yield during the first month of lactation since if forces the curve through the origin (Hohenboken *et al*. 1992).

There is evidence (Oftedal, 1984) that the calf's capacity to withdraw milk may be reduced in early lactation and that the residual milk left in the udder would stimulate mammary involution. An interaction exists between mammary evacuation and milk production, where cows suckled or milked more often produce higher levels of milk than those with infrequent mammary evacuation. Angus cows may be more sensitive to changes in mammary evacuation during early and late-lactation than the beef x dairy crossbred cows. The first drop in production seen in AA cows may be explained by the calf not being physically capable of fully evacuating the udder in the first few days due to physical consumption constraints and consequently, the residual milk left in the udder would stimulate the dam to reduce her milk production (Oftedal 1984). Then, as the calf grows and its ability to suckle increases, milk production stabilises at a level lower than peak to provide nutrients to the calf. Indeed, Blaxter (1961) suggested that milk yield is motivated towards the maximum possible growth rate of the offspring. Thus it is likely that a dams' milk production would respond to the stimulus from her calf, although as a non-dairy animal, AA cows may have limited capacity to produce more milk. Energy requirements increase with increasing age and there is evidence (Baker *et al*. 1976) that if forage availability is adequate, calves receiving less milk during lactation could increase their pasture intake to compensate for the low energy intake from milk. In this study during late-lactation, calves reared by AA cows may have not been receiving enough nutrients from a suckling event; therefore to compensate they reduced their milk consumption in favour of consuming more pasture. This may explain the second drop in milk production observed towards the end of the lactation of AA cows, which is typical in AA cows in other studies (Minick *et al*. 2001).

CONCLUSIONS

Results from this experiment confirmed the hypothesis that increasing the proportion of dairy genetics in the beef herd is accompanied with an increase in milk production. Under non-limiting pasture quality and availability, AF, AJ and AK cows produced more milk throughout lactation than AA cows.

REFERENCES

Baker R.D., Le Du Y.L.P. and Barket J.M. (1976) *J. Agric. Sci*. **87**: 187.
Blaxter K.L. (1961) In 'Milk: the mammary gland and its secretion', pp. 305-361, editors S.K.

Kon and A.T. Cowie, New York Academic Press, New York.
Chenette C.G. and Frahm R.R. (1981) *J. Anim. Sci*. **52**: 483.
Deutscher G.H. and Whiteman J.V. (1971) *J. Anim. Sci*. **33**: 337.
Gaskin C.T. and Anderson D.C. (1980) *J. Anim. Sci*. **50**: 828.
Hickson R.E., Laven R.L., Lopez-Villalobos N., Kenyon P.R. and Morris S.T. (2012) *Anim. Prod. Sci*.**52**: 478.
Hohenboken W.D., Dudley A. and Moody D.E. (1992) *Anim. Prod*. **55**: 23.
Jenkins T.G. and Ferrell C.L. (1984) *Anim. Prod*. **39**: 479
Minick J.A., Buchanan D.S. and Rupert S.D. (2001) *J. Anim. Sci*. **79**: 1386.
Montano Bermudez M., Nielsen M.K. and Deutscher G.H. (1990) *J. Anim. Sci*. **68**: 2279.
Morris S.T. (2008)'Unpublished review'. Meat and Wool. N.Z. pp. 31.
Oftedal O.T. (1984) *Sym. Zool. Soc. Lon*. **51**: 33.
Walker D.E.K. and Pos H.G. (1963) *N.Z. J. Agric*. **107**: 227.

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:147-150

POPULATION STRATIFICATION AND BREED COMPOSITION OF AUSTRALIAN TROPICALLY ADAPTED CATTLE

L.R. Porto-Neto[1], S.A. Lehnert[1], M.R.S Fortes[2], M. Kelly[2] and A. Reverter[1]

[1] CSIRO Food Futures Flagship and Animal, Health and Food Sciences,
306 Carmody Road, St Lucia, QLD 4067
[2] The University of Queensland, QAAFI Centre for Animal Science, St Lucia QLD 4072

SUMMARY

Population stratification and differences in individuals' ancestry can potentially bias genome-wide genetic analyses when they are not detected and included in the genetic model. This is especially important in situation where little is known about the extent and sources of stratification. Here a large sample of tropically adapted cattle, Brahman (BB) and Tropical Composite (TC), genotyped for more than 700K SNP loci were evaluated for population stratification using principal components and supervised hierarchical clustering analyses. The BB cattle were more homogeneous than the TC cattle in both analyses, reflecting the TC's more recent and complex origin. Nevertheless, within both breeds there were degrees of variability. The effect of farm of origin was also noticeable, particularly in TC. These analyses indicated that a simple breed designation, BB or TC, encompasses large variation in ancestry within breed. This opens the question whether ancestry composition should be included in downstream analyses. Appropriate use of information on ancestry composition could aid genome-wide association studies and genomic selection.

INTRODUCTION

The detection of population stratification and estimation of ancestry composition are *per se* a field of study that is fits within population genetics and dynamics. There are several factors that might create stratification of a population, some with real biological meaning and others due to experimental artifact. It has been shown that population stratification can cause spurious associations in genome-wide studies (Price *et al.* 2006; Ma *et al.* 2012). Therefore, information on both stratification and ancestry, are very relevant either as a final result or to be taken into account in genome-wide genetic analyses.

In Australia, most beef production operations are located in Northern regions, where the climate is warm, the environment is tropical and infested with parasites. Under these conditions, tropical adaptation is imperative for cattle to thrive. *Bos primigenius indicus* or Zebu cattle (e.g. Nelore) and *Bos primigenius taurus* or Taurine (e.g. Angus) evolved under different environmental pressures, and these natural adaptations are exploited by cattle breeders to improve herd productivity. A good example in Australia is the expansion of Brahman (BB), a Zebu breed that was graded up using Taurine cattle, and the Tropical Composite (TC), which involves crosses of Zebu, Taurine and, in some cases, African cattle. Given their formation, it is expected that both breeds would have a range of ancestry compositions.

In this study a large sample of tropically adapted cattle, BB and TC was evaluated for its potential population stratification and individual ancestry composition were estimated. Furthermore, the estimated ancestry composition was compared to farm origin of the animals.

MATERIAL AND METHODS

Animals. The tropically adapted breeds of BB (n=3,502) and TC (n=2,550), representing 21 and 7 different farms of origin were included in this study. These cattle were from the CRC for Beef Genetic Technologies, Beef CRC, which was described previously (Barwick *et al.* 2009;

Burns *et al.* 2013). To anchor the breed composition estimation a sample of Zebu cattle represented by the Nelore (n=29) and Taurine represented by the Angus (n=81) were included in the analysis.

Genotypes. Cattle were genotyped using either the BovineSNP50 or BovineHD array (Illumina Inc., San Diego, CA 2006). Animals that were genotyped using the smaller array were imputed to a higher density using Beagle 3.2 (Browning and Browning 2009). To reduce the potential bias in the analyses due to a large number of markers with high linkage disequilibrium (LD) and to reduce the computational time, the full dataset was pruned by LD using PLINK v.1.07 software (Purcell *et al.* 2007) to exclude one SNP of a pair that had $r^2>0.5$ calculated in a sliding window of 50 SNP. After pruning, the combined BB and TC dataset included 229,235 SNP genotypes.

Population structure and breed composition estimation. The structure of the population was explored by principal components analysis of the genetic relationship matrix based on the SNP genotypes, both calculated using GCTA (Yang *et al.* 2011). The breed composition estimation was performed using a supervised hierarchical clustering implemented in Admixture (Alexander *et al.* 2009) set at K=2 subpopulations, using the Nelore and Angus breeds as the two reference clusters.

RESULTS AND DISCUSSION

The genetic relationship matrix that considers the four different cattle breeds in a principal components analysis is expected to result in clusters that agree with breed designations. The main split in domestic cattle correspond to Zebu vs Taurine cattle, which can be argued to have occurred 330,000 years ago (Achilli *et al.* 2008). This split, here represented by the Nelore vs Angus distance, accounts for most of the variation resulting in extreme positions in the first principal component (Figure 1), with the main TC cluster positioned approximately half way between Nelore and Angus. This is in agreement with results previously described for other composite breeds (Harrison *et al.* 2009). Comparing the tightness of the Nelore and Angus clusters to TC and BB it is clear that there is more variation within the latter two breeds. This large variation within breed is particularly evident in the TC cluster. However, variation is also seen within BB, where a number of individuals are positioned closer to the TC and Angus clusters.

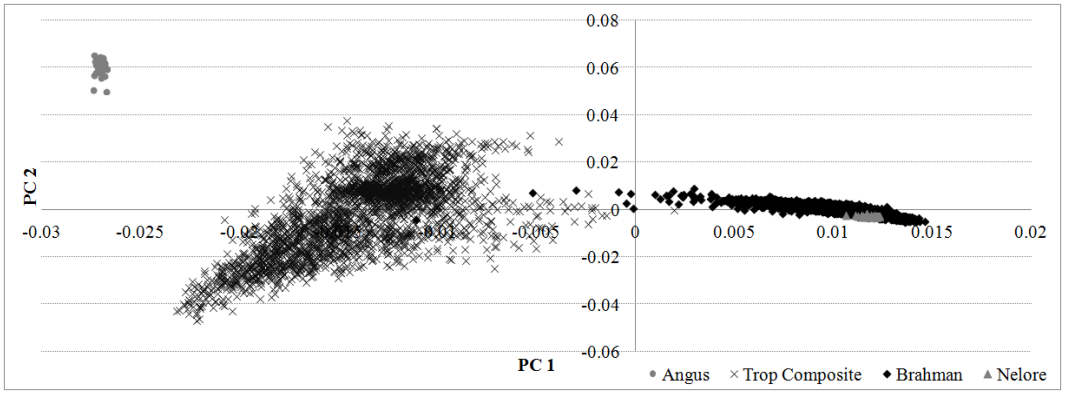


Figure 1. First Cartesian plan of principal components of the genetic relationship matrix between Angus, Tropical Composite, Brahman and Nelore animals.

In hierarchical clustering analyses TC estimated ancestry proportions attributed to Zebu varied from 0.00 to 0.79, and averaged 0.35 (Figure 2). In BB the estimated Zebu content varied from

0.46 to 1.00, and averaged 0.95. Once again, the BB cattle showed more homogeneity than TC. The Angus and Nelore reference breeds were chosen as proxies of original Taurine and Zebu cattle. However, other breeds also contributed to the formation of TC and BB. Whether the inclusion of other reference breeds in the analysis would improve the resolution of those estimates remains to be tested.

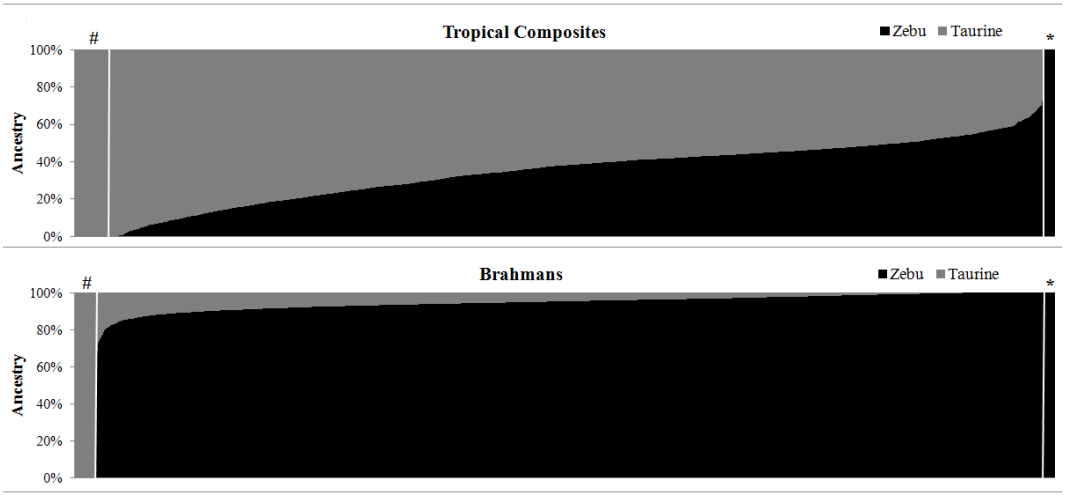


Figure 2. Breed composition (Zebu vs Taurine) estimated by supervised hierarchical clustering. A vertical bar represent each individual along the x-axis, Tropical Composites (n=2,550) and Brahmans (n=3,502) using Nelore (*) and Angus (#) as reference populations.

It is a reasonable assumption that within a breed designation, the farm of origin of an individual would reflect the Zebu vs Taurine proportion of its ancestry. However, as shown in Figure 3, the Zebu content varied between and within farm of origin with large standard deviations, and also within a breed designation.

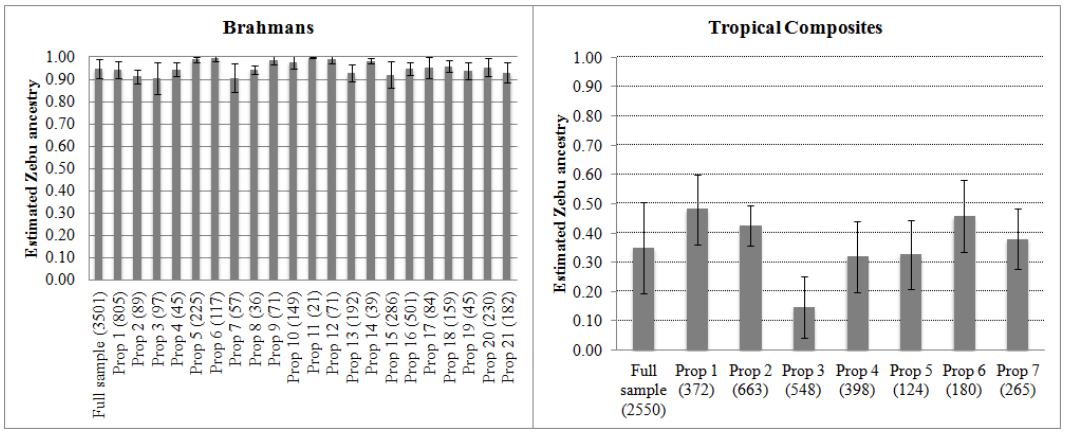


Figure 3. Estimated Zebu ancestry of Brahmans and Tropical Composites, averaged per farm of origin and its standard deviations. X axis: farm and number of animals sampled.

Both analyses demonstrated that there is genetic variation within TC and BB, although this was more evident within TC than in BB. The large spread within breed seen for TC and BB in the principal components analysis is strongly suggestive of differences in breed composition or stratification, and it could be partially attributed to differences in estimated Zebu to Taurine ancestry ratio of each individual. Population stratification on BB animals of the CRC was expected given previous results (Fortes *et al*. 2011). Importantly, the origin of the animal did not fully explain the population stratification; as large variation was also seen within each origin. Hence, using farm of origin or breed designation as factors in genetic analysis of these populations does not correct for the differences seen within. Further analyses are required to better explore and understand potential stratification of this population, to correlate the principal components and hierarchical clustering results, and to evaluate whether including estimated breed ancestry in genome-wide analyses improves the reliability of such analyses.

The large variation observed within and between BB origins and the relative high proportion of Taurine content estimated for some BB farms are interesting findings. Is the animal selection within those farms selecting "Taurine alleles" or "chromosomal segments" that were introduced long ago during the grading up of BB in Australia? On the other hand, is the TC variation in Zebu ancestry due different breeding strategies or due to selection decisions made in response to finding that animals with more or less Zebu content thrive in a particular environment? These open questions should be targeted in future research.

ACKNOWLEDGEMNTS

The authors thank the Cooperative Research Centre for Beef Genetic Technologies (Beef CRC) and its Legacy database for access to the data. We thank W. Barendse, R. McCulloch, B. Harrison and R. Bunch for generating genotypes, and T.S. Sonstegard for access to Nelore genotypes. Financial support for genotyping TC bulls was provided separately by Meat and Livestock Australia (project code B.NBP.0723).

REFERENCES

Achilli A, Olivieri A, Pellecchia M, Uboldi C, Colli L, Al-Zahery N, Accetturo M, Pala M, Kashani BH, Perego UA, Battaglia V, Fornarino S, Kalamati J, Houshmand M, Negrini R, Semino O, Richards M, Macaulay V, Ferretti L, Bandelt HJ, Ajmone-Marsan P and Torroni A (2008) *Curr. Biol.* **18**:R157.

Alexander D.H., Novembre J. and Lange K. (2009) *Gen. Res.* **19**:1655.

Barwick S. A., Johnston D. J., Burrow H. M., Holroyd R. G., Fordyce G., Wolcott M. L., Sim W. D. and Sullivan M. T. (2009) *Anim. Prod. Sci.* **49**:367.

Browning B.L. and Browning S.R. (2009) *Am. J. Hum. Gen.* **84**:210.

Burns B. M., Corbet N. J., Corbet D. H., Crisp J. M., Venus B. K., Johnston D. J., Li Y., McGowan M. R. and Holroyd R. G. (2013) *Anim. Prod. Sci.* **53**:87.

Fortes M.R.S., Bolormaa S., Porto Neto L.R., Holroyd R.G. and Reverter A. (2011) *Proc. Assoc. Advmt. Anim. Breed. Genet.* **19**:267.

Harrison B.E., Bunch R.J., McCulloch R., Williams P., Sim W., Corbet N.J. and Barendse W. (2012) *Anim. Prod. Sci.* **52**:890

Ma L., Wiggans G.R., Wang S.W., Sonstegard T.S., Yang J., Crooker B.A., Cole J.B., Van Tassell C.P., Lawlor T.J. and Da Y. (2012) *BMC Geno.* **13**:536.

Price A.L., Patterson N.J., Plenge R.M., Weinblatt M.E., Shadick N.A. and Reich D. (2006) *Nat. Gen.* **38**: 904.

Purcell S., Neale B., Todd-Brown K., Thomas L., Ferreira M.A.R., Bender D., Maller J., Sklar P., de Bakker P.I.W., Daly M.J. and Sham P.C. (2007) *Am. J. Hum. Gen.* **81**:559.

Yang J.A., Lee S.H., Goddard M.E. and Visscher P.M. (2011) *Am. J. Hum. Gen.* **88**:76.

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:151-154

INTEGRATED ASSEMBLY OF POSITIVELY SELECTED GENES IN CATTLE

I.A.S. Randhawa, M.S. Khatkar, P.C. Thomson and H.W. Raadsma

ReproGen - Animal Bioscience Group, Faculty of Veterinary Science, University of Sydney, Camden, NSW 2570, Australia

SUMMARY

In this study, we assembled genome-wide scans for selective sweeps in various breeds of cattle and constructed an integrated genomic map of positively selected genes on UMD3.1 assembly. Available studies have explored a variety of genetic diversity in the form of microsatellites, SNP genotypes and DNA sequences on animals from world-wide populations of pure bred and crossbred cattle. These studies tested for departure from neutrality using various tests, mostly based on estimates of population allele differentiation and haplotype homozygosity. Definite genomic regions harbouring genes associated with simple traits (e.g. coat colour, polledness, muscle hypertrophy etc.) have been identified through signatures of selection. The genes identified under selection for the polygenic traits (e.g. adaptation, production, reproduction, feed efficiency, immunity, behaviour etc.) have also been supported by gene networks, QTLs and genome-wide association studies. These diverse investigations highlight the advantages and limitations of the available bovine genomic resources and different methodologies and have been reviewed here.

INTRODUCTION

Recent advances in various fields of genetic research have increased the availability of high throughput molecular biology tools and analytical approaches for investigating genetic diversity of farm animals. This has led us to an early understanding of the origin of species, domestication, genetic control of adaptation and imprints for selection for health and production traits (Andersson and Georges 2004; Lenstra *et al.* 2011). Modern domesticated species are a result of positive selection for the traits of economic and social importance for efficient and sustainable production in the past ~10,000 years (Mirkena *et al.* 2010). Largely due to the diverse panel of ~ 800 breeds and mixture of factors shaping their high genetic diversity, the cattle genome has been extensively investigated for signatures of selection (Barendse *et al.* 2009; Flori *et al.* 2009; Qanbari *et al.* 2010; Stella *et al.* 2010). Here we present a survey of positively selected genes for various traits identified by many tests and data sets and integrated them on the genomic positions of UMD3.1 bovine assembly (http://www.cbcb.umd.edu/research/bos_taurus_assembly).

MATERIALS AND METHODS

Available studies have explored a variety of genetic polymorphism data in the form of microsatellites, SNP genotypes (10K and 50K Illumina's BovineSNP chip assays) and DNA sequences composed of thousands of animals of multiple populations (pure breeds and crossbred). We have selected those studies which used whole-genome high-density panels of SNP genotypes for characterization of positive selection across several major cattle breeds (Table 1). The studies which have used microsatellites, DNA sequences or restricted genotyping datasets are almost twofold of genome-wide scans (data and references are not shown) and have not been included in the present study. The populations in these studies were investigated using various methods to estimate parameters in support of historical or ongoing sweeps of beneficial mutations. An integrated genomic map of positively selected genes from previous bovine assemblies (Btau3.1 and Btau4.0) was constructed by placing them – along with unique indicators for the references, selection tests and number of reporting studies – on UMD3.1 genomic positions.

Table 1: Summary of selected studies on genome-wide scans of selection signatures in cattle

Study	Test	Data (SNPs) and genome assembly	Breeds and (samples)	Genes (N)	Selective sweeps examined
Hayes *et al.* (2009)	iHS AFD	10K (9,323) Btau3.1	4 (774)	4	Milk production
Chan *et al.* (2010)	F_{ST} EHH	10K (9,919) Btau4.0	13 (317)	33	Tropical adaptation: Tick resistance, Heat resistance, Immune system
Barendse *et al.* (2009)	F_{ST} iHS CLR	10K (8,859) Btau4.0	21 (385)	2	Residual feed intake, Beef yield (intramuscular fatness)
TBHMC (2009)	F_{ST} iHS CLR	TBHMC (37,470) Btau3.1	19 (497)	20	Domestication, Behaviour Immunity, MHC, Feed efficiency, Double Muscling, Milk yield & composition, Intramuscular fat content
Stella *et al.* (2010)	CLL	TBHMC (32,689) Btau4.0	19 (497)	55	Polledness, Coat color (Black, Piebald), Dairy production, Reproduction
Gautier *et al.* (2009)	BF	50K (36,320) Btau4.0	11 (437)	42	Adaptation (pathogens & climate), Trypanosomiasis tolerance, Immune response Nervous system, Skin and hair properties
Flori *et al.* (2009)	F_{ST}	50K (42,486) Btau4.0	3 (2803)	48	Milk production, Reproduction, Body coloration
Qanbari *et al.* (2010)	EHH REHH	50K (41,398) Btau4.0	1 (810)	44	Milk yield and composition, Reproduction, Behaviour, Dairy quality
Qanbari *et al.* (2011)	F_{ST} iHS	50K (42,600) Btau4.0	12 (3876)	26	Reproduction (fertility), Muscle formation, Feed efficiency, Productive life
Gautier and Naves (2011)	iHS Rsb	50K (44,057) Btau4.0	22 (725)	11	Reproduction, Metabolism, Immunity

AFD: Allele Frequency Difference, $\boldsymbol{F_{ST}}$: Fixation index, **BF**: Bays Factor, **TBHMC**: The Bovine HapMap Consortium, **iHS**: Integrated Haplotype-homozygosity Score, **CLR**: Composite of Likelihood Ratios, **CLL**: Composite of Log Likelihood, **EHH**: Extended Haplotype Homozygosity, **REHH**: Relative EHH, **Rsb**: a measure of across population haplotype homozygosity using single locus EHH

RESULTS AND DISCUSSION

A total of 285 genes declared as candidates under selection were assembled, of which only 11 genes (9 twice and 2 thrice) were identified in multiple populations (Table 1). The integrated map contains 272 genes underlying 236 selection regions of the bovine genome (Figure 1). At least 26 selection regions identified by different studies were less than 1 Mb apart . This discrepancy may either be due to different versions of gene annotation or the nature of selection test capturing slightly different patterns of genetic diversity shaped by selection, or could be due to different genetic factors. Evidence of selection was based on the measures of population differentiation, the allele frequency spectrum, linkage disequilibrium (LD) and haplotype structures. The most common tests used to analyse genomic regions under selection were estimates of population differentiation (F_{ST}) and haplotype homozygosity (EHH and iHS).

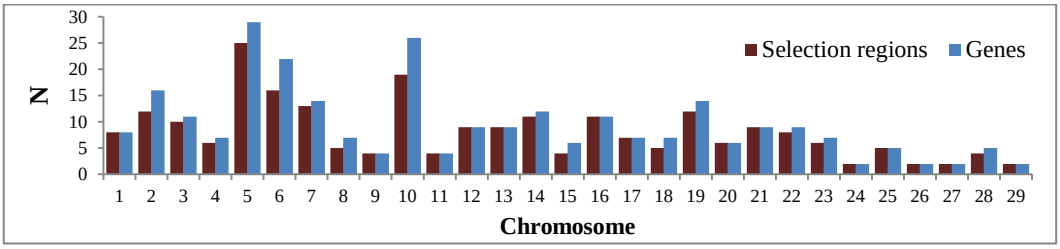


Figure 1: Distribution of chromosome-wise selection regions and genes in cattle genome.

Bovine chromosomes (BTA) 5, 6 and 10 have the highest number of identified selection regions and genes (Figure 1), whereas, BTA-2, 6 and 14 contain important candidate genes linked to various phenotypic traits in cattle (Figure 2). Cattle breeds undergoing directional or divergent selection for specific traits have shown a lack of concordance for genomic regions under selection when measured by different selection tests (Qanbari *et al.* 2011). Breed-wise sample composition, SNP panels and their density might have contributed to the differences in the results across studies (Barendse *et al.* 2009). Overall, poor concordance among studies and, selection tests within and across studies, especially in similar populations indicate the limitation of the available data sets and lack of power of selection tests. Signatures of selection harbouring genes associated with simple traits have been easily identified at the explicit genomic regions using outlier loci by applying simple genome-wide threshold strategies. For example, genes harbouring genetic mutations of major effect that control simple traits in cattle include; the polled gene on BTA-1 for absence of horns (Stella *et al.* 2010), MSTN on BTA-2 for double muscles (TBHMC 2009) and MC1R on BTA-18 for coat colour (Flori *et al.* 2009; Stella *et al.* 2010).

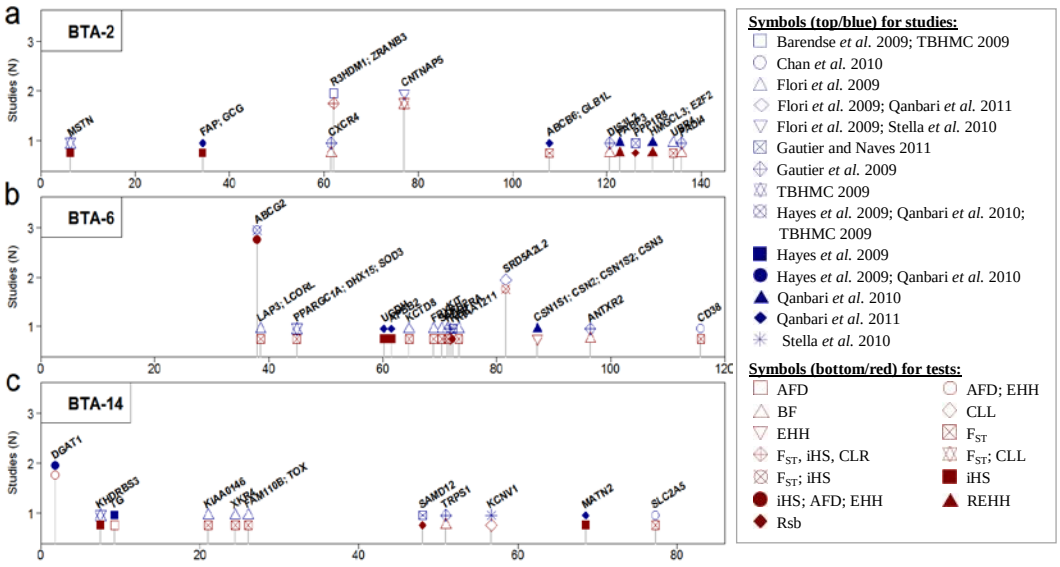


Figure 2: Genes underlying selection regions identified in various number (N) of studies by particular selection test(s) on a) BTA-2, b) BTA-6 and c) BTA-14.

The interpretation of selection signatures for complex traits is constrained by many factors, such as; limited availability of phenotypic records, variable selection pressure on polygenic alleles and inability of tests to capture selection by using conventional outlier loci approaches. The genes underlying the regions under selection for the polygenic traits have been generally linked to the phenotypic diversity in each study (see Table 1, e.g. adaptation, milk production, feed efficiency, reproduction, immune response, behaviour etc.) and in a few instances have also been supported by gene networks, QTL studies and genome-wide association studies.

Overall, the survey of genome-wide scans of selection illustrates several successful discoveries by using within and across population data sets of variable marker density. On the other hand, the disadvantages of previously available low-resolution and incomplete bovine genome maps might have provided restrictive insights. Hence, remapping previous results to the recently annotated UMD3.1 assembly and careful inspection along with new neighbouring genes can be useful. Meta-analysis of combined data from these studies can further improve the power for such analysis. Relative performance of several selection tests, as described above, has also shown differences in their power to localize a range of selection signals at varying magnitudes. A combination of multiple selection tests (Grossman *et al.* 2010; Randhawa *et al.* 2013) can be a robust approach to localize and fine-map selection regions, and link underlying genetic variation with phenotypic diversity. Moreover, the strength of signatures of selection can be improved by combining data sets and animals from multiple breeds which are phenotypically alike for the target traits.

REFERENCES

Andersson L. and Georges M. (2004) *Nature Reviews Genetics.* **5**:202.

Barendse W., Harrison B.E., Bunch R.J., Thomas M.B. and Turner L.B. (2009) *BMC Genomics.* **10**:178.

Chan E.K.F., Nagaraj S.H. and Reverter A. (2010) *Anim. Genet.* **41**:467.

Flori L., Fritz S., Jaffrézic F., Boussaha M., Gut I., Heath S., Foulley J.-L. and Gautier M. (2009) *PLoS ONE.* **4**:e6595.

Gautier M., Flori L., Riebler A., Jaffrezic F., Laloe D., Gut I., Moazami-Goudarzi K. and Foulley J.-L. (2009) *BMC Genomics.* **10**:550.

Gautier M. and Naves M. (2011) *Mol. Ecol.* **20**:3128.

Grossman S.R., Shylakhter I., Karlsson E.K., Byrne E.H., Morales S., Frieden G., Hostetter E., Angelino E., Garber M., Zuk O., Lander E.S., Schaffner S.F. and Sabeti P.C. (2010) *Science.* **327**:883.

Hayes B.J., Chamberlain A.J., Maceachern S., Savin K., McPartlan H., MacLeod I., Sethuraman L. and Goddard M.E. (2009) *Anim. Genet.* **40**:176.

Lenstra J.A., Groeneveld L.F., Eding H., Kantanen J., Williams J.L., Taberlet P., Nicolazzi E.L., Sölkner J., Simianer H., Ciani E., Garcia J.F., Bruford M.W., Ajmone-Marsan P. and Weigend S. (2011) *Anim. Genet.* **43**:483.

Mirkena T., Duguma G., Haile A., Tibbo M., Okeyo A.M., Wurzinger M. and Solkner J. (2010) *Livestock Science.* **132**:1.

Qanbari S., Gianola D., Hayes B., Schenkel F., Miller S., Moore S., Thaller G. and Simianer H. (2011) *BMC Genomics.* **12**:318.

Qanbari S., Pimentel E.C.G., Tetens J., Thaller G., Lichtner P., Sharifi A.R. and Simianer H. (2010) *Anim. Genet.* **41**:377.

Randhawa I.A.S., Thomson P.C., Khatkar M.S. and Raadsma H.W. (2013) *Proc. Assoc. Advmt. Anim. Breed. Genet.* **20**:(in press).

Stella A., Ajmone-Marsan P., Lazzari B. and Boettcher P. (2010) *Genetics.* **185**:1451.

The Bovine HapMap Consortium. (2009) *Science.* **324**:528.

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:155-158

GENETIC PARAMETERS FOR FAECAL WORM EGG COUNT AND OBJECTIVELY MEASURED WOOL TRAITS IN SOUTH AFRICAN MERINOS

P.A.M. Matebesi-Ranthimo[1,2], S.W.P. Cloete[3,4], J.B. van Wyk[1] and J.J. Olivier[4]

[1]National University of Lesotho, P.O. Roma 180, Roma, Lesotho
[2]University of the Free State, P.O. Box 339, Bloemfontein 9300, South Africa
[3]University of Stellenbosch, Private Bag X1, Matieland, 7602, South Africa
[4]Institute for Animal Production: Elsenburg, Private Bag X1, Elsenburg, 7609, South Africa

SUMMARY

Genetic parameter estimates for faecal worm egg count (FEC) and objectively measured wool traits were assessed, using data of Merino sheep from a selection experiment (four lines - a line selected for clean fleece weight, a fine wool line, an unselected control line and a line selected against rearing failure) maintained at the Tygerhoek research farm. Data consisted of between 3842 and 6822 (depending on the trait) records of animals born between 1989 and 2010. Rectal faecal samples were taken from individual sheep at 13 to 16 months of age after drenching was withheld for at least 10 weeks, generally in July to September. Nematode eggs were counted using McMaster technique, with a sensitivity of 100 eggs per gram of wet faeces. The heritability of FEC amounted to 0.16 after the data (with 100 added to account for zero counts) were transformed to logarithms to the power of 10. The genetic relationships of FEC with wool traits were favourable. Selection for a reduced FEC is unlikely to result in unfavourable correlated responses to wool traits in South African Merinos, in fact staple strength and the coefficient of variation of fibre diameter will benefit from it.

INTRODUCTION

Internal parasite infestations cost South African sheep producers hundreds of millions of rand each year arising from treatment costs, increased level of management and vigilance, a loss of production and even mortality in severe cases (Nieuwoudt *et al.* 2002). Resistance of gastro-intestinal parasites to anthelmintics has become more prevalent over the recent years to the stage that it has been described as rampant (Bath and Van Wyk 2009). In addition, consumers increasingly demand animal products that are free from contamination of chemicals (Khusro *et al.* 2004). Integrated parasite control measures (Nieuwoudt *et al.* 2002) may contribute to reduce the parasite burden in a variety of small ruminants. Several international authors reviewed genetic parameters for resistance to nematodes in sheep (Safari *et al.* 2005; Morris 2011). However, previous studies in South Africa is limited to those of Nieuwoudt *et al.* (2002), Cloete *et al.* (2007), as well as the work on the FAMACHA® system by Riley and Van Wyk (2009), The evidence of successful Australian (Woolaston and Piper 1996; Greeff *et al.* 2006) and New Zealand (Morris *et al.* 2005) breeding programs for resistance and resilience (Morris *et al.* 2010) to nematode infestation may also confer economic advantages in South African sheep. The objectives of this study were to estimate the (co)variance components and ratios as well as the genetic, phenotypic and environmental correlations between FEC and objectively measured wool traits in South African Merinos.

MATERIAL AND METHODS

Performance records were obtained from four lines (a line selected for clean fleece weight, a fine wool line, an unselected control line and a line selected against rearing failure) of Merino sheep maintained on the Tygerhoek experimental farm of the Western Cape Department of Agriculture, near Riviersonderend in the Western Cape Province of South Africa. Progeny born

between 1989 and 2010 were used, with a pedigree from 1969 to 2010. The data included between 3842 and 6822 (depending on the trait; Table 1) records, the progeny of 554 sires and 2483 dams. The origin and initial selection in the flock were described by Heydenrych (1975). There was no selection for a reduction of FEC in any of the lines, and line effects were thus not considered. Flock maintenance, husbandry, experimental design and sampling procedures for FEC are described by Cloete *et al*. (2007). Rectal faeces samples were obtained under natural challenge from 1995 to 2010 (with the exception of 2004 when FEC data were not collected) from individual animals between the ages of 13 to 16 months. This was conducted after drenching was withheld for at least 10 weeks, generally in July to September. The pathogens present during this time of the year at Tygerhoek farm are *Teladorsagia*, *Nematodirus* and *Trichostrongylus* spp (Reinecke, 1994). Individual faecal samples were assessed for FEC using the McMaster technique, with a sensitivity of 100 eggs per gram of wet faeces (Cloete *et al*. 2007). Mean (±s.d.) untransformed FEC amounted to 694±1232, clearly showing a non-normal distribution. Transformation to logarithms, as described below, resulted in a data set with a normal distribution for the analysis of FEC.

Traits included in the analyses were thus the logarithm to the power of 10 of FEC (after 100 were added to account for zero counts; hereafter referred to as FEC), clean fleece weight (CFW), clean scoured yield percentage, (CY), fibre diameter (FD), staple length (SL), staple strength (SS) and coefficient of variation of FD (CVFD). Greasy fleece weight (GFW) was recorded at shearing in August-September each year after a wool growth period of approximately one year. Information on GFW was combined with CY data to derive CFW. Measures of wool quality were determined on a midrib wool sample taken from each animal at 14-16-months of age. It is conceded that heterogeneous variances between years may affect the outcome of the analyses on FEC, but these effects were adequately dealt with by the transformation.

Data of animals with information on pedigree, sex (male or female) dam age (2-6 years) and birth status (single or multiple) were included The statistical analysis was conducted using ASREML (Gilmour *et al*. 2009). The significance of fixed effects of sex, year of birth, birth type, selection line, dam age in years and sex*birth year interaction was tested leaving only significant effects in the final model. The best random effects models involving direct and maternal genetic variances, the correlation between direct and maternal effects, as well as maternal permanent environmental variances were tested for significance with ASREML, using log likelihood ratios derived from single-trait analysis on all traits. Variance component and heritability estimates were derived from single-trait animal models. The correlations were estimated by fitting a series of two-trait models, as it was impossible to include all traits in a single multi-trait model. Parameters stemming from the different models were all within a 0.02 range, and only two-trait analyses were reported.

RESULTS AND DISCUSSION

The fixed effects of birth type (single/multiple, $P<0.05$), sex (male/female), year of birth (1989-2003, 2005-2010) and the sex*birth year interaction had a significant ($P<0.01$) effect on FEC. These results are consistent with those reported by Cloete *et al*. (2007) on the same Merino resource flock. A similar set of fixed effects for FEC, with the addition of selection line (1-4) and age of dam ($2\text{-}7^{+}$) significantly affected all objectively measured wool traits, and were included in the models used for subsequent analyses. Models with only the direct additive effect fitted the data best for FEC, CY, SL, SS and CVFD. Maternal effects in addition to direct additive effect were present in the FD analysis, while the covariance between animal effects as well as dam permanent environmental effect contributed to the variation of CFW. FEC was heritable in this investigation at 0.16 (Table 1). The h^2_a in this investigation for FEC is consistent with a value of 0.18 reported by Cloete *et al*. (2007) on the same Merinos, slightly lower than those of 0.19 to 0.23 reported by

Riley and Van Wyk (2009) and lower than the average value of 0.27 derived from a mixture of data sets from naturally and artificially challenged flocks in Australia (reviewed by Safari *et al*. 2005). Estimates of h^2_a for objective wool traits ranged from 0.21 for SS to 0.68 for CVFD. Estimates of maternal heritability amounted to 0.08 for CFW and to 0.04 for FD. The dam permanent environmental effect accounted for 0.04±0.02 of the total phenotypic variance for CFW. The correlation between direct and maternal effects was high and negative at -0.58±0.08 for CFW

Table 1 Descriptive statistics for the data from the Tygerhoek Merino resource flock, also with appropriate direct and maternal heritability estimates

Trait	n	Mean	SD	CV (%)	h^2 ± s.e	h^2_m ± s.e.
FEC	5473	2.59	0.51	19.69	0.16±0.02	-
Objective wool traits						
Clean fleece weight (kg)	6717	3.55	0.92	25.92	0.48±0.04	0.08±0.03
Clean yield (%)	6717	71.97	4.60	6.39	0.65±0.02	-
Staple length (mm)	6548	90.26	14.97	16.58	0.40±0.03	-
Staple strength (N/ktex)	3842	34.60	12.32	35.61	0.21±0.03	-
Fibre diameter (µm)	6822	19.33	1.99	10.29	0.66±0.04	0.04±0.01
CV of fibre diameter (%)	5683	19.62	2.84	14.48	0.68±0.03	-

n= number of records, SD=standard deviation and CV= coefficient of variation, FEC= log to the power of 10 transformed (FEC + 100), h^2_a= direct heritability, h^2_m= maternal heritability and se = standard error

Most of the correlations between FEC and objective wool traits were not significant (Table 2). The genetic correlation of FEC with CFW was unfavourable but did not reach a level of double the corresponding standard error. Animal model analyses by Pollott and Greef (2004) yielded an estimate of 0.13, which was in the same order of magnitude. The genetic relationships for FEC with CY, SL and FD were low to negligible. Pollott and Greeff (2004) accordingly reported genetic correlations of FEC that ranged from -0.02 to -0.05 for SL and from -0.03 to -0.08 with FD. Only staple strength and coefficient of variation of fibre diameter were significantly and favourably related to FEC on the genetic level. Similar findings were obtained previously on the current Merino resource flock (Cloete *et al*. 2007). The current genetic correlation of 0.33 between FEC and CVFD is higher than the values ranging from 0.06 to 0.13 reported for Australian Merinos (Greeff and Karlsson 1998). The negative, favourable genetic correlation of FEC with SS was comparatively high at -0.54 in the current study. Studies done on Australian Merinos yielded lower genetic correlation estimates, ranging from -0.05 to -0.17, depending on the model of analysis (Pollott and Greeff 2004). Cloete *et al*. (2007) also reported a similar genetic correlation estimate of -0.49 using a smaller data set of the current Merino flock. However, these results suggested that SS may be improved when animals are selected for a reduced FEC. Phenotypic and environmental correlations between wool traits and FEC were generally low and variable in sign.

CONCLUSIONS

The present study suggested that FEC is variable and heritable in South African Merinos, and that selection should result in additive gains in the ability of animals to resist natural nematode infestation. Selection for a reduced FEC is unlikely to result in unfavourable correlated responses in objectively measured wool traits in South African Merinos as suggested by low or negligible genetic relationships of most wool traits with FEC. The exceptions were favourable genetic correlations of FEC with SS and CVFD. Selection for resistance to nematode infestation may add to profitable sheep production in areas of Southern Africa with high levels of nematode challenge.

Table 2 Correlations (±s.e.) between log to the power of 10 transformed faecal worm egg count (FEC) and objectively measured wool traits in the Tygerhoek Merino flock

Trait	Genetic (r_g)	Environment (r_e)	Phenotypic(r_p)
Clean fleece weight	0.16±0.10	0.04±0.02	*0.06±0.02
Clean yield	0.05±0.07	-0.04±0.03	0.01±0.02
Staple length	-0.01±0.09	-0.01±0.03	-0.01±0.02
Staple strength	*-0.54±0.10	*0.08±0.03	*-0.04±0.02
Fibre diameter	-0.09±0.08	-0.00±0.03	-0.03±0.02
Coefficient of variation of fibre diameter	*0.33±0.07	*-0.06±0.03	*0.06±0.02

*** - significant (P<0.05) correlation**

`

REFERENCES

Bath G.F. and Van Wyk, J.A. (2009) *Small Rumin. Res.* **86**: 6.

Cloete S.W.P., Olivier J.J., Du Toit E. and Dreyer F.H. (2007) *S. Afr. J. Anim. Sci.* **37:** 237.

Gilmour A.R., Gogel B.J., Cullis B.R., Welham S.J. and Thompson R. (2009) 'ASREML-User Guide Release' 3.0 VSN International Ltd, Hemel Hempstead, HPI IES, UK.

Greeff J.C. and Karlsson L.J.E. (1998) *Proc. 6th World Congr. Gen. Appl. Livest. Prod*. **24:** 63.

Greeff J.C., Karlsson L.J.E. and Underwood N. (2006) *Proc. 8th World Congr. Gen. Appl. Livest. Prod.* 15-08.

Heydenrych H.J. (1975) PhD dissertation, University of Stellenbosch, Stellenbosch, South Africa.

Khusro M., Van der Werf J.H.J., Brown D.J. and Ball A. (2004) *Livest. Prod. Sci***. 91:** 35.

Morris C.A. (2011) *Proc. Assoc. Advmt. Anim. Breed. Genet.* **18:** 263.

Morris C.A., Wheeler M. and Shaw R.J. (2010) *N. Z. Soc. Sci*. **70:** 229.

Morris C.A., Wheeler M., Watson T.G., Hosking B.C. and Leathwick D.M. (2005) *N. Z. J. Agric. Res*. **48:** 1.

Nieuwoudt S.W., Theron H.E. and Krüger L.P. (2002) *J. S. Afri. Med. Assoc*. **73:** 4.

Pollott G.E. and Greeff J.C. (2004) *J. Anim Sci*. **82**: 2840.

Reinecke R.K. (1994) *Vet. Par.* **54**: 49.

Riley D.G.and Van Wyk J.A. (2009) *Vet. Par.* **164**: 44.

Safari E., Fogarty N.M. and Gilmour A.R. (2005) *Livest. Prod. Sci.* **92:** 271.

Woolaston R.R. and Piper L.R. (1996) *Anim. Sci*. **62:** 451.

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:159-162

IMPROVING MILK PRODUCTION AND LACTATION PERSISTENCY OF PHILIPPINE DAIRY BUFFALOES USING RANDOM REGRESSIONS

E.B. Flores and J.H.J. van der Werf

School of Environmental and Rural Science, University of New England Armidale NSW 2351

SUMMARY

Heritabilities and genetic correlations for milk production and lactation persistency were estimated from first parity test day records of 1,022 Philippine dairy buffalo cows using a random regression model. Varying orders of Legendre polynomials were combined with the Wilmink's function and were used in random regression models. Variance components for milk yield and various measures of lactation persistency were derived.

Heritabilities estimated by random regression for milk test day yields were moderate, ranging from 0.17 – 0.19 with a model that fitted a Wilmink's function for the random additive genetic and permanent environment effects. Two eigenvalues derived from the genetic covariance matrix explain 99% of the variation. The first eigenfunction was positive and constant while the second was negative at the beginning but increased and became positive halfway into the lactation. Selection emphasis on the second canonical variate can improve persistency. Optimal selection for increased milk yield and lactation persistency could be explored using the parameter estimates from a random regression model.

INTRODUCTION

Genetic evaluations for dairy cattle have shifted to the use of test day records directly rather than a single 305D lactation measure as test day yields can be adjusted for specific test day effects more accurately (Bilal and Khan 2009) and there is no need to adjust or standardize lactation yields to 305D. Schaeffer and Dekkers (1994) introduced a random regression test day model which involves the regression of merit on days in lactation to account for variation between cows in their performance across the lactation trajectory. This allows an individual cows' lactation curve to deviate from the average, making it possible to select for lactation persistency (Jamrozik *et al.* 1997). Functions frequently used in various studies to describe the shape of the lactation curve include among others, Woods's model, Legendre polynomials (Guo *et al.* 2002) and Wilmink's function (Schaeffer *et al*. 2000). Random regression models can also use higher order polynomial functions but these often have "end-of-range" problems resulting in erratic and extreme estimates of variance and genetic parameters (Meyer 2005).

Lactation persistency is defined as the rate of decline after peak lactation yield has been reached. With random regression models, estimated breeding values (EBV) can be calculated for any day within the lactation period. EBVs for lactation yields in the later part of lactation can be given more weight in selection thus; the shape of the lactation curve and persistency can be improved. A more persistent cow can be more profitable and may have less health and reproductive problems. Persistency could be especially useful in buffaloes that often suffer from too short lactations as well as negative energy balance in early lactation. Different measures of persistency have been proposed utilizing EBVs for daily yields or partial yields and these can be predicted from additive genetic effects estimated by the random regression test day model.

Information regarding the use of random regression models in dairy buffaloes is limited. Sesana *et al*. (2010) estimated genetic parameters for buffalo milk test days by random regression using Legendre polynomials and reported high genetic variance estimates at the beginning of lactation and negative genetic correlations between test days in early and mid to late lactations. The latter could be an indication of "end-of-range" problem which could be avoided with the use

of Wilmink's function. The objective of this study was to compare various random regression models for estimating genetic parameters for milk production traits in Philippine dairy buffaloes in terms of goodness of fit measures, genetic variance, genetic correlations between test days and heritabilities at different days in milk, and derive from such models breeding values for yield and persistency measure.

MATERIALS AND METHODS

Seven thousand eight hundred twenty five (7,825) test day records of 1,022 first parity Philippine dairy buffalo cows of 9 herds from 1997 to 2012 were used directly in a random regression (RR) model to estimate heritability at different days in milk (DIM) in a given lactation. The average test day milk, fat and protein yields as well as fat and protein concentration were 4.6 kg.±2.0, 0.34kg.±0.14, 0.20kg.±0.08, 7.22%±1.63 and 4.31%±0.61, respectively. The RR model is given as: $y_{ijkl} = HTD_i + \sum_{m=0}^{n} \beta_{km} z_{klm} + \sum_{m=0}^{n} \alpha_{jm} z_{jlm} + \sum_{m=0}^{n} pe_{jm} z_{jlm} + e_{ijkl}$ where y_{ijkl} is the test day record l of cow j made on DIM_{jl} of lactation; $\boldsymbol{HTD_i}$ is the fixed effect of herd-test date i; e_{ijkl} is random residual effect; $ß_{km}$, α_{jm}, and pe_{jm} are regression coefficients on days in milk (DIM) within sub-class $\boldsymbol{k}$ age-season of calving, random additive genetic and permanent environment effects of $\boldsymbol{m^{th}}$ order on days in milk, respectively. The Wilmink's function (Wil) and Legendre polynomial ($Leg_{\boldsymbol{m}}$) of varying orders describe the shape of the lactation curve. For Wilmink's function, let $Z_{jl0} = 1, Z_{jl1} = DIM, Z_{jl2} = exp^{-0.05DIM}$ whereas for Legendre polynomial, let $Z_{jl0} = 0.7071, Z_{jl1} = 1.2247 * DIM, Z_{jl2} = 2.3717 * DIM^2 - 0.7906, Z_{jl3} = 4.6771 * DIM^3 - 2.8062 * DIM$ The order of the RR functions can vary between components and the days in milk (DIM5 – DIM329) were standardized from -1 to 1 for all Legendre polynomial functions. Residual variances were allowed to vary for each of the ten TD periods in a lactation but residual covariance between TD periods were assumed to be zero. Various combinations of Wilmink's function and $Leg_{\boldsymbol{m}}$ of varying *(m)* order of fit were used for the fixed and random regression coefficient estimation. For all models, the F_1/F_2 format describes the combination of functions for $\boldsymbol{\alpha}$ (F_1) and $\boldsymbol{pe}$ (F_2) effect respectively. Average Information Residual Maximum Likelihood (ASREML) software (Gilmour *et al*. 2009) was used for variance component estimation. Random $\boldsymbol{\alpha}$ and $\boldsymbol{pe}$ regression coefficients were used to build the covariance matrix for different days in milk along the lactation period (Jamrozik *et al*. 1997).

Heritabilities for a particular DIM $\boldsymbol{i}$ in lactation were calculated by dividing the estimated genetic variance by the sum of genetic, permanent environment and appropriate residual variances for that particular DIM. Different models were compared based on heritability, log likelihood, Akaike's Information Criterion (AIC) and Schwarz' Bayesian Information Criterion (BIC). The lower value for both AIC and BIC indicates a better fitting model. Eigenfunctions related to eigenvalues of the genetic covariance matrix were estimated based on the method of Kirkpatrick *et al*. (1990) to analyze the pattern of variation across the trajectory and from this, to infer the variation in persistency. Transformation of the RR model to canonical scale was done according to the method of van der Werf *et al.* (1998). Response to selection from varying weights applied to canonical variates Z_1 and Z_2 was determined and plotted across the lactation period.

RESULTS AND DISCUSSION

Goodness of fit values of various RR models is shown in Table 1. Generally, given the same function for the random effects, models with Wil function (e.g. Leg_2/Leg_2, Leg_1/Leg_3, Wil/Wil) in fixed regression have better goodness of fit values compared to models with $Leg_{\boldsymbol{m}}$ (e.g. M3, M5, M6) functions. Top models based on AIC and BIC values were those with more than 12 random parameters. But the top models either have relatively high genetic variance in early lactation (Wil/Leg_3) or low variance in mid-lactation (Leg_2/Leg_3) except for Leg_1/Leg_3 (Fig. 1). The high

genetic variance in early lactation by Wil/Leg_3 resulted in relatively high heritabilities (Fig.2) that might not be realistic whereas the low genetic variance in mid-lactation by Leg_2/Leg_3 resulted in low estimates of heritability at that period. Heritability estimates by Wil/Wil and Leg_1/Leg_3 models were closer to those of the repeated measures TD model at 0.15 as reported by Flores *et al.* (2013). The Leg_1/Leg_3 model has slightly lower estimates of genetic variance and heritability in early lactation compared with the Wil/Wil model (Fig. 1 & 2). This might be an indication of inadequate fit to the random additive genetic effect for models with Leg_1 functions at that period.

Table 1. Measures of goodness of fit for various random regression models applied to first parity milk yield test day records of Philippine dairy buffalo cows

Model	Regression function			No. of random parameters	Log Likelihood	AIC	BIC	Rank
	α	***pe***	Fixed effect					
M3	Leg_2	Leg_2	Leg_2	12	-6327.3	12679	12701	9
M5	Leg_1	Leg_3	Leg_3	13	-6196.5	12419	12443	8
M6	Wil	Wil	Leg_3	12	-6145.3	12315	12337	7
Leg_1 / Wil	Leg_1	Wil	Wil	9	-6027.8	12074	12090	5
Leg_2 / Leg_2	Leg_2	Leg_2	Wil	12	-6027.1	12078	12100	6
Leg_2 / Leg_3	Leg_2	Leg_3	Wil	16	-6006.6	12045	12074	3
Leg_1 / Leg_3	Leg_1	Leg_3	Wil	13	-6000.8	12028	12051	2
Wil / Wil	Wil	Wil	Wil	12	-6023.7	12072	12093	4
Wil / Leg_3	Wil	Leg_3	Wil	16	-5994.2	12020	12050	1

*Leg_1 – first order Legendre polynomial; Leg_2 – 2nd order Legendre polynomial; Leg_3 – 3rd order Legendre polynomial; Wilmink – Wilmink's function. For all models described, the regression function used were always in the order **α** / **pe** effects. AIC – Akaike's information criterion; BIC – Bayesian information criterion*

Figure 1. Additive genetic variance for milk yield trait estimated from first parity test day records by random regression.

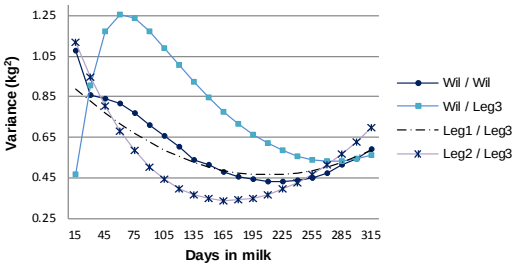
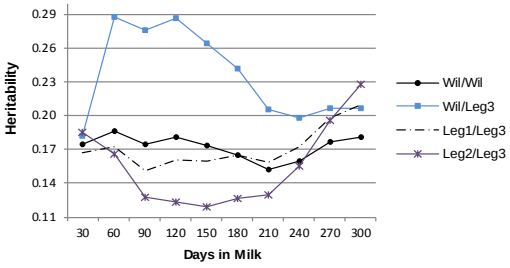


Figure 2. Estimates of heritability for milk yield trait estimated from first parity test day records by random regression.



***Wil -** Wilmink's function; **Leg1 - first order Legendre polynomial; Leg2** - 2nd order Legendre polynomial; **Leg3** - 3rd order Legendre polynomial. For models described, the regression function used were alsways in the order **α** / **pe** effect*

Genetic correlation between DIM5 and DIM_i showed positive but decreasing values as distance between days increased (Table 2). This is a pattern similar to dairy cattle but, correlations among adjacent test days were considerably lower for Wil/Leg_3 while they were unexpectedly high for Leg_1/Leg_3 and Leg_2/Leg_3 models. The Wil/Wil model had more realistic estimates for genetic correlation between test day periods. Overall, when considering formal test statistics, estimates of genetic variance and genetic correlations between test days, as well as model parsimony point of view, we conclude that the Wil/Wil is the preferred model.

Principal component analysis was done of the genetic covariance matrix from fitting the Wil/Wil model. The first and second principal components with eigenvalues EV_1 and EV_2 (Figure 3) were statistically significant (chi-square test, P<.0001) and explained 78% and 21% of the total genetic variance, respectively. The eigenfunction related to EV_1 was positive and constant

throughout the lactation. The result suggests that most of the variation in the test day milk yield is explained by a genetic component acting constantly throughout the lactation period. The eigenfunction related to EV_2 was negative in the first half of lactation but increased to positive values after DIM160. This eigenfunction may correspond to a genetic component for persistency (van der Werf *et al*. 1998) indicating it may be possible to select for lactation persistency.

Table 2. Genetic correlation between DIM5 and other days in milk estimated by different RR models

Model	Days in Milk									
	30	60	90	120	150	180	210	240	270	300
Wil / Wil	0.70	0.47	0.44	0.45	0.47	0.49	0.49	0.48	0.45	0.37
Wil / Leg_3	0.32	0.18	0.16	0.16	0.16	0.17	0.17	0.16	0.16	0.13
Leg_1 / Leg_3	1.00	0.99	0.97	0.93	0.88	0.70	0.70	0.59	0.47	0.25
Leg_2 / Leg_3	1.00	0.99	0.96	0.91	0.83	0.63	0.63	0.56	0.52	0.52

Figure 3. Eigenfunctions related to the three eigenvalues of the genetic covariance matrix from fitting Wil/Wil RR model to milk test day records of first parity Philippine dairy buffalo cows.

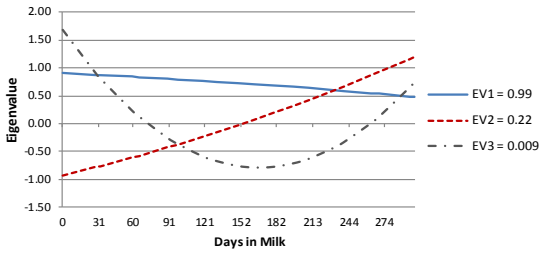
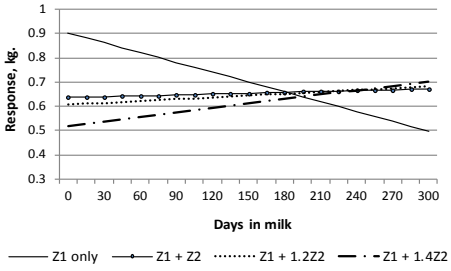


Figure 4. Response to selection on canonical variates Z_1 and Z_2 on milk yield trait



The transformation of the RR model to canonical scale with only the EV_1 and EV_2, enables selection on canonical varieties Z_1 and Z_2. Selection on Z_1 only will result to increase in milk yield mostly in the first trimester of lactation (Fig. 4). Equal weights applied to Z_1 and Z_2 will produce an even response across lactation but with a decrease in total milk yield. Increasing the weight applied to Z_2 further will increase milk production in the 3rd trimester of lactation. More emphasis on Z_2 results in a lower increase in total milk production but the increase rely less on a higher peak yield in first trimester of lactation. This may decrease stress to cows in this period. The relative economic weights of persistency and milk yield need to be known to optimally select for these traits simultaneously and genetic parameters from the RR model can be used for that purpose.

REFERENCES

Bilal G., and Khan M. S. (2009). *Pakistan Vet. J.* **29**: 34.
Flores E.B., Kinghorn B.P. and van der Werf J.H.J. (2013). *Proc. 10th World Buffalo Congress*
Gilmour A. R., Gogel B. J., Cullis B. R., and Thompson R. (2009). *ASReml User Guide 3.0*.
Guo Z., Lund M., Madsen P., Korsgaard I., and Jensen J. (2002). *J Dairy Sci.* **85:** 1596.
Jamrozik J., Schaeffer L. R., and Dekkers J.C.M. (1997). *J Dairy Sci.* **80:**1217.
Kirkpatrick M., Lofsvold D., and Bulmer M. (1990). *Genetics.* **124:** 979.
Meyer K. (2005). *Genet. Sel. Evol.* **37:** 473.
Schaeffer L. R., and Dekkers J. C. M. (1994). *Proc. 5th World Congr. Genet. Appl. Livest. Prod.*
Schaeffer L.R., Jamrozik J., and Kistemaker G. (2000). *J Dairy Sci.* **83:**1135.
Sesana R.C., Bignardi A.B., Borquis R.R., El Faro L., Baldi F., Albuquerque L.G., and Tonhati H. (2010). *J Anim Breed Genet*, **127**: 369.
van der Werf J. H. J., Goddard M. E., and Meyer K. (1998). *J Dairy Sci.* **81:** 3300.

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:163-166

THE PREDICTION OF GENETIC STRUCTURE OF EAST AFRICAN SMALLHOLDER DAIRY CATTLE

W.M.S.P. Weerasinghe[1], C. Gondro[1], M.G. Jeyaruban[2], O. Mwai[3], D. F. Mujibi[3] and J.P. Gibson[1]

[1]Centre for Genetic Analysis and Applications, University of New England, Armidale, NSW, 2351
[2]Animal Genetics and Breeding Unit, University of New England, Armidale, NSW, 2351
[3]International Livestock Research Institute, P.O. Box 30709, Nairobi, Kenya

SUMMARY

Identification of genetic structure and estimation of individual breed proportions in livestock species based on molecular data have become important tools in improvement of breeding programs in the developed world. In this study, we have applied high density SNP assays to understand genetic structure and breed composition of a developing world smallholder dairy system, which does not have pedigree records. Approximately 2000 East African smallholder crossbred dairy cattle were genotyped for 770k SNP. Principal component analysis (PCA) was used to identify the genetic structure and the ADMIXTURE program was used to estimate the proportional breed composition of individual animals. Genomic prediction and farmer prediction for breed proportions were compared. PCA revealed that the range of breed composition of small holder dairy cattle is much larger than commonly believed. The correlation between breed proportion estimated using ADMIXTURE and the farmer's assessment of breed proportion was only 0.4, revealing that in this system farmers do not have a good understanding of the breed composition of their animals. This will be a problem to be overcome if farmers are to make optimum breeding choices to produce replacement heifers.

INTRODUCTION

Dairy cattle play a major role in the economic life of over 5 million East African farmers. Starting over 50 years ago, smallholder dairy farmers have used a wide range of crossbred cattle, primarily crosses between indigenous breeds and European dairy breeds. But very few smallholder dairy farmers keep pedigree records and so the breed composition of individual animals is rarely known with certainty. High density single nucleotide polymorphism (SNP) assays have been widely used to map the genetic structure of admixed populations and can be used to estimate the breed composition of individual crossbred animals.

The indigenous population is very diverse and represents 77% of the total cattle population in East Africa (Rege *et al.* 2001). Moreover, through the extensive use of cross-breeding the purity of the indigenous population is believed to be compromised in many areas. To establish optimum breeding programs for genetic improvement and genetic conservation it is important to identify the genetic structure of existing populations and know the breed structure of breeding animals. The BovineHD Beadchip array (Illumina Inc.) includes SNP selected as informative in both *Bos taurus* and *Bos indicus* breeds and is suitable for use to determine population structure and breed composition in populations that include admixtures of *Bos taurus* and *Bos indicus* breeds. The aim of this study is to identify the genetic structure of smallholder crossbred dairy cattle in East Africa and to predict the breed proportions of individual crossbred animals, which do not have detailed pedigree records.

MATERIALS AND METHODS

Genotypic data for 2,051 indigenous (n=118) and cross bred (n=1,933) dairy cows were obtained as part of the Dairy Genetics East Africa project. The project includes data collection on

smallholder dairy cattle in Kenya and Uganda. Genotype data resulted from the BovineHD Beadchip array (Illumina Inc). Of the 777,962 SNP on the array, 566,056 were selected after quality control process (Gondro and Gibson, 2012). Since the crossbred dairy cows are studied, Y chromosomes were excluded. Only the SNP on the 29 autosomes were retained for analysis. The genotypes for the three main indigenous breeds Ankole (n=44), Nganda (n=16) and Small East African zebu (SEAZ) (n=58) were also included in the study. Since pedigree data is not available, 5 breeds from the international Bovine Hapmap 770k data set were selected as reference breeds: Nelore (reference *Bos indicus* breed; n=35), Brown Swiss (n=24), Holstein (n=66), Jersey (n=40) and N'Dama (reference African *Bos taurus* breed; n=24) and all the reference animals were used for all the analysis. Although Friesian and Ayrshire have been reportedly widely used in the East Africa region, these breeds were not available in the Hapmap data set.

A principal components analysis (PCA) was performed to help describe the genetic structure of our East African cattle sample. The analyses included the Hapmap animals to provide reference points for interpretation of the East Africa samples. PCA was based on the genomic relationship matrix, obtained using the allele frequency method (VanRaden, 2008). The dairy proportions for crossbred animals were estimated using the ADMIXTUTRE program (Alexander *et al.* 2009). The program was run setting the number of breed origins from K=2 to K=6. The farmer's prediction of exotic breed proportion was recorded and categorised into 7 groups: 0%, 25%, 50%, 65%, 75%, 85% and 100% dairy proportions. The correlation between farmer prediction and genomic prediction was estimated.

RESULTS AND DISCUSSION

Principal component 1 (PC1) separates *Bos taurus* from *Bos indicus* (Figure 1A) and explains 87% of the variation among animals (Table 1) PC2 separates the African *Bos taurus* N'Dama breed from other breed groups (*Bos indicus* and European *Bos Taurus ;*Figure 1A), and explains 1.84% of the variation (Table 1). Altogether the first five principal components explain 92% of the variation in the dataset.

Table 1. Summary of PCA for the first five principal components for the African dairy cattle and bovine Hapmap animals

Parameter	PC1	PC2	PC3	PC4	PC5
Standard deviation	3.677	0.533	0.474	0.347	0.217
Proportion of variance	0.877	0.018	0.014	0.009	0.003
Cumulative proportion	0.877	0.895	0.909	0.918	0.921

Ankole is one of the Sanga type breeds which are believed to include some African *Bos taurus* and *Bos indicus* ancestry. This is confirmed by their position in Figure 1A on the axis between the Nelore and N'Dama breeds. The Ankole animals do not cluster as tightly as the reference breeds. While a couple of Ankole outliers are clearly animals that have a small proportion of European *Bos taurus* content (they lie on the axis between Ankole and European taurine), it is not clear whether the remaining Ankole are simply more diverse than reference breeds or they might contain a low proportion of contamination with other breeds. The same picture is evident for the Nganda and SEAZ breeds, with diversity being substantially higher again. The SEAZ breed contains several animals that clearly have a substantial European taurine content. The SEAZ is not a single breed but consists of a range of small framed zebu breeds. The SEAZ position closer to the Nelore end of the Nelore to N'Dama axis, but not as distant from Ankole as their appearance might suggest: SEAZ look like *Bos indicus* while Ankole look substantially taurine.

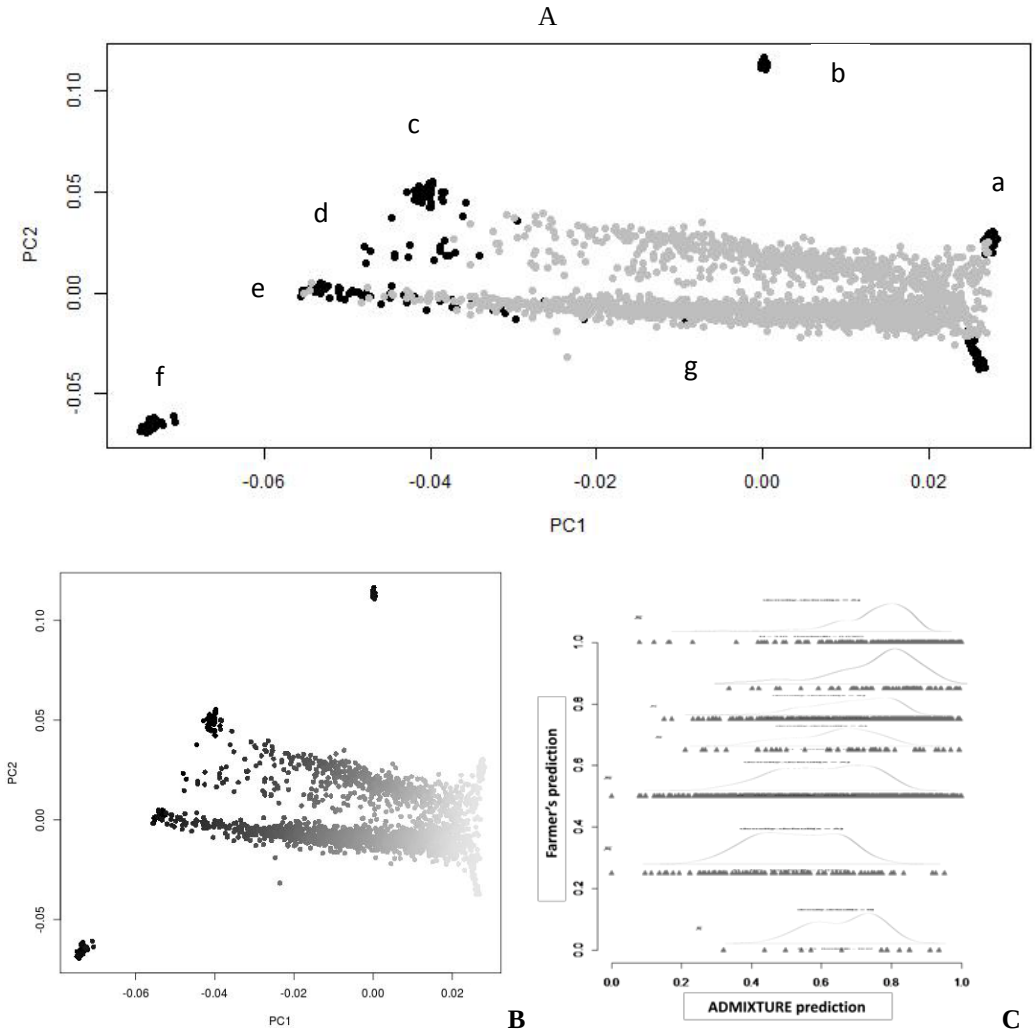


Figure 1. A. PC1 plotted against PC2 to visualise the genetic structure of East African cattle samples (a-European *Bos taurus*, b-N'Dama, c-Ankole, d-Nganda, e-SEAZ, f-Nelore, g-crossbred). B. Same as A, but with estimated dairy breed percentage from ADMIXTURE represented on a scale from black = 0% dairy to light grey = 100% dairy. C. Distribution of ADMIXTURE predictions of dairy breed proportion for individual animals shown for each class of farmer predicted breed composition.

The results for SEAZ agree with the study of Rege *et al*. (2001) where microsatellite allele data on autosomes indicated that all the Kenyan zebu breeds have a mixture of both taurine and indicine genes. Nganda animals are known to result from relatively recent (approximately 100 years) crosses between Ankole and SEAZ. Consistent with this origin, they lie in the middle of the axis between Ankole and SEAZ.

The estimated breed proportions of small holder crossbred dairy cattle vary from 0% to 100 %. Figure 1B shows the dairy proportion estimated using K=4 on a scale representing dairy

proportion from black (= 0% dairy) to light grey (=100% dairy). The Nelore, indigenous animals and N'Dama all are 0% dairy, while Holstein, Jersey and Brown Swiss are 100% dairy. The estimated proportion of dairy from ADMIXTUTE is shown for each of classes of farmer predicted breed composition in Figure 1C. The farmer's predictions were based on a combination of phenotypic appearance and varying degrees of knowledge of the ancestry of each animal. The class of 100% dairy includes all animals said to be of very high proportion dairy breed (>85%), rather than just absolutely pure animals; in practice the mean dairy breed proportion was 81%, and approximately 70% are less than 85% dairy. For cows predicted by farmers to be close to 50% dairy, the average dairy percentage is 61% and the range was from almost 0% to almost 100%. For the farmer's predicted classes of 65%, 75% and 85% dairy, the ADMIXTURE average estimates were 66%, 72% and 79%. The overall correlation of farmer's prediction and ADMIXTURE prediction is 0.41.

CONCLUSIONS

The population structure of East African smallholder dairy cattle is clearly illustrated using PCA and the results indicate that Ankole and SEAZ animals are quite similar composites of *Bos indicus* and *Bos taurus*, with Ankole having somewhat higher *Bos taurus* content but not so different as has been generally assumed given their physical appearance. All indigenous breeds have some animals that are clearly not purebred. Farmers' assessments/assumptions about the breed composition of their cows are poorly correlated with the cow's actual breed composition. This will be a problem when it comes to recommending the best breeding options for such cows, when the goal is produce a replacement female of a breed composition best suited to that farm environment.

AKNOWLEDMENTS

The authors gratefully acknowledge the contributions of other Dairy Genetics East Africa project team members, the collaborating farmers in Kenya and Uganda and the International Bovine Hap Map project for sharing their data. This work was funded by the Bill and Melinda Gates Foundation.

REFERENCES

Alexander D. H., Novembre J. and Lange K. (2009) *Genome Research*. doi: **10**.1101/gr.094052.109

Gondro C. and Gibson J. P. (2012) Centre for Genetic Analysis and Applications, University of New England. Armidale , Australia.

Rege J. E. O., Kahi A., Okomo-Adhiambo M., Mwacharo J. and Hanotte O. (2001) *Animal Genetic Resources Research 1* (pp. 103). Nairobi, Kenya: International Livestock Research Institute.

VanRaden P. M. (2008) *Journal of Dairy Science*, **91***;* 4414.

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:167-170

THE Fec*B* MUTATION INCREASES LAMB PRODUCTION IN SMALLHOLDER SUBSISTENCE FLOCKS IN MAHARASHTRA STATE OF INDIA

C. Nimbkar[1], R.S. Kataria[2], B.P. Mishra[2], B.K. Joshi[2] and P.M. Ghalsasi[1]

[1]Nimbkar Agricultural Research Institute, Phaltan 415 523, Maharashtra, India
[2]National Bureau of Animal Genetic Resources, Karnal 132 001, Haryana, India

SUMMARY

Introduction of the *FecB* mutation through heterozygous *FecB* carrier crossbred ewes was successful in increasing lamb production in nine smallholder sheep flocks managed traditionally. One copy of $FecB^B$ increased the number of live lambs born per lambing by 48% from 1.03±0.05 to 1.52±0.08, the number of lambs surviving at 3 months age from 0.96±0.07 to 1.42±0.13 and the weight of 3-months old lamb produced per lambing by 22% from 13.2±0.4 to 16.2±0.6 kg. Heterozygous ewes had a 36 days shorter lambing interval. Such strategies increase the efficiency and profitability of sheep production in the changing socio-economic and environmental milieu.

INTRODUCTION

About 100,000 smallholder sheep owners rear approximately 3.7 million non-prolific Deccani sheep on the dry, monsoonal Deccan plateau in Maharashtra State, India, as a source of livelihood. Lambs are usually sold at 3 months age in groups to butchers who pay a price based on a visual assessment. As a result, a sheep owner's income depends largely on the number of saleable lambs produced per ewe per year. The sheep production system is shepherded grazing on fallow and harvested fields, public (often degraded) lands, road and canal verges and farm bunds. Sale price of lambs has increased by 10 to 20 percent per year over the last 10 years due to the increases in the human population, urbanization, incomes and the gap between demand and supply.

A breeding program to increase the efficiency and profitability of lamb production by introgression of the *FecB* (Booroola) mutation which increases ovulation rate, from the small Garole breed into the Lonand Deccani breed type and a composite was established at the Nimbkar Agricultural Research Institute (NARI), Phaltan in south-western Maharashtra (Nimbkar *et al*. 2002). The new *FecB* carrier crossbred type comprised of only Deccani and Garole breeds was termed 'NARI Suwarna' (NS) and the composite comprised of Bannur and/or Awassi breeds additionally was termed 'NARI Composite' (NC).

The *FecB* mutation was introduced into nine smallholder flocks in January 2010 through the purchase of 94 pregnant heterozygous ($FecB^B/FecB^+$) crossbred ewes (comprising both NS and NC ewes) by flock owners from NARI with bank loans. Two of these flocks already had 27 $FecB^B/FecB^+$ ewes from earlier introductions of heterozygous and homozygous rams or semen. This study assesses the benefits of *FecB* carrier ewes in a largely traditional, low input subsistence farming system.

MATERIALS AND METHODS

Location. Performance records were collected in smallholder flocks in Bhadali village, 10 km south of Phaltan in Satara District, Maharashtra State, India (latitude 18^0 N and longitude 74^0 E).

Ewes. The smallest flock had 11 adult ewes; four flocks had 21 to 26 ewes each while the number of ewes in four flocks ranged from 35 to 46 each. Lambing and abortion records from the last quarter of 2009 to first quarter of 2012 were collected in nine smallholder flocks from 248 pregnant, non-carrier ($FecB^+/FecB^+$) ewes (482 records) and 114 crossbred $FecB^B/FecB^+$ ewes (244 records). Ewes purchased by flock owners from NARI in January 2010 provided 87 of the records of $FecB^B/FecB^+$ ewes and 26 of the records of non-carrier ewes while the remaining

records were from ewes already in the flocks at the end of 2009. Existing non-carrier ewes in the flocks were of the Lonand Deccani breed type with unknown admixture of the Madgyal breed, a hair sheep breed, taller, larger and faster growing than the Deccani, from southern Maharashtra and adjoining Karnataka state. The ewes purchased from NARI were crossbreds comprising of 40 to 94% Lonand Deccani, 6 to 28% Garole, 0 to 28% Bannur and 0 to 36% Awassi breeds. Seven $FecB^{B}/FecB^{+}$ ewes sourced from NARI had 50% Madgyal breed proportion. For about the last 10 years, Deccani sheep flock owners have been crossbreeding with Madgyal rams to improve lamb growth and adult size. The Deccani ewes already present in smallholder flocks were assumed to be non-carriers of *FecB* based on earlier studies (Pardeshi *et al*. 2005). The ewes introduced by NARI and those born in smallholder flocks from earlier carrier ram introductions were genotyped for the *FecB* locus at NARI using the PCR-RFLP direct DNA test (Wilson *et al*. 2001).

Animal management and records. All flocks were grazed by their owners under the traditional sheep-rearing system practised on the Deccan plateau. Some flocks migrated over approximately 50 km during the dry season between November and June. *FecB* carrier breeding rams were continuously run with the ewes. Lambs were not weaned; female lambs that were retained as replacements often suckled their dams until the dams naturally ceased lactating. Some flock owners provided supplementary feeding to lambs and lambed ewes. Cross-fostering was practised for twin-born lambs to other ewes in the flock, if the lambs' dam did not produce enough milk to maintain twin lambs. The only management interventions made by NARI were tagging, flock vaccination and treatment of sick ewes and lambs during routine visits. All ewes and lambs in the nine flocks were individually identified with ear tags. The flocks were visited once or twice a month and lambing (and abortions reported by flock owners), deaths, sales and lamb weights were recorded. Three month weight records were available for approximately half the total number of lambs surviving to 3 months as the flock owners sold lambs between NARI visits.

Traits analyzed. The ewe traits analyzed using the ASReml program (Gilmour *et al*. 2002) were:

1. Number of lambs born alive per lambing (NLBL): Zero if both lambs were stillborn (born dead on completion of full term) or one if one was alive and the other stillborn.
2. Number of lambs born alive per pregnancy (NLBP): Zeros for ewes that aborted (before term).
3. Number of lambs surviving to 3 months age per lambing (NL3M).
4. Weight of 3 months old lamb/s produced per lambing (WT3M).
5. Interval in days, between two lambings (LINT).

The traits NLBL, NLBP and NL3M were analyzed as Poisson variables with a log link while WT3M and LINT were analyzed as normal variables. Only fixed effects were fitted in univariate models for all variables. Fixed effects tested were flock or owner (9 classes), mating year-season for NLBL and NLBP (8 classes – summer, rainy and winter in 2009 and 2010 and summer and rainy in 2011), lambing year-season for NL3M (8 classes – rainy and winter in 2009 and summer, rainy and winter in 2010 and 2011), ewe breed type (Deccani vs. crossbred), ewe's *FecB* genotype (heterozygous carrier vs. non-carrier), the interaction of flock and ewe's *FecB* genotype and the covariable ewe's age in days. Garole breed proportion was additionally fitted for WT3M and LINT and age of the lambs at weighing was fitted for WT3M. An alternative model was fitted for WT3M using the fixed effect 'total number of lambs born per lambing' and for LINT using the fixed effect 'total number of lambs born in the ewe's previous lambing' instead of the ewe's *FecB* genotype. This was done because of the confounding between the ewe's *FecB* genotype and the two alternative fixed effects respectively. Least squares means (LSM) were estimated with only significant fixed effects in the model.

RESULTS AND DISCUSSION

Significance of fixed effects. The only fixed effect that was significant for the variables NLBL (P<0.001), NLBP (P<0.001) and NL3M (P=0.002) was the *FecB* genotype of the ewe. The P values of all other fixed effects and covariables were greater than 0.08 for these variables. However, for WT3M, the fixed effects of owner, lambing year-season, *FecB* genotype of the ewe, interaction between owner and ewe's *FecB* genotype, and the covariables age of the ewe, age of the lambs at weighing and Garole proportion of the ewe were all significant (P<0.012). Lambing year-season, *FecB* genotype of the ewe, Garole proportion of the ewe and age of the ewe were significant for LINT. Ewe's *FecB* genotype became non-significant for WT3M when the total number of lambs born was fitted as a fixed effect. Similarly, ewe's *FecB* genotype became non-significant for LINT when the number of lambs born in the previous lambing was fitted for LINT.

It thus appears that the differences in management among flock owners, annual or seasonal differences in feed availability did not influence ewe prolificacy or lamb survival significantly but they had a significant effect on WT3M and seasonal differences had a significant effect on LINT. Older ewes produced higher WT3M and had lower LINT. Garole breed proportion in the ewe had a negative effect on lamb weight but ewes with higher Garole proportion had lower LINT. Ewes that had twins in the previous lambing had a 12% shorter LINT than ewes that had singles. Almost all (89%) of these ewes were *FecB* carrier.

Effects of one copy of $FecB^B$. LSM of NLBL, NLBP, NL3M, WT3M and LINT for the fixed effect of ewe's *FecB* genotype, LSM of WT3M for the fixed effect of total number of lambs born per lambing and LSM of LINT for the number of lambs born in the previous lambing are reported in Table 1. One copy of $FecB^B$ increased NLBL by 48% from 1.03±0.05 to 1.52±0.08. Only two of the heterozygous ewes had triplets. The increase in NLBP was found to be 43% when the ewes that aborted were accounted for. Survival up to 3 months of both single and twin-born lambs was 94% leading to a 48% increase in NL3M with one copy of $FecB^B$ (1.42±0.13 vs. 0.96±0.07). The increase in WT3M due to one copy of $FecB^B$ was 22% and the reduction in LINT was 14%.

Table 1. Least squares means (LSM) and standard errors (SE) for all variables for the fixed effect of *FecB* genotype of the ewe and for the variables WT3M and LINT for the alternative models fitted using the fixed effects of total number of lambs born per lambing and total number of lambs born in the ewe's previous lambing respectively (n = number of records)

Variable	Fixed effect: Ewe's *FecB* genotype					
	$FecB^+/FecB^+$			$FecB^B/FecB^+$		
	n	LSM	SE	n	LSM	SE
NLBL	470	1.03	0.05	229	1.52	0.08
NLBP	482	1.00	0.05	244	1.43	0.08
NL3M	470	0.96	0.07	229	1.42	0.13
WT3M (kg)	228	13.23	0.36	122	16.20	0.60
LINT (days)	144	308.1	6.9	68	271.9	11.7
	Fixed effect: Total number of lambs born per lambing					
		1			2	
	n	LSM	SE	n	LSM	SE
WT3M (kg) (alternative model)	266	12.9	0.3	84	19.0	0.4
	Fixed effect: Total number of lambs born in the ewe's previous lambing					
		1			2	
LINT (days) (alternative model)	168	300.7	7.7	44	261.0	12.4

It therefore appears that smallholders managed to keep lamb mortality low and benefited from the moderate increase in NLBL. This is in contrast to the results of the ewe introduction in smallholder flocks in 2003 and 2004 (two of the flocks were the same as in 2010) when NLBL was 34% higher for heterozygous ewes compared to non-carrier ewes and NL3M was only 11% higher due to higher lamb mortality (Nimbkar *et al*. 2006). Some of the differences between the two introductions were that in the earlier introduction the ewes were given away while in 2010 the flock owners purchased ewes albeit at a subsidized rate and selected the ewes themselves. Additionally, the ewes introduced in 2003-04 were at least 25% Garole while in 2010, only 10% of the introduced *FecB* carrier ewes had more than 25% Garole proportion and 20% had less than 10% Garole proportion. Negative direct and maternal effects of the Garole on lamb survival and weight have been reported (Nimbkar 2006).

Ewes that had twin lambs produced 47% higher total weight of 3-months old lamb/s compared to ewes that had singles. It was unexpected that ewes producing twin lambs had a shorter lambing interval than single-bearing ewes as dams of twin lambs could be expected to undergo greater nutritional stress than dams of single lambs during lactation. Most (89%) of the ewes with twin lambs in their previous lambing and shorter lambing intervals were, however, heterozygous for $FecB^{B}$ suggesting a possible link between higher ovulation rate and quicker return to oestrus after lambing. Increased supplementary feeding to twin-bearing ewes before and after lambing and to their lambs is likely to maximize the benefits of the increased prolificacy and prove to be cost-effective.

CONCLUSIONS

Introgression of the *FecB* mutation into the Lonand Deccani strain of sheep while retaining the larger body size, hardiness, adaptation to harsh conditions and good mothering ability of the Deccani was found to be successful in sustainable improvement of sheep production in smallholder sheep flocks. This strategy is likely to be useful in the changing sheep production system due to declining grazing lands and increasing sedentarization.

ACKNOWLEDGEMENTS

The work described in this study was carried out under the project funded by the Department of Biotechnology, Ministry of Science and Technology, Government of India from April 2009 to March 2012. The funding from the Australian Centre for International Agricultural Research, Canberra, Australia, from 1998-2007 under projects AS1/1994/022 and AH/2002/038 for the *FecB* gene introgression project is also acknowledged gratefully.

REFERENCES

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. U.K.

Nimbkar, C., Ghalsasi, P.M., Walkden-Brown, S.W. and Kahn, L.P. (2002) *Proc. 7th Wld. Cong. Genet. Appl. Livest. Prod.* CD ROM Comm. No. 25-11.

Nimbkar C. (2006) PhD Thesis, University of New England, Australia.

Nimbkar C., Ghalsasi P.M., Chavan K.M., Nalawade M.H., Walkden-Brown S.W., Gupta V.S., Pardeshi V.C. and van der Werf J.H.J. (2006) Proc. 8th *Wld. Cong. Genet. Appl. Livest. Prod.* CD ROM Comm. No. 04-03.

Pardeshi V.C., Sainani M.N., Maddox J.F., Ghalsasi P.M., Nimbkar C. and Gupta V.S. (2005) *Curr. Sci*. **89**:887.

Wilson T., Wu X., Juengel J., Ross I., Lumsden J., Lord E., Dodds K., Walling G., McEwan J., O'Connell A., McNatty K. and Montgomery G. (2001) *Biol. Reprod.* **64**:1225.

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:171-174

ACCURACY OF GENOMIC PREDICTIONS IN NELLORE BEEF CATTLE

R. Carvalheiro[1], J. C. McEwan[2*], H.H.R. Neves[1], Y.T. Utsunomiya[1], A.M. Pérez O'Brien[3], S. A. Boison[3], J. Sölkner[3], F.S. Schenkel[4], C.P. Van Tassell[5], T.S. Sonstegard[5] and J.F. Garcia[6]

[1]UNESP - Univ Estadual Paulista. Faculdade de Ciências Agrárias e Veterinárias, Jaboticabal, São Paulo, Brazil. [2]AgResearch, Invermay, New Zealand. *Email: john.mcewan@agresearch.co.nz. [3]BOKU - University of Natural Resources and Life Sciences, Vienna, Austria. [4]Centre for Genetic Improvement of Livestock, University of Guelph, Canada. [5]ARS-USDA - Agricultural Research Service - United States Department of Agriculture. [6]UNESP - Univ Estadual Paulista. Faculdade de Medicina Veterinária de Araçatuba, São Paulo, Brazil.

SUMMARY

This study was carried out to assess the quality of genomic predictions in a Nellore beef cattle population, for 14 growth, carcass composition and reproduction traits, evaluated either at weaning or yearling. A forward prediction scheme was applied, so that information on a set of older animals (bulls and cows with accurate proofs in 2007) was employed to derive genomic prediction equations, while information on younger bulls (2012 proofs) was considered for validation purposes. The validation accuracies of genomic predictions averaged 0.47, consistent with the expectation for such statistics. Accuracies for two selection indexes including either weaning traits (WI) or both weaning and yearling traits (FI) were 0.44 and 0.58, respectively. For younger animals with own performance, genomic predictions increased by 10% (on average) the individual accuracies for both WI and FI.

INTRODUCTION

Since Meuwissen *et al.* (2001) showed that accurate predictions of the genetic merit of selection candidates could be obtained by using information from dense marker panels, the potential for incorporation of genomic selection (GS) in breeding programs has been an active research topic. It has been argued that GS schemes could increase the rates of genetic gain substantially, via increased accuracy of estimated breeding values and reduced generation intervals. It could also reduce costs for breeding organizations (Schaeffer 2006).

While cost-effectiveness of GS was evident for dairy cattle, justifying its current application in several countries, GS is still not employed by most beef cattle breeding programs. This is explained in part by the fact that GS methodology provides more modest benefit in beef cattle, especially for traits routinely recorded in early life, due to the differences in population structure, the smaller amount of information available to derive accurate and cost-effective prediction equations and the shorter generation intervals often found in beef cattle.

Despite these constraints, some potential for application of GS in *Bos indicus* (Nellore) beef cattle is envisaged. This breed plays an important role in beef production in Brazil. The large number of animals routinely recorded by breeding organizations and the fact that bulls are progeny-tested relatively late in life could justify GS, especially for important traits in which the progress achieved by conventional selection is currently limited. The aim of this study was to assess the quality of genomic predictions in a Nellore beef cattle population.

METHODS

Data. Phenotypic information consisted of EBVs obtained from routine evaluations of Conexão Delta G, a commercial beef cattle breeding program kept by an alliance of producers distributed across 12 states of Brazil. EBVs were based on records of 1,168,792 animals, collected between 1983 and 2012. Fourteen traits were considered, including weight traits, scrotal

circumference and carcass traits evaluated through visual scores (Table 1), as well as two selection indexes, including either weaning traits (WI) or both weaning and yearling traits (FI).

A total of 2,275 animals (influential bulls and cows) were genotyped with the Illumina® Bovine HD panel. Quality control of genotypes was carried out through an iterative process using the following SNP selection criteria: call rate > 0.98, minor allele frequency > 0.02, p-value for Hardy-Weinberg equilibrium test > 10^{-5}. SNPs meeting these criteria were further screened to interrogate their correlation with other syntenic SNPs located within a 100 marker window, allowing only one marker from each pair of highly correlated SNPs ($r^2 > 0.995$) to remain in the SNP dataset. Finally, samples with call rates < 0.90 were excluded from the analysis. The process was repeated until neither SNP nor samples were excluded, resulting in a final dataset of 995 bulls and 1,266 cows with 311,359 SNP called.

Study design. A forward prediction scheme was applied to compute genomic predictions. For each trait, the training population was composed of all genotyped animals with EBV's accuracy ≥0.50 in 2007, while younger bulls were included in the validation set (average accuracy of EBV of 0.88 in 2012). EBVs obtained in 2007 were considered as the response variable in model training, while EBVs obtained in the 2012 evaluation were used in the validation step, so information on validation animals (and their descendants) did not contribute to marker effect estimation.

Genomic predictions (DGV) were obtained using genomic BLUP (GBLUP), implemented using gebv software (Sargolzaei *et al.,* 2009). GBLUP equations included a modified genomic relationship matrix (G*), obtained as $G^{*}=0.8G+0.2A$, where G is the genomic relationship matrix computed similarly as in VanRaden *et al*. (2009), using observed allele frequencies. A is the regular numerator relationship matrix. The weights on G and A matrices were chosen based on previous analyses, which indicated better predictive ability with the adopted weights. A weighted analysis was conducted to account for differences in accuracies of EBVs of the training set. A diagonal matrix R was included in the GBLUP equations, whose entries were equal to $R_{ii}=(1/Rel_i)-1$, where Rel_i is the reliability of the i^{th} EBV.

In order to explore the importance of relatedness in the predictive ability, an alternative version of GBLUP was fitted. A modified genomic relationship matrix was computed as $G^{**}=0.999G_1+0.001A$, where G_1 is a genomic relationship matrix built considering exclusively the markers in BTA1, and A as described previously. The reasoning of this strategy is that information in one chromosome is expected to capture relationships, albeit this is subject to the assumption that LD is consistent across the genome, but is unlikely to contain all QTL (Daetwyler *et al.,* 2012). The relative influence of relationships on the accuracy of genomic predictions were assessed by contrasting the accuracies obtained using either G^{*} or G^{**} in GBLUP equations.

Analysis of results. The Pearson's correlation between DGV and EBV of the validation animals ($r_{EBV,DGV}$) was considered as a proxy for the accuracy of prediction. In order to evaluate the amount of uncertainty about such correlations, 95% confidence intervals for $r_{EBV,DGV}$ were obtained using Fisher's Z transformation. The intercept (b_0) and the slope (b_1) of the regression of EBV on DGV were evaluated to describe bias and inflation of the predictions, respectively. The expected values for accuracy of prediction $cor(g,\hat{g})$ were obtained using a deterministic formula (Daetwyler *et al.,* 2010), by assuming a value of 120 for the effective population size and that the average reliability of the EBVs would be equal to the heritability of the pseudo-phenotypes available for model training.

RESULTS AND DISCUSSION

The accuracies of the genomic predictions varied from 0.24 (birth weight and gestation length) to 0.68 (finishing precocity at yearling), with an average of 0.47 (Table 1), although the amount of uncertainty about such estimates was usually large, as a consequence of the relatively small

validation sets in this study. The expected values for these statistics, calculated using the formula proposed by Daetwyler *et al.* (2010), ranged from 0.48 (scrotal circumference) to 0.58 (weight gain from birth to weaning). In previous analyses using just sire genotypes, the use of deregressed EBVs (Garrick *et al.*, 2009) as pseudo phenotypes generated very similar results as using EBVs, while different statistical methods (e.g. BayesC and Bayesian LASSO) generated results similar to those of GBLUP (data not shown).

Table 1. **Summary statistics[2] of genomic predictions in Nellore beef cattle**

Trait[1]	Ntrain	accT	Nval	accV	rEBV,DGV (95% CI)	rPEV	cor(g,ĝ)	BTA1	b_0 (%)	b_1
WG	1,757	0.70	179	0.90	0.38 (0.25,0.50)	0.53	0.58	0.33	-23.2	0.80
CW	1,688	0.69	179	0.90	0.40 (0.27,0.51)	0.53	0.56	0.38	-9.3	1.12
PW	1,689	0.69	179	0.90	0.53 (0.41,0.62)	0.53	0.56	0.50	-21.0	0.99
MW	1,688	0.69	179	0.90	0.54 (0.42,0.63)	0.53	0.57	0.51	-20.5	1.02
NW	1,682	0.69	179	0.90	0.47 (0.35,0.58)	0.53	0.56	0.42	26.6	0.92
PWG	1,655	0.69	108	0.88	0.53 (0.38,0.65)	0.53	0.56	0.45	-19.6	0.95
CY	1,628	0.68	108	0.88	0.45 (0.29,0.59)	0.53	0.55	0.39	8.5	1.05
PY	1,628	0.68	108	0.88	0.68 (0.56,0.77)	0.53	0.55	0.63	-17.6	1.12
MY	1,628	0.68	108	0.88	0.64 (0.51,0.74)	0.53	0.55	0.60	-16.3	1.11
NY	1,625	0.68	108	0.88	0.47 (0.31,0.61)	0.53	0.55	0.43	24.7	0.99
SC1	1,089	0.70	56	0.83	0.46 (0.23,0.65)	0.50	0.49	0.39	-43.0	0.72
SC2	1,078	0.70	61	0.83	0.49 (0.28,0.66)	0.50	0.48	0.45	-33.9	0.55
BW	1,686	0.67	124	0.88	0.24 (0.07,0.40)	0.51	0.55	0.21	-3.4	0.38
GL	1,339	0.67	76	0.90	0.24 (0.01,0.44)	0.47	0.51	0.17	-15.5	0.67
WI	1,692	0.70	179	0.90	0.44 (0.31,0.55)	0.53	0.57	0.40	-25.4	0.90
FI	1,413	0.70	140	0.88	0.58 (0.46,0.68)	0.52	0.53	0.52	-21.0	0.92
Overall	1,560	0.69	129	0.88	0.47(0.32,0.60)	0.52	0.55	0.44	-13.1	0.89

[1]WG: weight gain from birth to weaning (about 205 days of age); CW, PW, MW, NW: visual scores taken at weaning for carcass conformation, finishing precocity, muscling and navel, respectively; PWG: weight gain from weaning to yearling (about 550 days of age); CY, PY, MY, NY: visual scores taken at yearling for carcass conformation, finishing precocity, muscling and navel, respectively; SC1 and SC2: scrotal circumference adjusted for age and for age and weight, respectively. BW: birth weight; GL: gestation length; WI: weaning index, composed by traits evaluated at weaning; FI: final index, composed by traits evaluated at weaning and yearling (FI).

[2]Ntrain (Nval): number of animals in the training (validation) set; accT (accV): average accuracy of EBVs of training (validation) animals; rEBV,DGV: validation accuracy, i.e. Pearson's correlation between EBV and genomic prediction (DGV); (IC95%): limits of the 95% confidence interval for rEBV,DGV; rPEV: average of the individual accuracies of DGV (obtained using elements of the inverse of the coefficient matrix); cor(g,ĝ): expected accuracy of prediction according to a deterministic formula (Daetwyler *et al.*, 2010). BTA1: validation accuracy obtained using only marker information from BTA1; b_0 and b_1: intercept and slope of the regression of EBV on DGV, respectively (b_0 is expressed relatively to the standard deviation of the EBVs for each trait, in %).

As a general rule, the validation accuracies matched well with the expected values (Table 1), although substantial departure from expectation was observed for birth weight and gestation length (accuracies about 50% lower than expected). Also, higher than expected accuracies were found for finishing precocity (PY) and muscling (MY) scores at yearling. This result seems to confirm the findings of Carvalheiro *et al.* (2012), who also obtained higher than expected accuracies for these traits, after fitting GBLUP with a smaller training set, and suggested that these traits were affected by the presence of genotype stratification associated to differences of the within group EBV means. Results of a principal component analysis based on the genomic relationship matrix

evidenced the existence of two subgroups of the sampled population. Further inspection corroborate the hypothesis that PY and MY are both affected by genotype stratification associated to a large difference of EBV subgroup means for these traits (data not shown).

The accuracies for two selection indexes, including just weaning traits (WI) or both weaning and yearling traits (FI) were about 0.44 and 0.58, respectively. Overall, the DGVs were slightly inflated (slope ~0.89) and overestimated (intercept < 0), although there was some variation in this pattern across traits. The most inflated estimates were verified for BW, SC2 and GL, while the predictions for SC1 and SC2 were those for which the overall mean of DGVs departed most from that of EBVs.

The averages of the individual accuracies of DGV (computed using the estimated Prediction Error Variance - rPEV) were more consistent across traits (average ~ 0.52) than rEBV,DGV and also matched more closely to the expected values for accuracy of prediction (Table 1). For younger animals with own performance recorded, the genomic predictions increased by 10% (on average) the individual accuracies for both WI and FI, when compared with the accuracies of traditional EBVs (data not shown).

This increase in accuracy was found to be associated to the relatedness of each young animal to the training set, so that individual accuracies of DGV for animals with a sire in the training set were increased by 20% (on average), compared to traditional evaluations. Similar association was also verified by Clark *et al.* (2012), who found strong correlations between different measures of relatedness to the training set and rPEV.

When averaged across traits, about 90% of the accuracy obtained when considering information from all chromosomes was recovered using only the information from BTA1, which reinforces the importance of relatedness (population structure) contribution to the accuracy of genomic predictions in this population, as is unlikely that all the QTLs affecting the studied traits are located on BTA1.

While the present study focused on routinely recorded traits, larger benefit is expected for traits in which the genetic progress achieved through conventional selection is limited (e.g. sexual precocity, productive longevity and meat quality), and this should be the subject of future research.

CONCLUSIONS

The current genomic predictions matched reasonably well the expectations and increased by 10% (on average) the individual accuracies of younger animals with own performance for two selection indexes, including either weaning traits or both weaning and yearling traits.

REFERENCES

Carvalheiro R., Perez O'Brien A.M., Neves H.H.R., Sölkner J., Utsunomiya Y.T., and Garcia J.F. (2012) *Proc of 63rd EAAP Annual Meeting*, EAAP, Bratislava, Slovakia.

Clark S.A., Hickey J.M., Daetwyler H.D., and van der Werf J.H.J. (2012) *Gen. Sel. Evol.* **44**:4.

Daetwyler H.D., Kemper K.E., van der Werf J.H.J, and Hayes B.J. (2012) *J. Anim. Sci.* **90**:3375.

Daetwyler H.D., Pong-Wong R., Villanueva B., and Woolliams J.A. (2010) *Genetics* **185**:1021.

Garrick D., Taylor J.F., and Fernando R.L. (2009) *Gen. Sel. Evol.* **41**:55.

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

Sargolzaei M., Schenkel F.S., and VanRaden P.M. (2009) In 'Technical report to the Dairy Cattle Breeding and Genetics Committee', University of Guelph, Guelph. Available at http://cgil.uoguelph.ca/dcbgc/Agenda0910/GEBV.pdf

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

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

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:175-178

THE IMPACT OF GENOMIC SELECTION ON GENETIC GAIN IN THE NEW ZEALAND SHEEP DUAL PURPOSE SELECTION INDEX

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

AgResearch, Invermay Agricultural Centre, Mosgiel, New Zealand

SUMMARY

To identify the impact of using molecular breeding values (mBVs) on the New Zealand sheep dual-purpose (DP) index, genomic selection (GS) accuracies were estimated using a training and validation data set consisting of 4,237 genotyped and pedigree recorded Romney animals. Molecular BVs and their accuracies for a range of DP production traits including live weight, fleece weight, faecal egg count, dagginess, reproduction and survival were estimated. The Romney mBV accuracies ranged from 0.16 to 0.52. For the majority of production traits the accuracies of the mBVs contributed information equivalent to having 1 to 8 measured progeny. For the traits: number of lambs born, lamb survival and lamb survival maternal the mBVs contributed between 11 to 145 measured progeny, albeit lamb survival maternal had a large error estimate. Combined with reducing the generation interval of rams used, from 2 years to 1 year, the potential increase in genetic gain in using mBVs in a New Zealand DP index was estimated to be 84%.

INTRODUCTION

The development of high density single nucleotide polymorphism (SNP) chips has allowed the development of GS which enables prediction of an animal's worth (via mBVs) from their genomic information at birth. In the dairy industry, GS has been implemented in many countries (Hayes *et al.* 2009). In the sheep industry, genomic information has been successfully implemented for many traits and breeds in New Zealand (Auvray *et al.* 2011). However, the impact of genomic selection on the New Zealand sheep industry has not been examined. The aim of this paper is to estimate the increase in genetic gain attainable for the Romney breed using the Sheep Improvement Limited (SIL) DP selection index plus resistance to internal parasites, dagginess and lamb survival.

MATERIALS AND METHODS

Data. Phenotypes (as estimated breeding values, eBVs) and pedigree data were downloaded from SIL; the export consisted of 3,535,557 animals born between 1990 and 2010 for 233 SIL flocks. Traits included in this analysis were direct and maternal weaning weight at 3 months (WWT, WWTm), carcass weight (CW), live weight at 8 months (LW8), adult ewe weight (EWT), lamb fleece weight (LFW), fleece weight at 12 months (FW12), adult ewe fleece weight (AFW), dag score at 3 and 8 months (DAG3, DAG8), number of lambs born (NLB), direct and maternal lamb survival (SURV, SURVm) and faecal egg count in summer (FEC1) and autumn (FEC2) and as an adult (AFEC).

There were 4,237 SIL recorded animals, mainly sires at least 70% Romney that were genotyped on the Illumina Ovine SNP50BeadChip (50K). Genotyping results were put through a quality control pipeline before analysis (Dodds *et al.* 2009); including removal of SNPs not retained as part of the Ovine HapMap study (Kijas *et al.* 2012). There were 48,327 SNPs which passed quality control. The animals were split into training and validation sets for each trait. Cut off years were chosen so at least 200 animals were used for validation.

Statistical analysis. Molecular breeding values (mBV) were calculated for each trait using genomic BLUP (gBLUP) model using the methods of Garrick *et al* (2009) and VanRaden (2008), fitting the G1 (VanRaden 2008) matrix. The first 6 principal components (PC), using G1 as a similarity matrix, were also fitted to adjust for breed effects.

The accuracies of the mBVs were derived from the validation animals using 2 different methods. For the first method (using G1); $r_A = cor(y,mBV)/h_g$ calculated using weights: $1/(1-r^2)$. The effective heritability from the GS analysis (h^2_g) was used and is equal to the average reliability (r^2) of the parent-average-removed, deregressed eBV, y. The second method uses the prediction error variance (PEV, Mrode 2005) from a gBLUP analysis (using a relationship matrix calculated using breed-specific allele frequencies) giving; $r_I = \sqrt{(1-(PEV_i/\sigma^2_u))}$, where σ^2_u is the genetic variance. This was calculated for all validation animals and averaged (weighted by $1/(1-r^2)$). The 'combined-accuracy' (r_C) was taken as the average of r_A and the r_I.

Comparison of genetic gain. The multiple trait selection index worksheet (van der Werf 2006a) was used to estimate the response per selection round for a given breeding scheme scenario in a DP Romney flock. The breeding scheme was simulated and assumed; a flock of 631 ewes; rams used once at a ratio of 1:90; number of lambs weaned/ewes mated was 141% (NLB: 1.71, lambs weaned/lambs born: 0.86 and ewes present at lambing/ewes mated: 0.98 (McEwan *et al.* 1992, Jopson *et al.* 2000, Pickering *et al.* 2012)); ewes lambed first at 2 years of age and retained to 5 years of age, with a 10% death and culling rate each year.

Selection was on a DP index with emphasis on increase kg of lamb, fleece weight, number of lambs, disease resistance, lamb survival and decrease dag score per ewe per ha. Heritability, repeatabilities and genetic and phenotypic correlations were from Pickering *et al* (2012) or were those used for SIL breeding value analysis (S. A. Newman, pers. comm.). This paper utilises the breeders equation: Genetic gain (ΔG) = $ir\sigma_a/L$, where i is the selection intensity, r is the accuracy, σ_a is the genetic standard deviation and L is the generation interval. The paper examines changes to r and L under the following scenarios:

- Scenario 1 assumed selection was on animal measurements either for a ram hogget (Scenario 1-A) or a 2 year old ram (Scenario 1-B) which is used only once.
- Scenario 2: Romney ram hoggets were genotyped with a 50K SNP chip, the number of equivalent progeny was estimated using r_C (van der Werf 2006b). This method assumes that the information of the mBVs and traditional eBVs are independent and this approach is equivalent to simple blending as outlined by Mrode (2005).

Dual-purpose economic weights for the traits were taken from Byrne *et al.* (2012). The FEC1, FEC2 and AFEC economic weights were converted from % to loge by multiplying by 100.

The maximum number of measurements available were; 1 measurement on the individual and sire, 2 on the dam (for NLB) and 126 on half sibs. For CW, LFW, AFW, FEC2 and AFEC no measurements were taken; these traits were estimated from their correlations with the other traits. Results were converted from selection response per 'selection round' ($r*\sigma_a$) to selection response 'per year' (ΔG) by multiplying by i/L. For scenario 1-B: i/L equalled 1.73/2.68 = 0.64 and for scenario 1-A and 2: i/L equalled 1.73/2.18 = 0.79.

RESULTS AND DISCUSSION

The accuracies (r_C) ranged between 0.16 and 0.52, equivalent to between 1 and 145 progeny (Table 1). For traits with low heritability the number of equivalent progeny equal to the accuracy of the SNP chip is large e.g. SURV and SURVm. For traits that are easy to measure and have moderate heritabilities, the SNP chip is equal to 1 or 2 equivalent progeny e.g. AFW and EWT.

The selection response per 'selection round' and 'per year' for each selection scenario is shown in Table 1. Scenario 1-B, resembling a farmer's normal decision using 2 year old rams and no SNP chip, had a genetic response of $1.43 per year (accuracy 0.34). Reducing the generation interval, by using ram hoggets (scenario 1-A) increased genetic response to $1.72 per year (accuracy 0.33). Scenario 2, selection of a hogget ram with SNP chip information had a genetic response of $2.63 (accuracy 0.51), an 84% increase compared to scenario 1-B.

Table 1. The response to each selection scenario as the unit change in trait (Δ units), the overall response per selection round (Response $) and accuracy, conversion factor (i/L, selection intensity/generation interval), and rate of genetic gain (ΔG), for each scenario[1]. Economic weighting (EW, $) for each trait and the accuracy of the SNP chip (r_C) and as equivalent progeny (E prog).

Trait		1-A	1-B	SNP chip		2
	EW	Δ units	Δ units	r_C	E prog	Δ units
Weaning weight	0.95	0.49	0.48	0.48	8	0.42
WWT maternal	0.84	0.25	0.26	0.34	3	0.32
Carcass weight	2.60	0.2	0.18	0.48	4	0.12
Live weight 8 months	0.00	0.56	0.54	0.50	3	0.45
Adult ewe weight	-1.04	0.12	0.04	0.48	2	-0.30
Lamb fleece weight	1.82	0.01	0.01	0.29	2	0.01
Fleece weight 12 months	0.79	0.02	0.04	0.50	3	0.02
Adult fleece weight	2.28	0.06	0.08	0.32	1	0.06
Number of lambs born	15.55	0.01	0.01	0.52	16	0.05
Survival	64.45	0.003	0.003	0.16	11	0.003
Survival maternal	58.40	0.0001	0.0001	0.48	145	0.005
Dag score 3 months	-0.34	-0.11	-0.11	0.40	2	-0.08
Dag score 8 months	-0.35	-0.09	-0.09	0.44	3	-0.07
Faecal egg count summer	-3.00	-0.07	-0.06	0.46	6	-0.07
Faecal egg count autumn	-3.00	-0.04	-0.04	0.50	7	-0.06
Adult faecal egg count	-2.00	-0.07	-0.07	0.41	2	-0.09
Response $		2.18	2.22			3.32
Accuracy		0.33	0.34			0.51
i/L		0.79	0.64			0.79
ΔG ($)		1.72	1.43			2.63

[1] selected on individual, dam, sire and half sib records as ram hogget (1-A), 2 year old ram (1-B), or ram hogget plus SNP chip based on breed combined-accuracies from genomic selection (2).

The majority of the gain was seen in the lowly heritable traits, or sex limited traits measured late in life. Also a proportion of the gain resulted from reducing the generation interval by using ram hoggets rather than 2 year old rams. The annual response in an Australian terminal index and a fine wool index after including genomic selection increased by 32% and 38% increase respectively (van der Werf, 2009). The results calculated here for scenario 2 are considerably larger than that presented by van der Werf (2009). They reflect differences in the economic weighting and accuracy of the genomic mBVs for the traits in the respective New Zealand and Australian breeding objectives. The example presented assumes that all animals in a flock are genotyped and that rams used are all of the same age. In practice, the actual response will vary by flock depending on the composition of the flock, breeding strategy and cost of SNP chips. Rams used will be a mixture of new untested rams, emerging rams used once before and mature tested rams. The current comparison also does not take into account the cost of genotyping. To maximise discounted financial returns, 2-stage selection would be used and only a proportion (10-20%) of ram lambs would be genotyped (Sise *et al.* 2011). This would effect a slight reduction on the

average mBV accuracies of the flock. In addition, costs would be reduced further by use of lower density chips, such as the 5K Ovine SNP chip, coupled with imputation. This would have minimal impact on the estimated mBV accuracies as shown by Berry and Kearney (2011) who estimated an average 97% correlation between mBVs estimated from imputed or real genotypes.

CONCLUSIONS

Genomic selection can provide a significant increase in the rate of genetic gain per year when selecting on the New Zealand dual purpose index. The majority of the benefit comes from the increased accuracy of breeding value for sex-limited and measurements recorded later in life. Additional benefits can be derived by reducing the generation interval via use of ram hoggets. This comparison did not include facial eczema, flystrike or adult ewe longevity which will also greatly benefit from use of genomic selection.

ACKNOWLEDGMENTS

Ovita and B+LNZ are acknowledged for their funding. Ovita and New Zealand sheep breeders are acknowledged for their permission for use of their data. Sheryl-Anne Newman is acknowledged for generation of eBVs and assisting with genetic parameters.

REFERENCES

Auvray B., Dodds K. G. and McEwan J. C. (2011) *Proc. N. Z. Soc. Anim. Prod.* **71**: 263.

Berry D. and Kearney J. (2011) *Animal*. **5**:1162.

Byrne T. J., Ludemann C. I., Amer P. R. and Young M. J. (2012) *Livestock Science*. **144**: 20.

Dodds K. G., Auvray B., Pickering N. and McEwan J. C. (2009) *Proc. Assoc. Advmt. Anim. Breed. Genet.* **18**: 296.

Hayes B., Bowman P., Chamberlain A. and Goddard M. (2009) *J. Dairy Sci.* **92**: 433.

Jopson N. B., Dodds K. G., Knowler K. J., Wheeler R. and McEwan J. C. (2000) *Proc. N. Z. Soc. Anim. Prod.* **60**: 47.

Kijas J. W., Lenstra J. A., Hayes B., Boitard S., Porto Neto L. R., San Cristobal M., Servin B., McCulloch R., Whan V., Gietzen K., Paiva S., Barendse W., Ciani E., Raadsma H., McEwan J., Dalrymple B. and other members of the International Sheep Genomics Consortium. (2012) *PLoS Biol.* **10**:e1001258.

McEwan J. C., Davis G. H., Dodds K. G., Fennessy P. F., Clarke J. N., Bruce G. D. and Hishon M. G. (1992) *Proc. N. Z. Soc. Anim. Prod.* **52**: 211.

Mrode R. A. (2005) Linear models for the prediction of animal breeding values. CAB international, Walingford, Oxfordshire, OX0 8DE, UK.

Pickering N. K., Dodds K. G., Blair H. T., Hickson R. E., Johnson P. L. and McEwan J. C. (2012) *J. Anim. Sci.* **90**: 1411.

Sise J. A., Linscott E. M. and Amer P. R. (2011) *Proc. N. Z. Soc. Anim. Prod.* **71**: 266.

van der Werf J. H. J. (2006a) Teaching software used for quantitative genetics at UNE: Multiple trait selection index, accessed 4 Apr 2012. http://www-personal.une.edu.au/~jvanderw/software.htm

van der Werf J. H. J. (2006b) Teaching software used for quantitative genetics at UNE: Single trait selection index, accessed 4 Apr 2012. http://www-personal.une.edu.au/~jvanderw/software.htm

van der Werf J. H. J. (2009) *Proc. Assoc. Advmt. Anim. Breed. Genet.* **18**: 38.

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:179-182

COMPARISON OF THE POWER OF POOLED GENOTYPING STRATEGIES TO DETECT SIGNIFICANT SNP EFFECTS FOR FLYSTRIKE RESISTANCE IN MERINO SHEEP

S. Dominik, J.M. Henshall and J. Smith

CSIRO Animal, Food and Health Sciences, New England Highway, Armidale, NSW 2350

SUMMARY

Genotyping pooled DNA is a cost-effective strategy to produce genotype information for whole genome association studies. The objective of this study was to compare the power to detect significant SNP effects for flystrike of simulated genotyping strategies in the Breechstrike Resource flock. A gene dropping approach was used to simulate allele frequencies. Individual genotyping was used to set a benchmark for comparison with three DNA pooling strategies. These included pooling before fixed effect adjustment, pooling after fixed effect adjustment and case-control pooling. The study showed that the highest power to detect significant associations between SNP allele frequencies and phenotypes can be achieved by individual genotyping. Case-control pooling and pooling after fixed effect adjustment had similar power to detect significant SNP effects, whereas pooling before fixed effect adjustment performed worst. The high power of detection of SNP effects of the individual genotyping strategy indicates that the Breechstrike Resource flock is a suitable resource for the detection of significant SNP effects for flystrike. However, pooling affects the power of detection of significant SNP effects, in particular when effects are small.

INTRODUCTION

Fly strike in Merino sheep, where fly larvae feed into the tissue, is a welfare issue in the Merino sheep industry. Traditionally, mulesing has been used as a strategy to prevent fly strike. Important and difficult to measure phenotypes, such as fly strike resistance, are ideal targets for genomic approaches. Whole genome association studies are most powerful where very large numbers of individuals are genotyped, however, individual genotyping is very costly. There is growing evidence that pooled DNA can be used successfully. This was mostly demonstrated for binomial phenotypes (Lee 2005, Huang *et al*. 2010), however, it has also been shown that DNA pooling is an effective strategy for quantitative traits (Henshall *et al.* 2012). Whilst pooled genotype strategies are more cost-effective than individual genotype strategies, the trade-off is a loss of power to detect SNP effects. The objective of this study was to compare simulated DNA pooling strategies with individual genotyping for their power to detect significant SNP effects for flystrike in the Breechstrike Resource flock.

MATERIALS AND METHODS

Data. Phenotype and pedigree data of the Breechstrike Resource flock (Smith 2009) were used for this study. The pedigree contained 3109 individuals, including 463 base animals born in 2003 and 2646 progeny born between 2005 and 2011, of which 2274 have flystrike phenotypes. Founder animals originated from three genetic groups: Ultrafine/Superfine (US), Fine/Fine Medium (FFM) and Medium/Strong (MS). The genetic groups provided the structure for the gene dropping approach used in this study. Contemporary groups were formed by sex (male or female), birth year (2005 – 2011) and mulesing status (mulesed or unmulesed), which were recorded on each animal that had a flystrike record. Phenotypes included flystrike (struck / not struck), wrinkle (high (H), moderate (M) and low (L)) and wool cover on the breech (high (H), moderate (M) and

low (L) breech cover). Flystrike, wrinkle and breech cover phenotypes formed phenotype classes. The combination of phenotype and contemporary groups was used for the DNA pool assignment.

Gene dropping. Allele frequencies for each genetic founder group (US, FFM and MS) were simulated using gene dropping (MacCluer et al. 1986). Single nucleotide polymorphism (SNP) genotypes were assigned at random to base animals, and then transmitted through the pedigree subject to Mendelian inheritance rules, i.e. progeny have equal chance of inheriting each of the two alleles carried by the parent. One hundred SNP were simulated and the gene dropping procedure was repeated 100 times. Simulations were run with and without an existing association between alleles and phenotypes. When an association between SNP alleles and phenotype was simulated, SNP effect sizes ranged from 0.1 to 1 phenotypic standard deviation (σ_p).

Genotype pooling strategies. Three pooling strategies were tested and compared to individual genotyping (Strategy 1). Individual genotyping of 2274 animals served as a benchmark for the other three strategies. In Strategy 2, animals were pooled for genotyping within phenotype (flystrike – struck/not struck, wrinkle, breechcover) and contemporary group (birthyear, sex and mulesing status). The numbers and size of pools resulting from Strategy 2 are shown in Table 1. In Strategy 3, flystrike phenotypes were adjusted for fixed effects (birthyear, sex and mulesing status). Individuals for genotyping were pooled with the objective of creating a balanced number of pools in the struck and not struck group and achieving even pool sizes (number of individuals per pool) across phenotype / contemporary groups (Table 1). Strategy 4 is a case-control pooling approach and uses a combination of individual and pooled DNA for genotyping. All struck animals are individually genotyped and matched, if possible, with an individual genotype of a not struck animal from the same contemporary / phenotype group. Not struck animals that were not paired with a struck animal are pooled within contemporary / phenotype group. Numbers and sizes of pools resulting from Strategy 4 are summarised in Table 1.

Table 1. Numbers of pools and pool size resulting from three pooling strategies; Strategy 1 being individual genotyping.

	Pool size	Number of pools		
	Number of individuals	Total	Struck	Not struck
Strategy 2				
	1	318	183	135
	5	174	20	154
	6	181	15	166
Strategy 3				
	1	373	373	0
	5	379	0	379
	6	1	0	1
Strategy 4				
	1	717	373	344
	$\leq$ 10	34	0	34
	11-20	21	0	21
	> 20	29	0	29

Analysis. Associations between flystrike phenotypes, “struck” and “not struck”, and allele frequencies were established by logistic regression. The phenotype and contemporary groups were included in the model as fixed effects. Analysis were conducted with software written in R (R Development Core Team 2008)

RESULTS AND DISCUSSION

For simulations without associations between phenotypes and allele frequencies, no more significant SNP than expected by chance were found with any of the genotyping strategies at significance levels P < 0.05, P < 0.01 and P < 0.001 (Table 2).

Table 2. Relative frequency of detection (%) of SNPs associated with the phenotype at three significance levels; no association between phenotype and allele frequency was simulated

Significance level	Relative frequency of detection (in %)			
	Strategy 1	Strategy 2	Strategy 3	Strategy 4
< 5%	3.60	3.30	3.60	3.40
< 1%	1.10	0.30	0.40	0.40
< 0.1%	0.03	0.01	0.02	0.01

When SNP effects of varying size were simulated to be associated with allele frequencies, associations were detected more often than just by chance (Figure 1). For all strategies the power to detect SNP effects increased with increasing SNP effect size. In reality most SNP effects are small and pooling strategies 2, 3 and 4 were not very powerful in detecting them. Strategy 1 (individual genotypes) had the highest power to detect small SNP effects ($0.1\sigma_p$), with a relative frequency of 18% (P < 0.05), whereas with Strategy 2, 3 and 4 ranged from 5.5% to 7%. Strategies 3 and 4 yielded similar results, with Strategy 3 being slightly more powerful (Figure 1). Strategy 2 was the least powerful approach at detecting significant SNP associations with phenotypes compared to all other genotype pooling strategies.

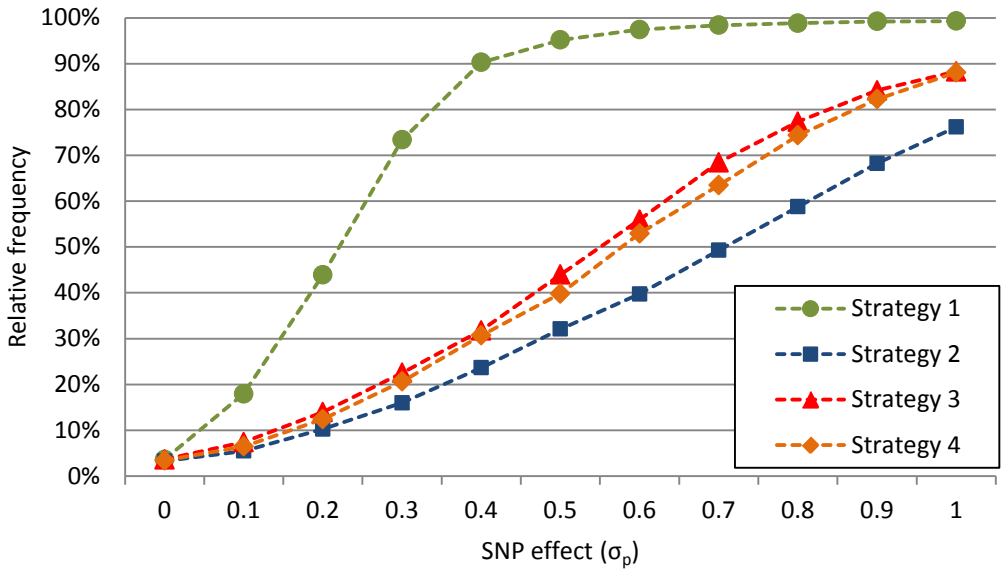


Figure 1. Relative frequency of detection of significant SNP effects (in %) of varying size (x axis in σ_p) with four different genotyping strategies at significance level P < 0.05.

Figure 2 shows the frequency of detection of SNP effects for Strategies 1 and 3 to demonstrate the difference of detection at different significance levels. The pooled genotyping strategies use

between 65-70% fewer assays compared to individual genotyping and the power of detecting significant SNP effects of small size reduced by 64-78% with small effects up to $0.4\sigma_p$ and 11-31% with large SNP effect sizes of $0.8\text{-}1\sigma_p$. Power was lower than expected based on results presented by Henshall *et al.* (2012). Huang *et al.* (2010) suggested increasing pool sizes as much as possible to estimate allele frequencies accurately, but this was not possible in this study due to small number of individuals in each phenotype / contemporary group class.

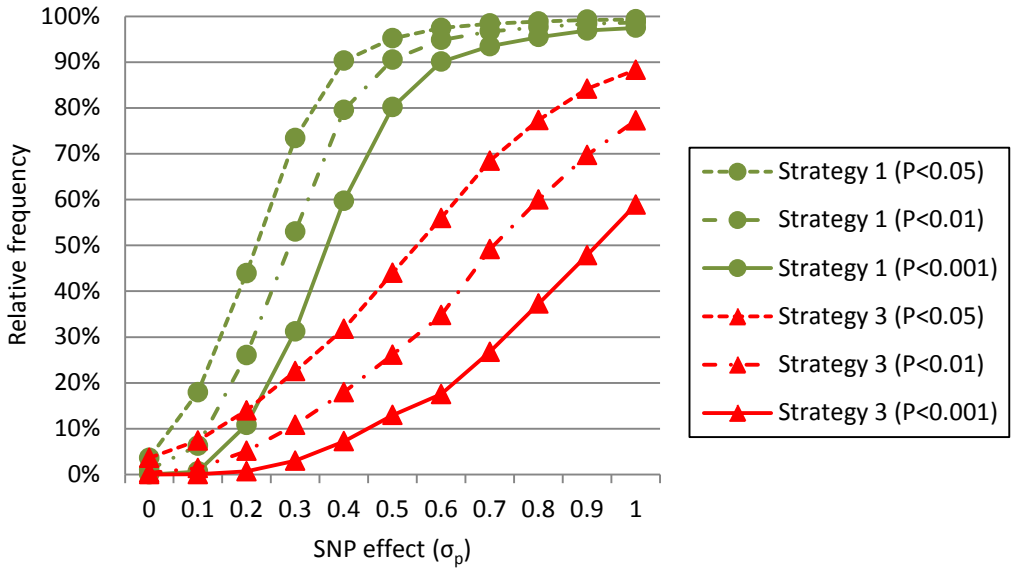


Figure 2. Relative frequency of detecting significant SNP effects of varying size (x axis in σ_p) with two pooling strategies 1 and 3 at three significance levels.

CONCLUSIONS

The power of DNA pooling approaches is affected by the nature of the phenotype and the number of contemporary groups in the data set. Pooling strategies lack power in the detection of small SNP effects; however, they could still provide a cost-effective alternative for the estimation of genomic breeding values. Pooling strategies should be designed and tested for their power prior to implementation.

ACKNOWLEDGEMENTS

This work was funded by Australian Wool Innovation Ltd.

REFERENCES

Henshall J.M., Hawken R.J., Dominik S. and Barendse W. (2012) *Genet. Sel. Evol*. 44: 12.
Huang W, Kirkpatrick BW, Rosa GJM, Khatib H. (2010) *Anim. Genet*. **41**: 570.
Lee W.C. (2005) *Cancer Epidem. Biomar*. **14**: 958.
MacCluer J.W., Vandeburg J.L., Read B. and Ryder O.A. (1986). *Zoo Biol.* **5**:149.
R Development Core Team (2008) 'R: A language and environment for statistical computing.' R Foundation for Statistical Computing, Vienna, Austria.
Smith J.L., Brewer H.G. and Dyall T. (2009) Proc. Assoc. Advmt. Anim. Breed. Genet. **18**:334.

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:183-186

INHERITANCE OF FLYSTRIKE RECORDED IN A NON-SEASONAL RAINFALL ENVIRONMENT

T.L. Bird-Gardiner[125], D.J. Brown[3], J.L. Smith[4] and S.I. Mortimer[5]

[1]Cooperative Research Centre for Sheep Industry Innovation, UNE, Armidale, NSW, 2350, [2]University of New England, Armidale, NSW, 2351, [3]Animal Genetics and Breeding Unit, UNE, Armidale, NSW, 2350, [4]CSIRO Animal, Food and Health Sciences, Armidale, NSW, 2350, [5]NSW Department of Primary Industries, Trangie, NSW, 2823

SUMMARY

Flystrike records from the Sheep CRC Information Nucleus Flock (IN) were evaluated to assess the influence of sex, birth and rearing type, dam age and sire and dam breed on yearling breech and non-breech strike. Heritability of breech and non-breech strike was estimated under a range of different models, on data comprising records from progeny of three genotypes, namely Merino, Maternal-Merino and Terminal cross, run in a variable non-seasonal rainfall environment. Observed heritability estimates ranged from 0.30 ± 0.10 to 0.43 ± 0.13 for breech strike and 0.16 ± 0.06 to 0.32 ± 0.16 for non-breech strike across genotypes and models. Heritability estimates for both traits on the underlying scale were all high (>0.6), with large standard errors. Flystrike was found to have a similar heritability to estimates obtained in other environments. The identification of key indicator traits for non-seasonal rainfall environments is warranted.

INTRODUCTION

The practice of mulesing to control breech strike in sheep is under scrutiny for social and ethical reasons (James 2006). Flystrike can be controlled through various other methods, such as shearing, crutching, chemical application and breeding for flystrike resistant sheep. Research into breeding breech strike resistant sheep is currently being conducted in 2 environments, which have either a summer (Armidale, NSW; Smith *et al*. 2009) or winter dominant rainfall (Mt. Barker, WA; Greeff and Karlsson 2009). Early results from these studies have shown breech strike to be heritable. While these 2 environments represent a large proportion of sheep production areas in Australia, they do not represent pastoral areas with variable non-seasonal rainfall.

Within the IN Flock of the Cooperative Research Centre for Sheep Industry Innovation (Sheep CRC), parasite resistance of progeny bred within the flock was assessed at all 8 sites (Fogarty *et al.* 2007) and this included flystrike observations on any part of the sheep. Most of the flystrike data collected was within a non-seasonal rainfall environment. The aim of this study was therefore to identify the importance of fixed effects on the expression of yearling breech and non-breech strike in a variable non-seasonal rainfall environment and compare heritability estimates derived from different trait expression models.

MATERIALS AND METHODS

The combination of environmental effects and achieving a balance between collecting information for the parasite program and other IN programs meant that some sites implemented regimes which reduced the expression of flystrike. This is evident in the variation in flystrike across sites (Table 1). Of the 8 IN sites, incidence of flystrike was greatest at the Trangie site and it is these records that were used in this analysis. Progeny were born in 2008 to 2011 inclusive and were not mulesed. Flystrike control at the Trangie site was specific to a contemporary group and included shearing or crutching at weaning and suitable preventative chemical application on the breech and body for progeny evaluated for meat or wool traits respectively (Fogarty *et al.* 2007). Flystrike was recorded on ewes and wethers from marking to shearing (approximately 10 months

of age) for wool evaluation progeny and until slaughter for meat evaluation progeny (5-10 months of age). Progeny genotypes were: Merino, Maternal-Merino and Terminal cross. Sires were only used in one year and dams across years. Sire breeds comprised: 4 Merino strains being fine, medium, strong, generic or strain cross; 5 Maternal (Dohne, South African Meat Merino, Coopworth, Corriedale, Border Leicester) and 4 Terminal (Poll Dorset, Texel, Black Suffolk, White Suffolk). SAMM, Texel and Black Suffolk sired progeny had no expression of flystrike to yearling age and were excluded from all analyses. No breech strike was recorded for White Suffolk progeny to yearling age, so these records were only included in the non-breech strike analysis. Dam breeds were condensed into 4 groups for analysis: fine Merino, medium Merino, strong Merino and crossbred. Two traits were analysed: breech strike and non-breech strike. Non-breech strike was defined as flystrike other than the breech area (Watts *et al*. 1979). Two data sets were analysed within each trait. The full data set (All) contained all Merino, Maternal and Terminal cross breeds, the Merino data set (Mer) only comprised Merinos retained for wool evaluation. The full data set contained 1321 and 1353 animals for the breech and non-breech strike respectively, whereas the Merino data set comprised 580 animals for both traits. The data were analysed as: Strikes, sum of strikes between marking and shearing for each animal; and Struck, presence of flystrike (not struck, 0; struck, 1) for each trait.

Statistical analysis. Data were analysed using ASReml (Gilmour *et al.* 2009) fitting generalised linear mixed models to evaluate fixed effects and sire or animal fitted as a random effect. Sex, year of birth x contemporary (CG; management group within birth year), birth rearing type, dam age, sire and dam breed and their interactions were fitted to each trait and were included in the models if significant (P <0.05). Variance components from the different models were used to estimate heritabilities on the observed scale and heritabilities on the underlying normal distribution scale. Data for Strikes were analysed with a sire and pedigreed animal model on the observed scale. Struck data was analysed using a sire model on the observed scale and on the underlying scale using a logit link function in a sire threshold model. The method described by Hill and Smith (1977) was used to transform heritability estimates.

RESULTS AND DISCUSSION

The incidence of flystrike was greatest at the Trangie site, with 18.2% of total animals affected. The higher incidence of flystrike at the Trangie site was due to optimal fly wave conditions occurring in 2010 and 2011 when excessive (226mm above average) rainfall was recorded.

Table 1. Flystrike incidence recorded in IN flocks, expressed as a percentage of total animals

Site	Animals	Strike Type		Breed		Age	
		Breech (%)	Non-breech (%)	Merino (%)	Non-Merino (%)	0-12m (%)	>12m (%)
Kirby	3325	0.6	0.0	0.8	0.4	0.6	0.0
Trangie	1695	14.5	6.9	36.7	6.9	9.1	12.8
Cowra	2110	3.5	1.8	11.2	1.5	1.6	3.6
Rutherglen	2213	3.6	0.6	11.2	0.8	3.7	0.6
Hamilton	1970	2.4	0.4	5.6	1.0	2.7	0.1
Struan	1980	0.9	0.6	2.7	0.7	0.9	0.6
Turretfield	2476	4.3	1.4	10.1	2.6	3.2	2.4
Katanning	4181	4.9	0.5	8.5	3.2	5.1	0.2
All sites	19950	4.0	1.2	9.4	2.1	3.3	1.9

Fixed effects. The fixed effects were of greater significance in models fitted to Strikes data than those fitted to Struck data. Only fixed effects for Strikes data are reported.
Breech strike. The expression of yearling breech Strikes was influenced (P <0.001) by sex, CG and sire by dam breed interaction for both data sets. Females had a higher incidence of breech Strikes than males. For the Merino data set, progeny born in the 2009 CG had a significantly lower expression of yearling breech Strikes than progeny born in other CG. The sire by dam breed interaction indicated that progeny from the same sire breed had varying levels of breech Strikes depending on the dam breed. The fixed effects fitted to breech Strikes were similar to previous studies (Smith *et al.* 2009).
Non-breech strike. Sex had no effect on the expression of non-breech Strikes, irrespective of breed. Significant fixed effects for the all breed data were CG with a sire by CG interaction (P <0.001) and sire by dam breed interaction (P <0.05). Only sire breed by CG (P <0.001) influenced non-breech Strikes in the Merino data.

Heritability estimates. The various models produced observed heritability estimates of similar magnitude across genotypes for breech Strikes 0.32 ±0.08 to 0.43 ± 0.13, but varied more for non-breech Strikes 0.16 ± 0.06 to 0.30 ± 0.15 (Table 2). Despite the lower incidence of non-breech Strikes compared to breech Strikes the standard errors were similar for both traits on the observed scale. Merino breech Strikes heritability estimates were similar to those reported by Smith *et al*. (2009) (0.32 ± 0.11 on the observed scale) and Greeff *et al.* (2013) (0.58 ± 0.16 on the underlying scale). The all breeds heritability estimates for Struck derived on the underlying scale were high for both traits, with large standard errors. Observed heritability estimates transformed to the underlying scale were higher than those derived from the logit link function model and some were above one (not reported). The scaling factor in the model described by Hill and Smith (1977) can

Table 2. Phenotypic variance (σ^2p) and heritability (±s.e.) estimates for yearling breech and non-breech flystrike traits; all breeds (All), Merinos (Mer), flystrike incidence (%)

Traits and Model	Incidence %	Breed	Observed scale		Transformed	Underlying scale
			σ^2p	h^2	h^2	h^2
Breech						
Strikes - Sire	8.6	All	0.11	0.42 ± 0.12	-	
Strikes - Animal	8.6	All	0.11	0.32 ± 0.08	-	
Struck - Sire	8.6	All	0.07	0.30 ± 0.10	0.95 ± 0.32	
Struck - Sire threshold	8.6	All	4.10			0.79 ± 0.26
Strikes - Sire	14.0	Mer	0.17	0.42 ± 0.17	-	
Strikes - Animal	14.0	Mer	0.16	0.43 ± 0.13	-	
Struck - Sire	14.0	Mer	0.11	0.33 ± 0.15	0.80 ± 0.36	
Struck - Sire threshold	14.0	Mer	3.85			0.58 ± 0.27
Non-breech						
Strikes - Sire	3.7	All	0.06	0.19 ± 0.08	-	
Strikes - Animal	3.7	All	0.06	0.16 ± 0.06	0.87 ± 0.33	
Struck - Sire	3.7	All	0.03	0.22 ± 0.09	-	
Struck - Sire threshold	3.7	All	3.98			0.69 ± 0.40
Strikes - Sire	4.3	Mer	0.07	0.30 ± 0.15	-	
Strikes - Animal	4.3	Mer	0.07	0.18 ± 0.10	0.88 ± 0.49	
Struck - Sire	4.3	Mer	0.04	0.32 ± 0.16	-	
Struck - Sire threshold	4.3	Mer	4.16			0.83 ± 0.53

produce inflated estimates when incidence levels within the population are less than 30% (Gilmour *et al.* 1985), as was the case in this study for both traits. Merino heritability estimates for non-breech Struck on the observed scale were similar to estimates reported by Raadsma *et al.* (1989) of 0.27 ± 0.12, but these estimates were only for body strike. Standard errors obtained on the underlying scale were larger than other published work but were similar on the observed scale (Greeff and Karlsson 2009; Smith *et al.* 2009). The size and structure of the data and incidence of flystrike would have contributed to the large standard errors obtained (van der Werf *et al.* 2010).

Heritability of binomial traits such as flystrike (when observed as absence or presence of flystrike) is related to the incidence in the population (Atkins 1979), making prediction to selection response difficult when the expression varies from year to year. Using the underlying scale of susceptibility to flystrike, the mean and heritability is not dependent on the incidence within the population. The underlying scale maps flystrike susceptibility on a continuous normal distribution scale which is expressed on the observed scale once a certain threshold is reached. Results from this study and other research indicate that susceptibility to breech strike on the underlying scale is moderately heritable and comparable to other heritability estimates derived for traits which have a continuous distribution of phenotypes such as greasy fleece weight, indicating potential for improvement through selection. But flystrike is highly dependent on environmental conditions making it difficult to include in routine selection practices. The practice of indirect selection is used to improve traits that are difficult to measure such as susceptibility to flystrike. Therefore it is important to identify easy to measure correlated traits that are associated with the expression of flystrike, which can be incorporated into breeding programs. Identifying key correlated traits (to allow selection when flystrike is not expressed) would provide selection tools to breed for flystrike resistant sheep. As research has identified that some key indicator traits are environment specific (Greeff *et al.* 2013; Smith *et al.* 2009), it is important to establish indicator traits for non-seasonal rainfall environments.

ACKNOWLEDGEMENTS

Funding for this project was provided by the Cooperative Research Centre for Sheep Industry Innovation. Drs Arthur Gilmour and Gavin Melville are gratefully acknowledged for their statistical advice.

REFERENCES

Atkins K.D. (1979) National symposium on the sheep blowfly and flystrike in sheep. Department of Agriculture, New South Wales.

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.

Gilmour A.R., Anderson R.D. and Rae A.L. (1985). *Biometrika* **72:** 593.

Gilmour A.R., Gogel B.J., Cullis B.R. and Thompson R. (2009) 'ASReml User Guide release 3.0' VSN International Ltd: Hemel Hempstead, UK.

Greeff J.C. and Karlsson L.J.E. (2009) *Proc. Assoc. Advmt. Anim. Breed. Genet.***18:** 272.

Greeff J.C., Karlsson L.J.E. and and Schlink A.C. (2013) *Anim. Prod. Sci.* http://dx.doi.org/10.1071/AN12233

Hill W.G. and Smith C. (1977) *Biometrics* **33:** 234.

James P.J. (2006) *Aust. J. Exp. Agric.* **46:**1.

Raadsma H.W., Gilmour A.R., and Paxton W.J. (1989) *Aust. J. Agric. Res.* **40:** 207.

Smith J.L., Brewer H.G. and Dyall T. (2009) *Proc. Assoc. Advmt. Anim. Breed. Genet.* **18:** 334.

Van der Werf J.H.J., Kinghorn B.P. and Banks R.G. (2010) *Anim. Prod. Sci.* **50:** 998.

Watts J.E., Murray M.D. and Graham N.P.H. (1979) *Aust. Vet. J.* **55:** 325.

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:187-190

THE EFFECT OF BREED, EWE AGE AND SEASON ON TICK COUNTS OF INDIGENOUS AND COMMERCIAL SHEEP IN SOUTH AFRICA

J.J.E. Cloete[1,2], S.W.P. Cloete[1,3], A.J. Scholtz[3] and S. Matthee[4]

[1]Department of Animal Sciences, University of Stellenbosch, Matieland 7602, South Africa
[2]Elsenburg Agricultural Training Institute, Elsenburg 7607, South Africa
[3]Directorate: Animal Sciences: Elsenburg Elsenburg 7607, South Africa
[4]Department of Conservation Ecology, University of Stellenbosch, Matieland 7602, South Africa

SUMMARY

Repeated udder health and tick counts were recorded on ewes belonging to the indigenous Namaqua Afrikaner (NA) fat-tailed breed, as well as the commercial Dorper and SA Mutton Merino (SAMM) breeds. Udders were scored subjectively on a 1-5 scale and ticks were counted on three locations. Udder score (US) increased (i.e. became worse) with age from 2 to 6+ years, an effect that was accentuated in the commercial ewes compared to the NA. NA ewes generally had lower tick counts than the commercial breeds on their front (FTC) and hind (HTC) parts, but had more ticks on the breech, perineum and tail (BPTTC) than the Dorper. Repeatability estimates amounted to 0.75±0.03 for US, 0.19±0.05 for FTC, 0.58±0.04 for HTC and 0.24±0.05 for BPTTC. Significant correlations between animal effects amounted to 0.47±0.07 between US and HTC, and to 0.58±0.04 between FTC and HTC. The results suggest an advantage in favour of the indigenous NA breed for udder health and body tick infestation compared to the commercial breeds.

INTRODUCTION

Sheep form an integral component of most livestock production systems in South Africa, particularly in the arid, pastoral regions. The various sheep breeds are able to survive and to produce in a wide range of ecotypes and in many cases can exploit the scarce feed resources available. However, ectoparasites such as ticks are considered to be of veterinary and economic importance (Fourie *et al*. 1988). Some tick species transmit diseases (Howell *et al*. 1978), while others, because of their long mouthparts and tendency to form clusters, cause severe tissue damage (MacIvor and Horak 1987) and necrosis (Howell *et al*. 1978). In addition, certain species transmit toxins that cause paralysis (Fourie *et al*. 1989), while others cause tissue damage in feet giving rise to foot abscesses (MacIvor and Horak 1987). Extreme cases of blood loss can also drain the nutrients and "tick worry" can irritate animals, resulting in lower production. This paper reports on tick burdens and udder damage on mature ewes from three South African breeds, in the absence of literature on this topic. It is assumed that resilience to tick infestation will provide an indication of hardiness under extensive, free ranging conditions.

MATERIALS AND METHODS

The study was conducted at Nortier research farm near Lamberts Bay in the Western Cape Province of South Africa. The farm is situated on the Western seaboard of the country where winter rainfall occurs. The location expects a long-term annual precipitation of 220 mm per annum, 78% of which is recorded between April and September. The experimental animals grazed natural shrub pasture typical of the region.

A total of 635 repeated udder health and tick count records were available on reproducing ewes belonging to the indigenous Namaqua Afrikaner (NA; n=275) fat-tailed breed, as well as to the commercial Dorper (n=366) and South African Mutton Merino (SAMM; n=94) breeds. The animals were examined and ticks counted and removed during austral summer (December 2011; n=255), autumn (May 2012; n=188) and spring (September 2012; n=192). Individual ewes were

upended and recordings on ewes included the following: Subjective udder damage score (US; 1-5), where a complete smooth and healthy udder with undamaged teats was awarded a score of one. A severely damaged udder with scar tissue and misformed/malformed teats due to tick damage, validating the culling of the individual, was awarded a score of five. Provision was made for half marks in cases where US was situated between two categories. The whole body was divided into three areas, namely the front part (FTC; including head, ears and front legs up to the navel), the hind part (HTC; posterior of the navel, including the udder, thighs, hind legs and feet) as well as the breech and perineum area (BPTTC; including the tail in the NA). Care was taken to ensure that ticks were counted on the tip as well as in the twist of the tail of NA ewes. Tick count data were transformed to square roots to stabilize variances. Random ewes within breeds were identified to have all ticks removed from one side of the body for identification during each sampling session.

Repeated records of each trait on the same animal at different times of the year were accounted for by fitting a repeatability model to the data, using ASREML (Gilmour *et al*. 2006). It was not attempted to partition the between animal variance in direct genetic and animal permanent environmental effects given the small size of the data set. Repeatability was estimated by expressing the between animal variance component as a ratio of the phenotypic variation, after the known fixed effects have been accounted for. The fixed effects considered were the breed of animal (NA, Dorper or SAMM), age (2 years, 3-5 years or 6+ years) and time of the year (Desember 2011, May 2012 or September 2012). Interactions between fixed effects were also considered and reported where they occurred. Geometric means and appropriate standard errors pertaining to significant fixed effects were predicted in ASREML (Gilmour *et al*. 2006). Initially single-trait analyses were fitted to each trait, to obtain operational models. Subsequent analyses involved the fitting of two- and three-trait models to obtain correlations between animal effects (hereafter referred to as between-animal correlations), as well as phenotypic correlations among traits.

RESULTS AND DISCUSSION

A total of 3980 ticks (including males, females, nymphs and larvae) were recovered and identified from 73 ewes. Of these ticks, 2001 (50.3%) belonged to the species *Rhipicephalus evertsi evertsi*, 1051 (26.4%) to the species *R. gertrudi* and 890 (22.4%) to the species *Hyalomma truncatum*. The remaining 0.9% of ticks consisted of four species (*H. rufipes*, *H. truncatum parma*, *H. glabrium* and *R.* g*labroscutatum*) that were of minor importance. These figures give an indication of the species that were involved. Detailed information on the site of attachment and the species distribution across season, breed, gender and age class falls beyond the scope of this paper.

US was independent of the month of recording ($P>0.10$), but were affected by the interaction between breed and age (Table 1). US did not differ appreciably between breeds in two-tooth maiden ewes, but deteriorated with age ($P<0.01$) (i.e. became higher). The rate of deterioration was markedly faster in Dorper and particularly SAMM ewes when compared to NA ewes. Ewes of 6+ years had an average US of 1.42±0.18 for the NA, 2.21±0.10 for the Dorper and 3.03±0.19 for the SAMM (All $P<0.05$). No other information on the three sheep breeds pertaining to the impact of tick infestation on udder health could be sourced.

FTC was dependent upon an interaction between breed and month of recording. Although higher in SAMM ewes ($P<0.05$), average geometric means for FTC during December 2011 were below one in all cases (Table 2). FTC differed markedly ($P<0.01$) between breeds during May 2012, with counts for SAMM ewes being more than twofold that of Dorpers. FTC in the latter breed was also approximately double those in NA ewes. FTC in September 2012 was lower ($P<0.05$) again while the breeds were re-ranked to an extent. Counts were highest in the Dorper, followed by the SAMM, while FTC in the NA remained the lowest. FTC was independent of ewe age ($P=0.38$).

Table 1. Least squares means (± s.e.) depicting the interaction of breed with ewe age for udder score (1 –smooth and healthy; 5 – severely damaged, validating the culling of ewe)

Effect		Breed	
Ewe age	NA	Dorper	SAMM
2 years	$1.01\pm0.16^{a\,1}$	$1.39\pm0.13^{b\,1}$	$1.29\pm0.16^{a,b\,1}$
3-5 years	$1.14\pm0.12^{a\,1,2}$	$1.57\pm0.08^{b\,1}$	$2.26\pm0.16^{c\,2}$
6+ years	$1.41\pm0.18^{a\,2}$	$2.21\pm0.10^{b\,2}$	$3.02\pm0.19^{c\,3}$

^{a,b,c} – significant ($P < 0.05$) differences in rows ^{1,2,3} – significant ($P < 0.05$) differences in columns

HTC increased with ewe age, geometric means being 2.90±0.59 for two-tooth ewes, 5.73±0.63 for 3-5 year old ewes and 10.1±1.15 for ewes of 6+ years. HTC was affected by an interaction between breed and month of recording ($P<0.05$; Table 2). It was evident that HTC in Dorper ewes increased roughly linearly from December 2011 to September 2012 ($P<0.05$). No differences between months were evident for SAMM ewes. No differences were found between Dorper and SAMM ewes ($P>0.05$). HTC of NA ewes were consistently below half of the other breeds ($P<0.05$). BPTTC was independent of ewe age ($P>0.50$), but was affected by an interaction between breed and month of recording ($P<0.05$; Table 2). Dorper ewes consistently had a lower BPTTC than the other breeds, although the advantage in favour of Dorper ewes differed in magnitude between months. It needs to be stated that only the NA breed had tails, with the tails of the other breeds being docked.

Table 2. Geometric means (± approximate standard errors) derived from square-root transformed data and depicting the interaction of breed with month of recording for tick counts on the front part, the hind part as well as on the breech, perineum and tail of ewes

Effect		Breed	
Month of record	NA	Dorper	SAMM
		Front part	
December 2011	$0.07\pm0.07^{a\,1}$	$0.23\pm0.08^{a\,1}$	$0.82\pm0.33^{b\,1}$
May2012	$2.78\pm0.48^{a\,3}$	$5.92\pm0.51^{b\,3}$	$17.55\pm1.73^{c\,2}$
September 2012	$0.44\pm0.20^{a\,2}$	$2.46\pm0.33^{c\,2}$	$1.52\pm0.44^{b\,1}$
		Hind part	
December 2011	$1.61\pm0.47^{a\,1}$	$6.34\pm0.61^{b\,1}$	$6.96\pm1.23^{b\,1}$
May2012	$3.75\pm0.76^{a\,2}$	$8.26\pm0.79^{b\,2}$	$8.35\pm1.50^{b\,1}$
September 2012	$2.49\pm0.64^{a\,1,2}$	$10.22\pm0.88^{b\,3}$	$8.62\pm1.38^{b\,1}$
		Breech, perineum and tail	
December 2011	$11.49\pm0.94^{b\,3}$	$3.76\pm0.36^{a\,2}$	$10.51\pm1.22^{b\,2}$
May2012	$6.45\pm0.77^{b\,1}$	$2.88\pm0.37^{a\,1}$	$5.64\pm1.02^{b\,1}$
September 2012	$9.07\pm0.95^{b\,2}$	$4.49\pm0.47^{a\,2}$	$7.45\pm1.03^{b\,1}$

^{a,b,c} – significant ($P < 0.05$) differences in rows ^{1,2,3} – significant ($P < 0.05$) differences in columns

The interaction of breed with month of assessment pertaining to tick counts may be associated with the ecology of the ticks, as the abundance of different tick species are known to depend on the season in the dominant species at the research site (Fourie *et al.* 1988; Fourie and Horak 1991). However, further studies on the species distribution of the ticks that were removed from the experimental animals are needed to gain a better understanding of mechanisms involved.

Single-trait repeatability estimates for the respective traits were all within 0.01 of the repeatability estimates derived from a series of two-trait and three-trait analyses involving all

possible trait combinations. Results for the two-trait analyses are thus presented in Table 3. All traits were repeatable with all estimates exceeding a level of twice the corresponding standard error. At 0.75, US were highly repeatable. Repeatability coefficients for FTC were lower at 0.19. HTC were also highly repeatable (0.58), with an estimate of 0.24 for BPTTC. Between-animal correlations were significant between US and HTC (0.47±0.07) and between FTC and HTC (0.58±0.04). Only one reference was found where ticks were implicated in ovine udder damage (Fourie *et al*. 2001).

Table 3. Repeatability estimates for udder score (US) as well as tick counts on the front part (FTC), hind part (HTC), as well as on the breech, perineum and tail (BPTTC) of the animals assessed in bold italics on the diagonal. Between-animal correlations are provided above the diagonal and phenotypic correlations below the diagonal

Trait	US	FTC	HTC	BPTTC
US	***0.75±0.03***	0.01±0.13	0.47±0.07	-0.13±0.11
FTC	-0.03±0.05	***0.19±0.05***	0.58±0.04	-0.00±0.18
HTC	0.40±0.05	0.15±0.04	***0.58±0.04***	-0.18±0.12
BPTTC	-0.07±0.05	0.02±0.04	-0.10±0.05	***0.24±0.05***

Significant repeatability estimates augmented with breed differences in tick counts indicate a heritable component for resistance to tick infestation. In contrast with the present results, Fourie and Kok (1996) found that Merino sheep had a lower *Ixodes rubicundus* burden than Dorpers when grazing the same Karoo shrub veld. They suggested this difference may be related to different grazing patterns between the two breeds. MacLeod (1932) also reported that Cheviot ewes were more resistant than Blackface ewes to infestation with female ticks. Further studies are needed to more fully characterize the genetic basis of tick resistance in South African sheep.

CONCLUSIONS

The repeatable nature of tick counts and observed breed differences indicate a genetic basis for tick tolerance in sheep. The advantages in favour of the indigenous NA breed for udder health and hind tick infestation indicates that the NA would be more robust than the other two breeds on natural pasture where there are high tick burdens. However, this must be balanced against any breed differences for important performance traits that have not been considered here.

REFERENCES

Fourie L.J. and Horak I.G. (1991) *J. S. Afr. Vet. Assoc.* **62**: 104.
Fourie L.J., Horak I.G. and Marais L. (1988) *J. S. Afr. Vet. Assoc.* **59**: 191.
Fourie L.J. and Kok D.J. (1996) *Onderstepoort J. Vet. Sci.* **63**: 273.
Fourie L.J., Kok D.J. and Peter R.J. (2001) *J. S. Afr. Vet. Assoc.* **72**: 143.
Fourie L.J., Petney T.N. and Horak I.G. (1989) *Vet. Parasitol.* **33**: 319.
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)
Howell C.J., Walker J.B. and Nevill E.M. (1978) *Dept Agric. Tech. Serv. RSA Sci. Bull.* **393**: 69.
MacIvor K.M.De F. and Horak I.G. (1987) *J. S. Afr. Vet. Assoc.* **58**: 113.
MacLeod J. (1932) *Parasitology* **24**: 382.

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:191-194

IMPROVING FACIAL ECZEMA TOLERANCE IN NEW ZEALAND DAIRY CATTLE

N. G. Cullen[1], P. R. Beatson[2], R. T. Courtney[2] and C. A. Morris[3]

[1]AgResearch, Ruakura Research Centre, PB 3123, Hamilton, New Zealand
[2]CRV Ambreed, PO Box 175, Hamilton, New Zealand
[3]AgResearch (retired)

SUMMARY

After implementing genomic selection for milk and production traits in dairy cattle, a New Zealand breeding company (CRV Ambreed) has been encouraged to expose a selection of their young bulls to a challenge with the toxin responsible for causing the facial eczema (FE) disease. Resistance to FE is heritable in dairy cattle, and breeding values (BVs) for it are freely available. For young bulls, BVs are based on parent-average predictions, but there is now also the option of phenotyping. This process can identify the more tolerant bulls which can be used by those farmers who suffer major production and cow losses on their farms as a result of this disease. In susceptible animals, FE causes severe liver damage, and may also cause severe and painful skin damage resulting from photosensitisation. This paper describes a protocol which, by removing sources of chlorophyll from the diet, allows animals to develop a sufficient range of liver damage to rank bulls for FE tolerance but does not cause animals to present with photosensitisation.

INTRODUCTION

Facial Eczema (FE) is caused by the toxic spores of the fungus *Pithomyces chartarum.* The toxin (sporidesmin) is produced by spores which are released during warm, humid weather over mid-summer until late-autumn (December – May). It occurs particularly in ryegrass/white clover pastures in almost all the North Island and the upper South Island of New Zealand (NZ). While FE is of major concern in NZ, the same disease is reported elsewhere in regions of similar latitude (both North and South) and particularly in coastal southern regions of Australia. Di Menna *et al*. (2009) have reviewed FE research in NZ since 1939.

Damage to the liver and bile ducts was identified in early research as the most important pathology, although sporidesmin can affect many organs. As a result of bile duct blockage, a breakdown product of chlorophyll, phylloerythrin, can build up in the body causing sensitivity to sunlight and visible skin damage (hence the common name of FE). Chronic wasting and/or death may occur at the time of damage or months later when the animal is under stress, e.g. parturition.

Early research in sheep showed that gamma-glutamyltransferase (GGT) levels in blood measured 2-3 weeks after dosing with the toxin are positively correlated to post-mortem liver damage scores and to losses in live weight (Towers and Stratton 1978). GGT levels (with a natural log transformation) are now used as the proxy for measuring responses to the toxin, in animals not suspected of having suffered any other liver-damaging process.

The first research in dairy cattle involved the daughters of young progeny test sires subjected to a natural challenge in 1989. This demonstrated that FE was also heritable in dairy cattle, as in sheep, and was subsequently followed by limited evaluation of dairy sires by both direct performance testing and/or progeny testing. However, due to the costs and risks involved (possible losses of potentially very valuable bulls), this work ceased after a short period. Morris *et al.* (2013) provide a review of the genetic work performed in both sheep and cattle since the early 1980s.

Sampling of all cows from clinically-affected dairy herds in the upper North Island from 2004-11 provided GGT and pedigree data on ~15,000 cows from 70 herds. These data plus pedigree information on sires allowed AgResearch to update the heritability estimates and provide FE breeding values (BV) for all animals including sires. The current estimate of the heritability for

$\log_e$ GGT is 0.34 ± 0.02 (Cullen *et al*., 2011).

With the introduction of genomic selection, it is now possible to short-circuit the process of identifying elite sires for milk production instead of using the traditional sire-proving approach which takes a minimum of 5 years. Genomic selection allows the bulls that will progress to progeny-testing to be identified by 6 months of age. The remaining group of bulls with reduced genomic predictions is then challenged with sporidesmin to identify the most resistant bulls amongst this group. Farmers could then compare an FE-tolerant bull with its breeding worth (BW – the NZ national breeding objective for dairy cattle) and make informed decisions about bull selection for breeding. A joint bid by two farmers (Mr and Mrs Burt) and AgResearch to DairyNZ's 'On Farm Innovation Fund' in 2010 was successful in obtaining funds to test the concept of sequential selection of dairy bulls (BW prediction, then FE tests). CRV Ambreed Ltd. (CRV), given their past associations in this work, was approached to be a partner in this process.

A protocol was devised for FE testing in the dairy bull industry, and evaluated here. It allows artificial challenge of young bulls with the FE toxin, in order to rank them for tolerance without causing photosensitisation and its severe effect on animal welfare.

METHODS

New Testing Protocol. Animal ethics approval was obtained from the Ruakura Animal Ethics Committee, and a pilot trial was run at CRV's facility near Hamilton in November 2010. This tested a new protocol with an artificial sporidesmin challenge, using a dose rate of 0.25 mg sporidesmin (in suspension in water) per kg of bull live weight, to be administered by stomach tube. Eleven surplus 2009-born bulls (6 Jersey and 5 Holstein-Friesian) were challenged with sporidesmin in this manner at 16 months of age. The agreed protocol was that bulls would be housed in stalls with 24-hour access to shade and fed solely with silage and supplements. The premise was that, if sources of chlorophyll were removed from the diet, phylloerythrin levels would be kept low and thus not cause animals to develop photosensitivity as a result of liver and bile duct damage. GGT levels were measured on these 11 bulls before dosing, to ensure there was no existing damage. The initial sporidesmin dose rate chosen was not sufficiently high to identify the targeted 10% FE-tolerant bulls, so 5 of the non-responders (one bull was removed because of poor semen production) were re-dosed at a higher rate (0.30 mg/kg) 6 weeks later. GGT levels for all bulls were measured at 14, 21 and 28 or 35 days post-dosing, and then every 14 days for those bulls with levels remaining above 200 iu/l. All animals were inspected daily by CRV's veterinary and farm staff for signs of ill health. Live weight was measured 3-5 days before dosing to prepare individual sporidesmin doses and again when bulls were sampled for GGT at day 21 post-dosing.

This protocol and dose rate was to be the basis of future work with ~50 young bulls per year, starting with 'Year 1' in 2011. Although the initial aim was to test non-elite bulls (on genomic BW), CRV were confident enough with the results of this Pilot Trial that 90% of bulls in Year 1 were elite animals. In 'Year 2', 2012, all bulls were elite animals. The risk was reduced by the ability to pre-screen bulls by predicting parent-average FE BVs.

Year 1. In March 2011, after 51 8-month old bulls (2010-born) were pre-selected on the basis of having the most negative (favourable) parent-average BVs, and the ethical exclusion of 6 bulls because of evidence of a previous 'natural' FE challenge, the main trial began. Forty-five young bulls were challenged with a sporidesmin dose of 0.30 mg/kg at the same CRV facility and using the same feeding regimes as previously.

Year 2. In March 2012, the programme continued with another 50 8-month old bulls (2011-born), with the same pre-screening procedure and the same protocol. The sporidesmin dose rate was reduced slightly to 0.27 mg/kg to reduce potential risks to animals from unique blood-lines which were included in the group.

RESULTS AND DISCUSSION

Pilot Trial. The normal range for GGT is 0-40 iu/l and levels >5000 iu/l have been observed in FE-affected cows on dairy farms. The maximum GGT level attained in the initial dosing was 853 iu/l with 5/11 bulls having levels greater than 40 iu/l. Morris *et al*. (2009) observed that elevated phylloerythrin levels were only detected when GGT levels were greater than ~600 iu/l and not all animals above this threshold exhibited clinical signs of FE. Only one bull surpassed this threshold so it was not possible to draw any conclusions about the protocol protecting animals from severe FE. For the 5 non-responders redosed at 0.30 mg/kg, the highest GGT level was 188 iu/l. No photosensitisation was observed.

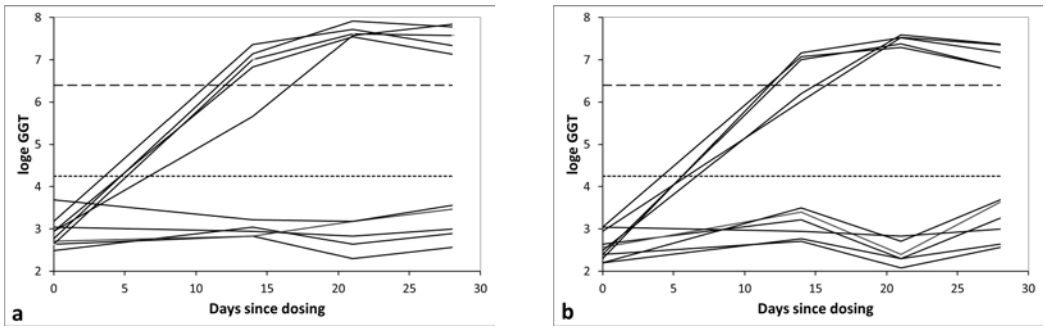


Figure 1: $\log_e$ GGT levels recorded over the 28 days from dosing for the 5 bulls with the lowest and highest GGT levels at 21 days post-dosing for both the a) 2011 and b) 2012 dosing rounds (dotted line is equivalent to a GGT value of 70 iu/l, a level below which any liver damage may be minimal; and the dashed line to 600 iu/l – the level above which clinical signs of FE due to photosensitisation might be observed).

Year 1. 38 of 45 of bulls had elevated GGT levels (maximum 2736 iu/l). Figure 1a shows the GGT time series for each of 5 bulls having the lowest and highest GGT levels at 21 days post-dosing. The trends for the remaining 31 bulls with intermediate values for GGT are not shown but follow similarly; the repeatability for $\log_e$ GGT in serially sampled animals is 0.86 ± 0.004 (Cullen *et al*. 2011). The repeatability over days 14, 21 and 28 for these data were 0.85. The high GGT levels were in the range where it was expected that some bulls would show clinical FE signs due to photosensitisation (but none observed provided support for the chosen protocol).

Year 2. The second year of dosing 8-month old bulls showed similar results (Figure 1b) to Year 1; 43 of 50 bulls had elevated GGT levels. No photosensitisation was observed.

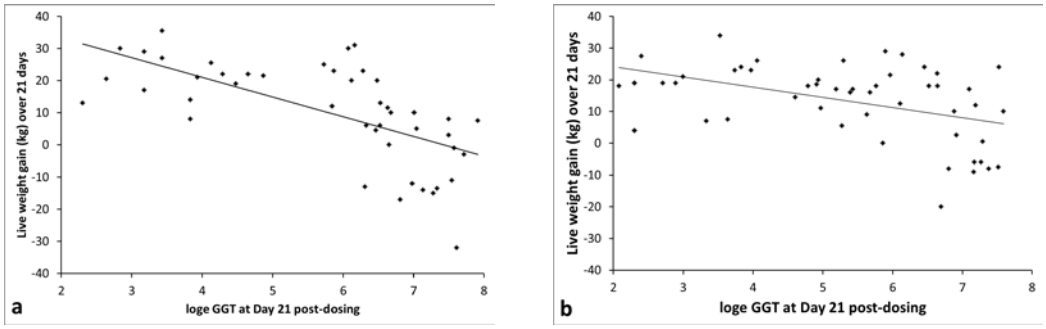


Figure 2: Live weight gain plotted against $\log_e$ GGT levels at 21 days post-dosing for the a) 2011 and b) 2012 dosing rounds.

Although no clinical signs of FE were observed in the 106 bulls challenged over the 3 batches of animals, it was evident that the liver damage, as measured by increased GGT levels did have a dramatic effect on live weight. This is probably due to the suppression of appetite through the bull feeling unwell. This was more evident in 2011 with the slightly higher dose rate. The regression estimates of live weight gain/loss on $\log_e$ GGT at 21 days post-dosing were -6.1 ($P < 0.0001$) and -3.2 ($P < 0.01$) kg/$\log_e$ GGT unit for 2011 and 2012 respectively.

The data and pedigrees previously collected from the on-farm sampling of dairy cows has allowed the calculation of reliable (> 60%) breeding values for $\log_e$ GGT for approximately 200 industry sires. Incorporating the results from this work into the data has provided the comparative ranking of sires from this study (reliabilities ranging from 36 to 49%) to be compared to other sires which have been used widely in industry. CRV are now marketing teams of both Jersey and Holstein-Friesian FE-Tolerant bulls to industry. The FE BVs for the 13 bulls in CRV's 2013 FE team (born in 2009-11) along with the ~200 bulls with FE reliabilities greater than 60% have been plotted against their birth-year in Figure 3. The team average BV $\log_e$ GGT of -0.24 is superior to 87% of the widely-used industry sires born since the 1980s.

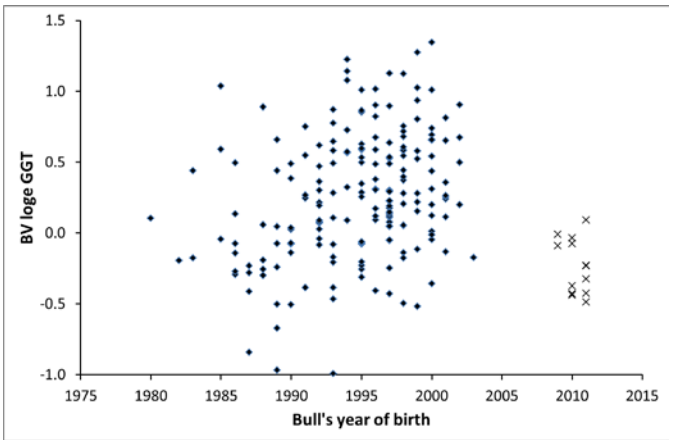


Figure 3: Breeding values for $\log_e$ GGT for industry sires (reliability > 0.60) (symbol = ♦) and bulls in CRV Ambreed's 2013 FE Team (symbol = ×).

This method of genetic improvement of a disease trait like FE is long and slow, as demonstrated by the sheep industry progress with FE over 25 years. It is anticipated that genomic selection will play a role to expedite more rapid gains in the future.

ACKNOWLEDGEMENTS

The initial funding to commence this work was obtained by farmers IR and EA Burt of Matata from DairyNZ's On-Farm Innovation Fund in 2010. This work now continues as a commercial service provided by AgResearch.

REFERENCES

Cullen N.G., Morris C.A., Hickey S.M. and Henderson H.V. (2011) *Proc. NZ Soc. Anim. Prod.* **71**: 117.

Di Menna M.E., Smith B.L. and Miles C.O. (2009) *NZ J. Agric. Res.* **52**: 345.

Morris C.A., Hickey S.M. and Phua S.H. (2009) *Proc. NZ Soc. Anim. Prod.* **69**: 118.

Morris C.A., Phua S.H., Cullen N.G. and Towers N.R. (2013) *NZ J. Agric. Res.* **56**: 170.

Towers N.R. and Stratton G.C. (1978) *NZ Vet. J.* **26**: 109.

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:195-198

WHOLE GENOME ASSOCIATION ANALYSIS OF SUSCEPTIBILITY TO PARATUBERCULOSIS IN NEW ZEALAND DAIRY CATTLE

R.G. Sherlock[1], S. Loker[1], P.J. Back[2], H. Voges[1] and R.J. Spelman[1]

[1]Livestock Improvement Corporation, Hamilton 3284, New Zealand
[2]IVABS, Massey University, Palmerston North 4442, New Zealand

SUMMARY

Johne's disease (JD) is a chronic, inflammatory gastrointestinal disease caused by *Mycobacterium avium* ssp. *paratuberculosis* that can be difficult to detect and control. The herd-level prevalence of JD in the New Zealand (NZ) dairy industry is high, making it a costly disease due to reduced milk production and premature culling. The objective of this study was to identify single nucleotide polymorphisms (SNP), associated with susceptibility to JD in NZ dairy cattle, that could be included in a selection programme. Dairy herds with high levels of JD were identified by testing vat milk. From these herds, individual cows were confirmed as being affected (JD+) by ELISA of milk and serum samples. Approximately 1,800 JD+ cows were genotyped using the Illumina SNP770 Bead Chip. Control genotypes were sourced from a pool of 23,000 cow genotypes (Illumina SNP50 Bead Chip) representing the general population. Approximately 6,800 Control cows were chosen to match the breed proportion profile of the JD+ cows and their genotypes were imputed to 770k for analysis. The association between SNP and JD status was determined using a multi-SNP (Bayes B ($\pi = 0.99$)) approach. Results suggest several regions of interest across the genome, which are moderately associated with the susceptibility of NZ dairy cattle to JD.

INTRODUCTION

Johne's disease (JD) is a chronic, inflammatory gastrointestinal disease, particularly affecting cattle, sheep and deer (Purdie *et al*. 2011). Johne's disease is characterized by lesions in the distal part of the ileum, hindering nutrient uptake (van Hulzen *et al*. 2012), resulting in chronic diarrhea, emaciation, decreased milk production, and eventually death (Gonda *et al*. 2006). The causative agent of JD is *Mycobacterium avium* ssp. *paratuberculosis* (MAP) (Gonda *et al*. 2006; Pant *et al*. 2010; Purdie *et al*. 2011), an infectious bacterium that can be spread through faecal shedding, and can persist in the environment for many months. Commercial vaccinations are available, although they tend to delay the onset of clinical signs rather than prevent the disease, and these vaccinations cause a false-positive reaction to tuberculosis tests.

Apparent herd prevalence has been reported to range between 7 % in Austria to 60 % in New Zealand (Grant 2005). The presence of JD in the national herd represents a large economic loss to the dairy industry, mainly due to reduced milk production and premature culling (Ott *et al*. 1999).

Previous studies have demonstrated the presence of genetic variation for susceptibility to JD (Gonda *et al*. 2006; Attalla *et al*. 2010) and some have identified genomic regions associated with increased susceptibility to JD (Kirkpatrick *et al.* 2010; Minozzi *et al.* 2012; van Hulzen et al. 2012). The genomic signals identified vary between these previous studies, however these are in different populations (both genetics/breed and production systems).

The objective of this study was to identify genetic markers associated with susceptibility for JD in the NZ dairy cow population. Markers identified by this study could form part of a predictive test for potential incorporation into NZ dairy breeding schemes.

METHODS

Cows identified as Johne's disease positive (JD+) from within the NZ dairy population were compared to a Control group representing the general population to identify genomic regions associated with susceptibility to JD.

Johne's disease diagnosis. Diagnostic testing of milk and blood samples employed an enzyme-linked immunosorbent assay (ELISA) marketed as the IDEXX Paratuberculosis Screening Ab Test (www.idexx.com).

Herds were initially prioritised for individual cow screening by ELISA on bulk milk samples. Subsequently, routine herd test milk samples from individual cows in these herds were tested by ELISA used to identify potential JD+ case cows. A blood plasma sample was collected from milk reactor cows to confirm the ELISA positive status. The ELISA sample to positive control optical density ratio thresholds were set at 0.4 and 0.7 for milk and plasma respectively, as per kit instructions prior to 2010. Only cows testing positive on milk as well as plasma ELISA were classified as JD+.

Genotypes. DNA for genotyping was extracted from the blood samples that were used to confirm cows as JD+. Genotyping was performed with the Illumina Bovine SNP770 Bead Chip and resulted in 1833 valid JD+ genotypes with a sample call rate of 95% or greater.

Genotypes from 23,097 cows, representing the general NZ dairy cow population, were made available to the study and formed the Control group following the approach taken by the Wellcome Trust Case Control Consortium (2007). Genotypes for Control cows were obtained using the Illumina Bovine SNP50 Bead Chip and were imputed to the 770,000 SNP using Beagle v3.3.2 (Browning and Browning 2009). SNP with a minor allele frequency of less than 1%, an imputation R^2 of less than 90% in the reference, or with poor clustering characteristics were removed from the analysis. In addition, any SNP common to both the SNP50 and SNP770 Bead Chips were removed to minimize the effects of between-panel differences on the analysis. The remaining 626,033 SNP were included in subsequent analyses.

Analysis. To reduce breed stratification, JD+ cows were grouped into 10 Holstein-Friesian/Jersey breed classes. Control cows from these same classes were chosen at random to generate a matched control of 6,849 cows. The total number of animals in the matched control was determined by the number of Control cows available in the limiting breed class. A multi-SNP, genome-wide association study (GWAS) using the Bayes B method ($\pi = 0.99$) (Meuwissen *et al.* 2001) was performed using the software GenSel v4.53R (Fernando and Garrick 2008). Year of birth, and proportions of Jersey, Holstein-Friesian and overseas' genetics were fitted as covariates. A total of 50,000 iterations were used, with the first 5,000 excluded as the burn-in.

RESULTS AND DISCUSSION

The GWAS detected multiple signals across the genome associated with susceptibility to Johne's disease (Figure 1).

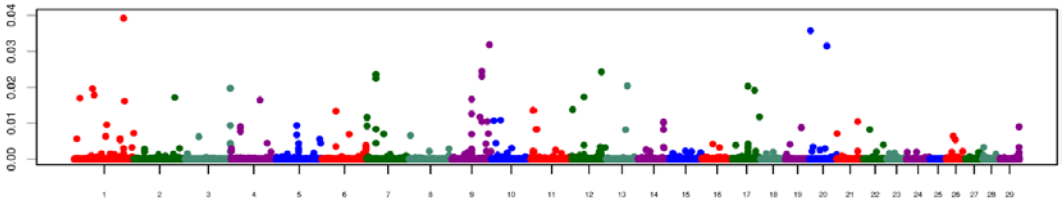


Figure 1. Bayes effect variances for an association of single nucleotide polymorphisms (SNP), with Johne's status, from a multi-SNP, Bayes B genome-wide association study

A number of the signal locations are within 1Mbp of immune-related genes. Of particular note are signals on Chromosome 3 and 7 (Figure 2) that correspond to the receptor (CSF2RA) and ligand (CSF2) for Colony Stimulating Factor 2 respectively. CSF2RA has been previously linked to response to infection by *Mycobacterium bovis* (Meade *et al.* 2008). There was no significant overlap between the major signals identified in this study and those reported in previous studies (Kirkpatrick *et al.* 2010; Minozzi *et al.* 2012; van Hulzen *et al.* 2012).

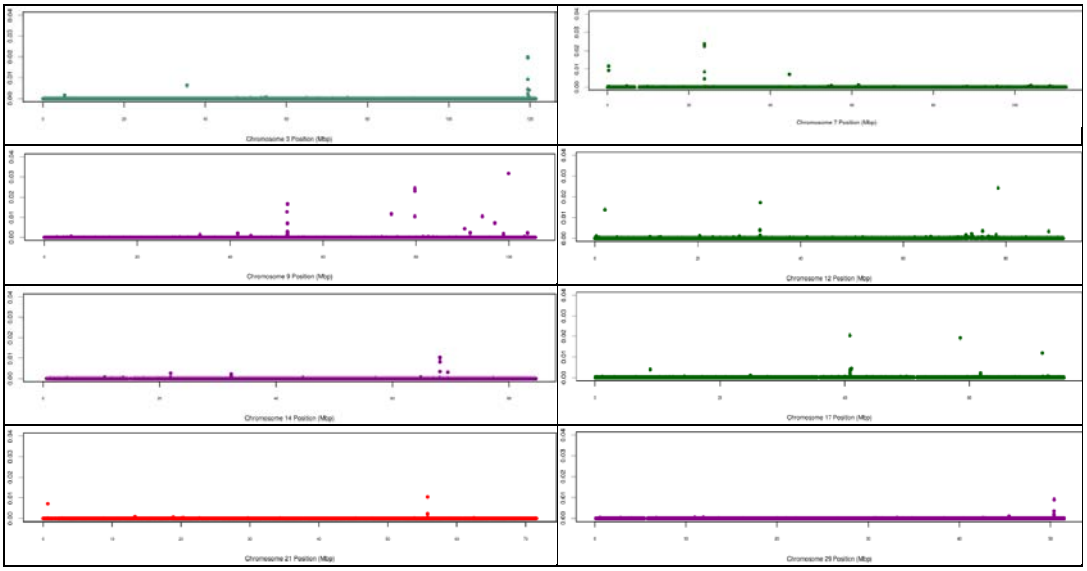


Figure 2. Bayes effect variances for an association of single nucleotide polymorphisms (SNP) with Johne's status on Chromosomes 3, 7, 9, 12, 14, 17, 21 and 29, from a multi-SNP, Bayes B analysis

That these signals tend to be associated with meaningful regions suggests that the study is targeting biologically relevant genetic structures and deeper investigation of these markers may help further illuminate the biological pathways contributing to susceptibility to Johne's disease. More immediate benefit may be gained by using the data to develop a predictive test that could be applied to an animal's genomic profile to predict the susceptibility of the animal (and its progeny) to MAP infection.

REFERENCES

Attalla S.A., Seykora A.J., Cole J.B. and Heins B.J. (2010) *J. Dairy Sci.* **93**: 1729.

Browning B.L. and Browning S.R. (2009) *Am. J. Hum. Genet.* **84**: 210.

Fernando R. and Garrick D.J. (2008) 'GenSel: User Manual for a Portfolio of Genomic Selection Related Analyses'. Animal Breeding and Genetics, Iowa State University.

Gonda M.G., Chang Y.M., Shook G.E., Collins M.T. and Kirkpatrick, B.W. (2006) *J. Dairy Sci.* **89**: 1804.

Grant I.R. (2005) *J. Appl. Microbiol*. **98**: 1282.

Kirkpatrick B.W., Shi X., Shook G.E., and Collins M.T. (2010) *Anim. Genet.* 149.

Meade K.G., Gormley E., O'Farrelly C., Park S.D., Costello E., Keane J., Zhao Y., and MacHugh D.E. (2008) *BMC Genomics* **9**: 447.

Meuwissen T.H.E., Hayes B.J. and Goddard M.E. (2001) *Genetics*. **157**: 1819.
Minozzi G., Williams J.L., Stella A., Strozzi F., Luini M., Settles M.L., Taylor J.F., Whitlock R.H., Zanella R. and Neibergs, H.L. (2012) *PLoS ONE*. **7**:e32578.
Norton S. (2007) PhD Thesis, Massey University.
Ott S.L., Wells S.J. and Wagner B.A. (1999) *Prev. Vet. Med*. **40**: 179.
Pant S.D., Schenkel F.S., Verschoor C.P., You Q., Kelton D.F., Moore S.S. and Karrow N.A. (2010) *Genomics*. **95**: 176.
Purdie A.C., Plain K.M., Begg D.J., de Silva K. and Whittington R.J. (2011) *Comp. Immunol. Microbiol. Infect. Dis.* **34**: 197.
Wellcome Trust Case Control Consortium (2007) *Nature*. **447**: 661.
van Hulzen K.J.E., Schopen G.C.B., van Arendonk J.A.M., Nielen M., Koets A.P., Schrooten C. and Heuven H.C.M. (2012) *J. Dairy Sci.* **95**: 2740.

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:199-202

GENETIC EVALUATION FOR RESISTANCE TO METABOLIC DISEASES IN CANADIAN HOLSTEINS

F. Miglior[1,2], A. Koeck[3], J. Jamrozik[3], D.F. Kelton[4], S. Loker,[3,5] K. Stachowicz,[3,6] and F.S. Schenkel[3]

[1]Guelph Food Research Centre, Agriculture and Agri-Food Canada, Guelph, ON, Canada, N1G 5C9
[2]Canadian Dairy Network, Guelph, ON, Canada, N1K 1E5
[3]Centre for Genetic Improvement of Livestock, University of Guelph, Guelph, ON, Canada, N1G 2W1
[4]Department of Population Medicine, Ontario Veterinary College, Guelph, ON, Canada, N1G 2W1
[5]Livestock Improvement Co, Newstead, Hamilton 3284, New Zealand
[6]AbacusBio Limited, PO Box 5585, Dunedin, New Zealand

SUMMARY

The overall objective of this study was to investigate if a genetic evaluation for resistance to metabolic diseases (**METAB**), which jointly includes cases for ketosis, displaced abomasum and milk fever, would be feasible in Canada. Health data recorded by producers were available from the national dairy cattle health recording system. Heritability estimates for METAB were 0.03 and 0.02 in first and later lactations, respectively. METAB in first lactation was a different trait than METAB in later lactations (genetic correlation = 0.76). Moderate genetic correlations were found between METAB and body condition score, fat to protein ratio and milk ß-hydroxybutyrate in first lactation cows. Pearson correlations between breeding values for METAB resistance and other routinely evaluated traits were computed, which revealed noticeable favorable relationships to direct herd life and fertility. The present study showed that a genetic evaluation for resistance to METAB based on producer-recorded health data would be feasible in Canada. Selection for METAB would also have a positive impact on cow's fertility and longevity.

INTRODUCTION

In Canada, a national dairy cattle health and disease data management system was started in 2007. The main objectives of this initiative are to provide information to dairy producers and their veterinarians for herd management and to establish a national genetic evaluation system for genetic selection for disease resistance. Eight diseases that are known to affect herd profitability are recorded by producers on a voluntary basis: mastitis, displaced abomasum, ketosis, milk fever, retained placenta, metritis, cystic ovaries and lameness. The feasibility of using producer recorded health data for genetic evaluations for disease resistance in Canada has been shown previously by Neuenschwander *et al.* (2012). In this study the first results of a genetic evaluation for resistance to metabolic diseases (**METAB**) in Canadian Holsteins are presented.

MATERIALS AND METHODS

Data. Health data recorded by dairy producers from April 2007 to August 2012, body condition score (**BCS**) records, as well as test-day records of milk, fat and protein yield between 5 and 55 DIM, were obtained from the Canadian Dairy Network (Guelph, Ontario). Data on milk ß-hydroxybutyrate (**BHBA**), which is only recorded in some herds in Quebec, was available from January 2011 onwards. In order to ensure that all cows were from herds with reliable recording of METAB, several editing criteria were applied. Only herds having at least two records of METAB were considered. The first and last record had to be at least 180 d apart to remove herds which had

done recording just for a short time period. In addition, a minimum disease frequency (reported cases per herd and year) of 1% was applied to ensure continuous data recording within individual herds. For genetic analyses, only records from first to fifth lactation Canadian Holstein cows were considered. The sire pedigree file was generated by tracing the pedigrees of sires and maternal grandsires back as far as possible.

Traits. The trait METAB was defined as a binary trait (0 = no disease case, 1 = at least one disease case) based on whether or not the cow had at least one case for ketosis, displaced abomasum or milk fever in the period from calving to 100 d after calving. Three traits that are indicators of energy balance and may subsequently be related to the metabolic status of an animal were included: BCS, fat to protein ratio (**F:P**) and milk BHBA. BCS was routinely recorded by professional type classifiers on a scale from 1 to 5 in increments of 0.25. Only first classifications within 365 DIM were analyzed. For BCS only information from first lactation cows was available. For F:P and milk BHBA the first-day record between 5 and 55 DIM was considered as almost all cases of METAB occur during this time period. For the traits METAB, F:P and milk BHBA first and later lactation records were considered as different traits. Summary statistics of the analyzed data is given in Table 1.

Table 1. Summary statistics of the data set used [metabolic diseases (METAB), body condition score (BCS), first test-day fat to protein ratio (F:P) and milk ß-hydroxybutyrate (BHBA)]

Trait[1]	Number of records	Mean	SD
$METAB_1$, %	141,297	3.82	0.19
$METAB_{2+}$, %	266,677	7.99	0.27
BCS, points	124,259	2.82	0.36
$F:P_1$	134,100	1.32	0.27
$F:P_{2+}$	251,835	1.33	0.28
$BHBA_1$, mmol/l	7,701	0.15	0.08
$BHBA_{2+}$, mmol/l	14,894	0.16	0.09

[1]1 = first lactation cows; 2+ = second and higher lactation cows

Model. Linear sire models were fitted using the AI-REML procedure in the DMU package (Madsen and Jensen, 2008).

The model used for $METAB_1$, $F:P_1$ and milk $BHBA_1$ was: $\mathbf{y = X\beta + Z_h h + Z_s s + e}$
where **y** is a vector of observations; **β** is a vector of systematic effects, including fixed effects of age at calving for all traits, year-season of calving for all traits and days in milk for $F:P_1$ and milk $BHBA_1$; **h** is a vector of random herd-year of calving effects for all traits; **s** is a vector of random additive genetic sire effects; **e** is a vector of random residuals; and **X**, $\mathbf{Z_h}$, and $\mathbf{Z_s}$ are the corresponding incidence matrices.

The model for BCS was: $\mathbf{y = X\beta + Z_s s + e}$
where **y** is a vector of observations; **β** is a vector of systematic effects, including fixed effects of herd-round-classifier and age at calving-stage of lactation; **s** and **e** are as defined above; and **X**, and $\mathbf{Z_s}$ are the corresponding incidence matrices.

The model for $METAB_{2+}$, $F:P_{2+}$ and milk $BHBA_{2+}$ was: $\mathbf{y = X\beta + Z_h h + Z_s s + Z_p p + e}$
where **y** is a vector of observations; **β** is a vector of systematic effects, including fixed effects of parity, year-season of calving and days in milk; **p** is a vector of random permanent environmental effects; **h**, **s** and **e** are as defined above; and **X**, $\mathbf{Z_h}$, $\mathbf{Z_s}$, and $\mathbf{Z_p}$ are the corresponding incidence matrices. Bivariate models were carried out for each combination of 2 traits considered in the present study. Assumptions were that:

[**h**' **s**' **p**' **e**']' ~ N[**0**, **V**] with $\mathbf{V} = \sum_{i=1}^{4} \mathbf{V}_i$, where

$\mathbf{V}_1 = \mathbf{I} \otimes \mathbf{H}$, **I** is an identity matrix, **H** is a covariance matrix for HY effects;
$\mathbf{V}_2 = \mathbf{A} \otimes \mathbf{G}$, **A** is an additive relationship matrix, **G** is a genetic covariance matrix;
$\mathbf{V}_3 = \mathbf{I} \otimes \mathbf{P}$, **P** is a covariance matrix for permanent environmental effect;
$\mathbf{V}_4 = \sum_{i=1}^{N} \mathbf{E}_i$, $\mathbf{E}_i$ is a residual covariance matrix. Residuals for all traits were assumed to be correlated.

Breeding value estimation. Breeding values of sires with at least 30 daughters for METAB in first and later lactations were obtained from a bivariate analysis. Estimated breeding values were reversed in sign. Thus, higher breeding values indicate sires with daughters more resistant to METAB.

RESULTS AND DISCUSSION

The frequency of METAB was 3.8 and 8.0% in first and later lactations, respectively. Heritability estimates and genetic correlations for all traits are given in Table 2. Heritability estimates for METAB were 0.03 and 0.02 in first and later lactations, respectively. In agreement with previous studies, a heritability of 0.23 was obtained for BCS. Heritabilities for F:P were 0.15 and 0.12 in first and later lactations, respectively. For milk BHBA a heritability of 0.12 was obtained in first and later lactations. METAB in first lactation was a different trait than METAB in older cows (genetic correlation = 0.76). In first lactation cows, METAB was moderately correlated with BCS (-0.44). Moderate correlations of 0.29 and 0.32 were also found between METAB and F:P and METAB and milk BHBA, respectively, in first lactation cows. Genetic correlations between METAB in later lactation cows and F:P and milk BHBA were lower.

Table 2. Heritabilities (on the diagonal) and genetic correlations (above the diagonal) for metabolic diseases (METAB), body condition score (BCS), first test-day fat to protein ratio (F:P) and milk ß-hydroxybutyrate (BHBA)

Traits	$METAB_1$	$METAB_{2+}$	BCS	$F:P_1$	$F:P_{2+}$	$BHBA_1$	$BHBA_{2+}$
$METAB_1$	**0.03**	0.76	-0.44	0.29	-0.05	0.32	-0.08
$METAB_{2+}$		**0.02**	-0.13	0.01	0.12	0.25	0.11
BCS			**0.23**	-0.33	-0.03	-0.36	-0.02
$F:P_1$				**0.15**	0.75	0.32	0.02
$F:P_{2+}$					**0.12**	0.12	0.19
$BHBA_1$						**0.12**	0.68
$BHBA_{2+}$							**0.12**

(1 = first lactation cows; 2+ = second and higher lactation cows)

Correlations of sire breeding values for resistance to METAB with other routinely evaluated traits are shown in Table 3. Routinely evaluated traits in Canada, with the exception of SCS, are scored so that a higher breeding value is favorable. Higher angularity was genetically linked with more cases of METAB. Favorable associations were found with fertility and longevity, which indicate that selection for resistance to METAB would lead to selection for cattle with improved fertility and longer herd life. Also, a higher resistance to METAB in later lactations was associated with a better Lifetime Profit Index (**LPI**) and the LPI-Production component.

Table 3. Pearson correlations between breeding values of sires with at least 30 daughters for resistance to metabolic diseases in first ($METAB_1$) and later lactations ($METAB_{2+}$) and other routinely evaluated traits (n=number of sires)

Trait[1]	$METAB_1$ (n=525)	$METAB_{2+}$ (n=1,084)
LPI (Lifetime Profit Index)	0.04	0.21***
LPI – Production	-0.03	0.23***
LPI – Durability	0.10*	-0.02
LPI – Health & Fertility	0.08	0.06*
Milk yield	-0.02	0.09**
Direct herd life	0.31***	0.18***
Somatic cell score	0.02	-0.04
Calving to first service	0.20***	0.09**
First service to conception (cows)	0.09*	0.03
Days open	0.15***	0.06
Angularity	-0.32***	-0.14***

[1]$METAB_1$, $METAB_{2+}$ and routinely evaluated traits, with the exception of SCS, are scored so that a higher breeding value is favorable. Significant effects: *P<0.05, **P<0.01, ***P<0.001.

CONCLUSIONS

The present study showed that a genetic evaluation for resistance to METAB based on producer-recorded health data would be feasible in Canada. Selection for METAB would also have a positive impact on cow's fertility and longevity.

ACKNOWLEDGMENTS

All dairy producers recording health data are acknowledged. This study was funded by the DairyGen council of CDN (Guelph, ON, Canada) and the NSERC of Canada (Ottawa, ON, Canada).

REFERENCES

Madsen P. and Jensen J. (2008) Danish Institute of Agricultural Sciences, Tjele, Denmark.

Neuenschwander T.F.-O, Miglior F., Jamrozik J., Berke O., Kelton D.F. and Schaeffer L.R. (2012) *Animal* **6**:571.

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:203-206

BOVINE NEURONAL CEROILD LIPOFUSINOSIS IN AUSTRALIAN DEVON CATTLE

H. Okazaki[1], E. Jonas[1,2], P.C. Thomson[1] and I. Tammen[1]

[1] ReproGen, Faculty of Veterinary Science, The University of Sydney, Camden, NSW, 2570
[2] Swedish University of Agricultural Sciences, Department of Animal Breeding and Genetics, 750 07 Uppsala, Sweden

SUMMARY

Neuronal ceroid lipofusinoses (NCL) are a group of lethal inherited progressive neurodegenerative disorders and more than 20 different genes have been associated with these diseases. In Australian Devon cattle there is strong evidence that NCL is caused by a single nucleotide insertion in the *CLN5* gene (c.662dupG). The aim of the present study was to estimate the frequency of the disease-causing allele in the Australian Devon cattle population. Samples from 300 randomly selected animals were requested, 190 samples were received and genotyped using a previously described DNA test. All animals were homozygous normal and the allele frequency of NCL in Australian Devon cattle was therefore estimated to be zero.

INTRODUCTION

Neuronal ceroid lipofusinoses (NCL) are a group of inherited lethal progressive neurodegenerative disorders in humans and many other animal species (Mole *et al.* 2011). Affected individuals show a progressive loss of visual, motor and mental function, often suffer from seizures and die prematurely. Characteristic findings are the accumulation of auto-fluorescent storage bodies comprised of either subunit C of mitochondrial ATP synthase or saponins in lysosomes, especially in neurones (Mole *et al.* 2011). NCL are mostly recessively inherited and at least 21 different genes have been associated with these diseases in humans and various other animal species (Bond *et al.* 2013; Mole 2013). Currently no cure is available but clinical trials for gene therapy, stem cell therapy and various pharmacological approaches are in progress (Mole *et al.* 2011, Mole 2013).

Both naturally occurring and artificially induced mutations can cause NCL in various animal species (Bond *et al.* 2013). Naturally occurring bovine NCL has been reported in Beefmaster (Read and Bridges 1969), Devon (Harper *et al.* 1988) and in a single Holstein-Friesian bull (Hafner *et al*. 2005). So far, a disease causing mutation has only been proposed for the NCL variant in the Devon breed and a direct DNA test was developed identifying a nucleotide duplication in exon 4 (c.662dupG) of the bovine *CLN5* gene on bovine chromosome 12 (Houweling *et al.* 2006). *CLN5* codes for a protein with unknown function (Kollmann *et al.* 2013). The previous study by Houweling *et al.* (2006) was only focused on a single herd within New South Wales (NSW), Australia, and it was therefore not possible to make a statement about the allele frequency or the prevalence of the disease in the Australian Devon population.

As Devon cattle are raised nationwide we describe here large scale sampling and genotyping in order to estimate the allele frequency of NCL in the Australian Devon breed. The results of this study are aimed to assist the Devon breed organisation to make informed decisions about the impact of NCL in Devons.

MATERIALS AND METHODS

Random sampling. The Devon Cattle Breeders' Society of Australia approved access to their electronic herd book and a data extract was obtained from ABRI, University of New England. The extract compromised of 65,535 animals born between 1922-2011 and listed more than 100 owner

identifiers in Australia and New Zealand. A random sampling approach was applied to the subset of 4880 animals born after 2005. The sampling size was calculated using an estimated relative targeting allele frequency for the disease. Considering an estimated allele frequency in the original herd (p = 0.03) as the worst case for NCL in the Devon population, sample size (n) was determined using the following formula, $se(p) = \sqrt{p(1-p)/(2n)}$ where p is the allele frequency, se(p) is the standard error of the estimate. A relative error se(p)/p less than 0.3 was chosen as a minimum requirement suggesting that a minimum of 200 animals are needed to be genotyped to be reasonably confident that a true allele frequency greater than zero can be detected (Table 1). In consideration of an incomplete response rate 300 samples were requested.

Table 1. Effect of varying sample sizes (n) for different allele frequencies p. The results are shown as the relative error se(p)/p. Relative errors less than 20% and 30% of the true allele frequency are shown in italics and bold, respectively.

Frequency	Sample size (n)								
P	100	150	200	250	300	350	400	450	500
0.005	1.00	0.81	0.71	0.63	0.58	0.53	0.50	0.47	0.45
0.01	0.70	0.57	0.50	0.44	0.41	0.38	0.35	0.33	0.31
0.015	0.57	0.47	0.41	0.36	0.33	0.31	*0.29*	*0.27*	*0.26*
0.02	0.49	0.40	0.35	0.31	*0.29*	*0.26*	*0.25*	*0.23*	*0.22*
0.025	0.44	0.36	0.31	*0.28*	*0.25*	*0.24*	*0.22*	*0.21*	**0.20**
0.03	0.40	0.33	*0.28*	*0.25*	*0.23*	*0.21*	**0.20**	**0.19**	**0.18**
0.035	0.37	*0.30*	*0.26*	*0.23*	*0.21*	**0.20**	**0.19**	**0.18**	**0.17**
0.04	0.35	*0.28*	*0.24*	*0.22*	**0.20**	**0.19**	**0.17**	**0.16**	**0.15**
0.045	0.33	*0.27*	*0.23*	*0.21*	**0.19**	**0.17**	**0.16**	**0.15**	**0.15**
0.05	*0.31*	*0.25*	*0.22*	**0.19**	**0.18**	**0.16**	**0.15**	**0.15**	**0.14**

Letters including labelled sampling bags, a consent sheet, an animal ethics information statement, an NCL information sheet, a hair sampling guideline, a support letter from the Devon Cattle Breeders' Society of Australia and a stamped return envelope were sent to the 41 owners, who were asked to provide tail hair samples for requested animals (or equivalent replacement animals) and additionally semen straws of any available artificial insemination (AI) bulls. The study was approved by the University of Sydney animal ethics committee (N00/8-2011/3/5581).

DNA test. For tail hair DNA extractions, 3-4 hair roots per animal were boiled for 15 minutes with 50 µl of 200 mM NaOH. After brief centrifugation, the solution was neutralized with 50 µl of 200 mM HCl and 100 mM Tris-Cl pH 8.5. The mixture was vortexed and centrifuged 2 minutes at 13,000 rpm and diluted 1:10 with Milli-Q water. DNA was extracted from AI semen straws using the method described by Heyen *et al.* (1997). The DNA test was conducted as described by Houweling *et al.* (2006) and involved PCR and visualisation of products on 8% polyacrylamide gels using a LI-COR 4200 sequencer. The normal allele is expected to produce a product of 62 bp whereas the disease allele yields a product of 63 bp. DNA samples of homozygous affected and carrier animals were available and used as positive controls (Houweling *et al.* 2006)

RESULTS

Out of 300 hair samples requested from 41 farmers, only 54.3% (163 animals) of the hair samples were received from 36.6% (15 owners) of the owners. These included replaced samples (72.3%, n=118) selected by owners due to inaccessibility to the originally targeted animals and 13

voluntarily donated samples. In addition, 27 semen samples were donated. All hair and semen samples were processed for DNA extraction, PCR amplification and gel electrophoresis. All 190 samples were successfully genotyped and tested as homozygous normal.

DISCUSSION

NCL has been previously identified in a single herd of Devon cattle in NSW (Harper *et al.* 1988). Advanced pedigree analysis could not identify a common founder (Tammen *et al.* 2002) but molecular characterisation identified that these cattle are a model for the Finnish variant late infantile disease in humans and a direct DNA test was developed to eradicate the disease allele within the initial herd (Houweling *et al.* 2006). The present study suggests that the allele frequency of bovine NCL in registered Devon cattle in Australia was zero or very close to zero and thus a management program is currently not indicated.

However, there are several limitations in the present study. In relation to the sampling size ideally a relative error of 10% should have been considered. However, this would have required sampling of a large proportion of the whole population. The statistic used does assume that all samples are unrelated, which is unlikely to be the case. Furthermore, the samples received were less than requested (a total of 190) and the majority of these were not randomly sampled. This could have been related to the relatively short return time line of 1 month. Herdbook data might not have been updated in regards of 'death' or 'sales', which resulted in a high number of replacement animals. Sampling for this study was totally depending on the willingness of the owners to participate and a support letter from the breed organisation might be the reason for relative high response rates. Access to additional semen samples would have been of interest as they could have provided a more historic view and would have allowed testing of animals with a broader impact on the population.

A follow up study using advanced pedigree analysis such as the GeneProb software (Kerr and Kinghorn 1996) where DNA test results of a subset of animals (including carrier and affected animals identified by Houweling *et al.* (2006)) and pedigree information of all animals in the population are used to predict genotypes for all animals is recommended. This can be used to identify the founder animal and to verify that the risk for the current population is extremely low. Such a study might also provide us with some insight in why the mutation that was only ever found in Australian Devon cattle in a single herd in NSW was not found in this population screen. Considering that the owner of the initial affected animals acted very responsible and immediately notified the breed society and other breeders, the possibility of a *de novo* mutation in this herd that was then contained within the herd and eradicated (initially by culling affected animals and obligate carriers and after DNA testing of the whole herd by culling remaining carriers) might be a possible explanation.

Any suspected clinical case of NCL disease should be reported and investigated.

ACKNOWLEDGEMENT

We acknowledge the support received from the Devon Cattle Breeders' Society of Australia and are thankful to farmers for submitting samples for this study. We are very grateful for the technical support provided by Ms. Gina Attard.

REFERENCES

Bond M., Kleine Holthaus S.-M., Tammen I., Tear G. and Russell C. (2013) *BBA – Mol. Basis Dis.* (in print) http://www.sciencedirect.com/science/article/pii/S0925443913000136

Hafner S., Flynn T.E., Harmon B.G. and Hill J.E. (2005) *J. Vet. Diagn. Invest.* **17:** 194.

Harper P.A.W., Walker K.H., Healy P.J., Hartley W.J., Gibson A.J. and Smith J.S. (1988) *Acta Neuropathol*. **75**: 632.

Heyen D.W., Beever J.E., Da Y., Evert R.E., Green C., Bates S.R., Ziegle J.S. and Lewin H.A. (1997) *Anim. Genet.* **28**: 21.
Houweling P.J., Cavanagh J.A.L., Palmer D.N., Frugier T., Mitchell N.L., Windor P.A., Raadsma H.W. and Tammen I. (2006) *BBA – Mol. Basis Dis*. **1762**: 890.
Kerr R.J. and Kinghorn B.P. (1996) *J. Anim. Breed. Genet*. **113**: 457.
Kollmann K., Uusi-Rauva K., Scifo E., Tyyneläe J., Jalanko A. and Braulke T. (2013) *BBA – Mol. Basis Dis.* (in print) http://dx.doi.org/10.1016/j.bbadis.2013.01.019.
Mole S.E. (2013) NCL Resource – a gateway for Batten disease: <http://www.ucl.ac.uk/ncl/>
Mole S.E., Williams R.E. and Goebel H.H. (2011) 'The Neuronal Ceroid Lipofuscinoses (Batten Disease)' 2nd ed. Oxford University Press, Oxford.
Read W.K. and Bridges C.H. (1969) *Pathol. Vet.* **6**: 235.
Tammen I., Kurts E., Tier B., Gibson A.J., Windor P.A. and Raadsma H.W. (2002) *Proc. 7th WCGALP* **31**: 771.

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:207-210

ACCURACIES OF GENOMIC PREDICTIONS IN US BEEF CATTLE

M. Saatchi[1] and D. J. Garrick[1,2]

[1]Department of Animal Science, Iowa State University, Ames, IA
[2]Institute of Veterinary, Animal and Biomedical Sciences, Massey University, Palmerston North, New Zealand

SUMMARY

Molecular breeding values (MBV) derived from genomic information holds promise to increase the accuracy of genetic evaluation in young animals. The objective of this study was to derive and evaluate the accuracies of MBV for economically relevant traits for routine implementation in several beef cattle breeds. Within-breed application of genomic predictions improved the accuracies of genetic evaluations compared to traditional methods. Accuracies of MBV ranged from 0.37 to 0.68 in American Hereford, 0.37 to 0.85 in American Red Angus, and from 0.19 to 0.73 in American Simmental using within-breed genomic predictions. Within-breed genomic predictions had less utility when applied to other breeds. Within-breed genomic predictions were less accurate for animals that were less closely related to the training population, such as those bred in other countries. The accuracies of MBV improved slightly for some breeds when predictions were derived using multi-breed reference populations obtained by simply pooling the genotypic and phenotypic data from different breeds. Genomic information has now been implemented into routine genetic evaluation for breeders of American Angus, Hereford, Limousin, Red Angus and Simmental beef cattle and will soon be extended to other US breeds.

INRODUCTION

Now it is possible to genotype beef cattle for more than 50,000 single nucleotide polymorphisms (SNP) at relatively low cost. The resulting SNP genotypes can be used to produce molecular breeding values (MBV) for selection candidates that do not necessarily have phenotypes (Meuwissen *et al*. 2001). Selection of young animals using MBV could reduce generation intervals and increase genetic progress (Schaeffer 2006). The accuracies of resultant MBV are key to successful application of this new technology as genetic gain is directly proportional to the accuracy achieved. The objective of this study was to compare accuracies of genomic predictions using within- or multi-breed reference populations for American Hereford, Red Angus and Simmental beef cattle.

MATERIALS AND METHODS

Genotype and phenotype data. A total of 9,931 animals (3,550 Hereford, 3,178 Simmental, 1,766 Black Angus, 1,274 Red Angus, 124 Gelbvieh, 37 Brangus and 2 Charolais) were genotyped, mainly at GeneSeek (Lincoln, NE). Most animals were genotyped with the BovineSNP50 BeadChip (Illumina, San Diego, CA) but some animals (less than 5%) were genotyped with the BovineHD BeadChip (Illumina, San Diego, CA). For those animals, genotypes for markers present on the BovineSNP50 BeadChip were extracted. Deregressed estimated breeding values (DEBV) free of parent average effects, derived following Garrick *et al*. (2009), were used as response variables in weighted analyses to estimate SNP effects.

In total, 12 traits were analyzed (some traits were not recorded in all breeds, Table 1). The number of genotyped animals with DEBV varied among traits (2395 for scrotal circumference to 9443 for birth weight) because some animals had no individual or offspring information contributing to their expected progeny difference (EPD) and therefore had no information in their DEBV.

Statistical model. Habier *et al.* (2011) showed that the method "BayesC" (Kizilkaya *et al*. 2010) is less sensitive to prior assumptions than the method "BayesB" (Meuwissen *et al*. 2001). So, method BayesC was used to estimate marker effects for Red Angus and Simmental animals. However, the method BayesB was used for Hereford animals (all traits except fat thickness and marbling, which used BayesC) as the higher accuracies reported by Saatchi *et al*. (2013) for a subset of the data using BayesB method. Both methods assume that some known fraction of markers (π) have zero effect. For each trait, the following model was fit to the DEBV data for training: $y_i = \mu + \sum_{j=1}^{k} z_{ij} u_j + e_i$,

where y_i is the DEBV for animal i, μ is the population mean, k is the number of marker loci in the panel, z_{ij} is allelic state (i.e., number of B alleles from the Illumina A/B calling system) at marker j in individual i, u_j is the allele substitution effect for marker j, with $u_j \sim N(0, \sigma_u^2)$ (with probability $1 - \pi$) or $u_j = 0$ (with probability π) as described by Habier *et al*. (2011) for BayesB and BayesC methods, and e_i is a residual with heterogeneous variance, depending on the reliability of the DEBV information for animal i (Garrick *et al*. 2009). Parameter π was assumed to be 0.95 for all analyses. Markov chain Monte Carlo (MCMC) methods with 41,000 iterations were used to provide posterior mean estimates of marker effects and variances, after discarding the first 1,000 samples for burn-in.

The MBV for individual i within a validation set was derived as the sum over all k markers of posterior means of predicted SNP effects, as estimated in the training set, multiplied by the number of copies of the B allele. All analyses were performed using GenSel software (Fernando and Garrick 2010).

Within breed genomic predictions. The accuracies of MBV were evaluated by pooling estimates from either 5-fold or 6-fold cross-validation strategies for 2,980 American Hereford, 1,274 American Red Angus or 2,703 American Simmental animals. The K-means clustering method (Saatchi *et al*. 2011) was used for partitioning animals with the aim of increasing within-group and decreasing between-group relationships. Within-breed training analyses were performed by excluding one group when estimating marker effects, which were then used to predict MBV of individuals from the omitted group (validation set). Bivariate animal models were used for each trait to estimate the genetic correlation between DEBV and MBV as a reflection of the accuracy of genomic prediction (Saatchi *et al*. 2012).

Accuracies of genomic predictions were also evaluated for Red Angus animals using a multi-breed reference population (consisted of 3,178 Simmental, 1,766 Black Angus, 124 Gelbvieh, 37 Brangus, 31 Hereford and 2 Charolais plus 1,274 Red Angus animals). In multi-breed cross-validation, the same four Red Angus groups for each of the five training runs were used, except that animals from all the other breeds were always included in the training analyses. In multi-breed analyses only the accuracies of Red Angus predictions were of interest.

Across countries and across breed genomic predictions. The accuracies of genomic predictions were evaluated for 100 Argentine, 75 Canadian and 395 Uruguayan Hereford, 3,178 American Simmental and 1,274 American Red Angus animals using marker estimates from training on American Hereford animals. Simple correlations between MBV and DEBV were used as estimates of the accuracies of MBV in non American Hereford animals because pedigree information was not available to estimate genetic correlations between DEBV and MBV.

RESULTS

Accuracies of MBV ranged from 0.37 to 0.68 (average 0.53) in American Hereford, 0.37 to 0.85 (average 0.64) in American Red Angus, and from 0.19 to 0.73 (average 0.50) in American Simmental using within-breed genomic predictions (Table 1). Genomic predictions were more accurate in Red Angus using multi-breed rather than the single-breed reference population for all

traits except calving ease and weaning weight maternal (ranged from 0.32 to 0.90 with the average of 0.69, Table 1).

Table 1. Accuracies[1] of genomic predictions (±SE) in American Hereford, Simmental and Red Angus beef cattle using single- and multi-breed reference populations.

Trait	Single-breed			Multi-breed
	Hereford	Simmental	Red Angus	Red Angus
Birth weight	0.68±0.03	0.67±0.03	0.66±0.04	0.75±0.04
Calving ease direct	0.68±0.04	0.46±0.02	0.59±0.03	0.60±0.04
Calving ease maternal	0.51±0.04	0.31±0.02	0.37±0.03	0.32±0.04
Carcass weight	-	0.61±0.04	0.62±0.04	0.75±0.04
Fat thickness	0.48±0.04	0.19±0.02	0.85±0.16	0.90±0.15
Marbling	0.43±0.04	0.60±0.04	0.77±0.10	0.85±0.09
Rib eye muscle area	0.49±0.03	0.55±0.05	0.71±0.07	0.75±0.06
Scrotal circumference	0.43±0.04	-	-	-
Weaning weight direct	0.52±0.03	0.56±0.04	0.55±0.04	0.67±0.04
Weaning weight maternal	0.37±0.03	0.32±0.03	0.54±0.04	0.51±0.04
Yield grade	-	0.73±0.09	0.81±0.18	0.83±0.12
Yearling weight	0.60±0.03	0.45±0.02	0.57±0.04	0.69±0.03

[1]Measured as the genetic correlation between tderegressed estimated breeding values and molecular breeding values estimated from a bivariate animal model.

Table 2. Accuracies[1] of genomic predictions in non American Hereford, American Simmental and American Red Angus beef cattle using marker estimates from training on American Hereford.

Trait	Argentine Hereford	Canadian Hereford	Uruguayan Hereford	American Simmental	American Red Angus
Birth weight	0.15	0.48	0.24	0.29	0.28
Calving ease direct	-	0.58	0.28	0.30	0.27
Calving ease maternal	-	0.46	0.19	0.17	-0.11[2]
Fat thickness	-0.19[2]	0.30	0.12	0.09	0.15
Marbling	0.35	0.58	0.23	0.16	0.05
Rib eye muscle area	0.17	0.43	0.20	0.12	0.15
Scrotal circumference	-0.16	0.26	0.17	-	-
Weaning weight direct	0.10	0.31	0.28	0.22	0.21
Weaning weight maternal	0.00	0.24	0.20	0.05	0.07
Yearling weight	0.04	0.14	0.33	0.23	0.24

[1]Simple correlations between deregressed estimated breeding values (DEBV) and molecular breeding values (MBV).
[2]By definition, the accuracy cannot be a negative value. However, the negative value obtained here as the simple correlation between DEBV and MBV used as a measure of accuracy.

Within-breed genomic predictions were less accurate for animals that were less closely related to the training population, such as those bred in other countries (Table 2). Genomic predictions for Argentine, Canadian and Uruguayan Hereford were less accurate than those obtained for American Hereford animals (simple correlations between DEBV and MBV have not shown for American Hereford). Among non American Hereford animals, genomic predictions were most

accurate for Canadian Hereford (Table 2). Across-breed genomic predictions were less accurate than those obtained from within breed genomic predictions (Table 2). Among all traits, across-breed genomic predictions were higher for birth, weaning and yearling weights; and calving ease direct, than for other traits (Table 2).

DISCUSSION

This study applied genomic prediction to American and non American Hereford, American Red Angus and American Simmental beef cattle using single-breed or multi-breed reference populations. The accuracies of genomic predictions were more accurate than those reported by Saatchi *et al*. (2013) for American Hereford animals (the average of accuracies increased from 0.37 to 0.53). This is due to the larger training populations (2,980 vs. 1,081 animals) used in this study. Goddard and Hayes (2009) showed that the size of training population is a crucial factor influencing the accuracies of genomic predictions. Among non American Hereford animals, genomic predictions were more accurate for Canadian Hereford, which reflects the higher degree of genetic relationship between American and Canadian Hereford population in comparison with Argentine and Uruguayan Hereford populations (Saatchi *et al*. 2013).

Genomic predictions were more accurate using multi-breed than single-breed reference populations for Red Angus animals for all traits except calving ease and weaning weight maternal (the average of accuracies increased from 0.64 to 0.69). This may reflect the fact that some registered Simmental animals have a heterogeneous genetic background being admixed with other beef cattle breeds including Red Angus, as American Simmental Associations allow registration of crossbred animals with other beef cattle breeds. This demonstrates the benefit of using multi-breed reference population for American Red Angus beef cattle.

Across breed genomic predictions had less utility when applied to other breeds for most traits. This may reflect differences in linkage phase between markers and quantitative trait loci (QTL) or differences in causative mutations at the same QTL for these traits across different breeds. However, across-breed genomic predictions had some utility for birth, weaning and yearling weights; and calving ease direct traits due to the segregation of common QTL with large effects among these breeds (unpublished data).

Genomic information has now been implemented into routine genetic evaluation for breeders of American Angus, Hereford, Limousin, Red Angus and Simmental beef cattle and will soon be extended to other US breeds.

REFERENCES

Fernando R.L. and Garrick D.J. (2010) http://taurus.ansci.iastate.edu/ Accessed Sep. 1, 2010.

Garrick D.J., Taylor J.F. and Fernando R.L. (2009) *Genet. Sel. Evol.* **41**: 55.

Goddard M.E. and Hayes B.J. (2009) *Nat. Rev. Genet.* **10**: 381.

Habier D., Fernando R.L., Kizilkaya K. and Garrick D.J. (2011) *BMC Bioinformatics* **12**: 186.

Kizilkaya K., Fernando R.L. and Garrick D.J. (2010) *J. Anim. Sci.* **88**: 544.

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

Saatchi M., Ward J. and Garrick D.J. (2013) *J. Anim. Sci.* **91**: 1538.

Saatchi M., McClure M.C., McKay S.D., Rolf M.M., Kim J., Decker J.E., Taxis T.M., Chapple R.H., Ramey H.R., Northcutt S.L., Bauck S., Boodward B., Dekkers J.C.M., Fernando R.L., Schnabel R.D., Garrick D.J. and Taylor J.F. (2011) *Genet. Sel. Evol.* **43**: 40.

Saatchi M., Schnabel R.D., Rolf M.M., Taylor J.F. and Garrick D.J. (2012) *Genet. Sel. Evol.* **44**: 38.

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

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:211-214

ACCURACY OF IGENITY DIRECT GENOMIC VALUES IN AUSTRALIAN ANGUS

V. Boerner and D. J. Johnston

Animal Genetics and Breeding Unit[1], University of New England
Armidale, 2351, NSW, Australia

SUMMARY

The quality of Igenity[2] direct genomic values (GEBVs) derived by two different prediction procedures for 12 traits of 1032 Angus bulls was estimated as the genetic correlation to their phenotypic target traits. In addition, the effect of a decreasing genetic relationship between validation and training population was inferred by subdividing the set of 1032 GEBVs accordingly. Genetic correlations estimated were medium to high even when all training individuals were excluded from the analysis, and well in line with those already published. Thus blending Australian Angus breeding values with Igenity GEBVs can be beneficial for breeders.

INTRODUCTION

GEBVs, calculated by applying previously derived prediction equations to known SNP genotypes, are available for Australian Angus beef cattle from at least two commericial suppliers (www.pfizer.com,www.igenity.com). The value of this additional information to breeders depends on the genetic correlation (accuracy, r_g) to their phenotypic target traits. An analysis of GEBVs from both providers by the American Angus Association found such correlations between 0.65 and 0.29 depending on the trait (Northcutt 2011). Evaluations of Pfizer Molecular Value Predictions done in the Australian Angus population resulted in r_gs between 0.45 and 0.2 (Johnston *et al.* 2010). For Igenity molecular breeding values r_gs of 0.8 for scan intra-muscular fat content of yearling bulls and 0.38 for carcase marbling score were found in American Angus (MacNeil *et al.* 2010).

This paper presents results of a correlation analysis of Angus GEBVs supplied by Ingenity for 12 different traits of which phenotypic target traits are also recognised in the usual breeding value estimation for this breed. As the training individuals were part of the GEBV set, and for each trait GEBVs from two different prediction procedures were supplied, we have also analysed the effect of an increasing genetic distance between training and validation population on r_gs and how differently derived prediction equations affect these correlations.

METHODS

The accuracy of GEBVs was determined as the genetic correlation between GEBVs (modelled as traits) and the corresponding phenotypic target traits estimated using REML or Gibbs sampling in a bi-variate approach.

Direct genomic values of 1032 Angus bulls for birth weight (d.BWD), 200 day weight direct (d.WWD), 200 day weight maternal (d.WWM), 400 day weight (d.YWD), mature cow weight (d.MCW), scrotal circumference (d.SC), carcase weight (d.CWT), carcase intra-muscular fat content (d.CIM), carcase ribeye area (d.CEA), direct calving ease (d.CED), maternal calving ease (d.CEM) and docility (d.DOC), predicted by two different procedures (50K3 and 50KGB), were supplied by Igenity (http://www.igenity.com/). For both

[1]A joint venture of the NSW Department of Primary Industry and the University of New England
[2]Igenity is a registered trademark of Neogen Corporation.

prediction procedures, the underlying genotype was obtained from an Illumina 50K Bead Chip, but 50K3 GEBVs were calculated from prediction equations derived on 392 SNP individually chosen for each trait, whereas 50KGB GEBVs were calculated from prediction equations derived in a GBLUP approach. Across prediction procedures, GEBVs were supplied in two sets, A: 736 GEBVs of American Angus individuals genotyped in the US and used in the Igenity training set, and B: 355 GEBVs of Australian Angus individuals genotyped in Australia. To analyse the effect of an increasing genetic distance (decreasing genetic relationship) between training and validation population, sets A and B were united and then subdivided as follows: FULL: all genotyped individuals of set A and B (n=1032). AU: only set B individuals (n=345). AUS: as AU, but direct progeny of individuals in set A were excluded (n=188).

Phenotypic traits included in the analysis were birth weight (p.BWD, n=248562), 200 day weight (p.WW, n=234087), 400 day weight (p.YWD, n=156893), mature cow weight (p.MCW, n=90795), carcase weight (p.CWT, n=4535), carcase intramuscular fat percentage (p.CIM, n=3434), carcase eye muscle area (p.CEA, n=2732), scrotal circumference (p.SC, n=159171), calving ease (p.CE, n=161172) and docility (p.DOC, n=13050). Records were obtained from the Australian Angus Society database. Note that in Australian Angus phenotypic calving ease and docility are recorded as calving difficulty and wildness, respectively, so negative correlations were expected for these traits.

The linear model was $y = Xb + Z_d u_d + Z_m u_m + Z_q p_q + Z_r p_r + e$, where y is a vector phenotypes, b is a vector of fixed effects, u_d is a vector of random direct genetic effects, u_m is a vector of random maternal genetic effects, p_q is a vector of random maternal environmental effects, p_r is a vector of random permanent environmental effects and e is a vector of random residual effects. X, Z_d, Z_m, Z_q and Z_r are incidence matrices linking the effects to their respective phenotypes. Note that for GEBVs, X is a vector of ones. It was assumed that traits $\sim N(Xb, Z_d A Z_d' \sigma_d^2 + Z_m A Z_m' \sigma_m^2 + Z_d A Z_m' \sigma_{d,m} + Z_q I Z_q' \sigma_q^2 + Z_r I Z_r' \sigma_r^2 + I\sigma_e^2)$, where A is the numerator relationship matrix built from a pedigree such that every individual with an observation had at least, if available, three generations of ancestors and I is an identity matrix. u_m and p_q were modelled only for p.BWD, p.WW, p.YWD and p.CE, and p_r only for p.MCW.

The software used to estimate parameters of continuously distributed phenotypic traits and their related GEBVs was WOMBAT(Meyer 2007). Parameters of categorically distributed phenotypic traits and their related GEBVs (p.CE, p.DOC, d.CED, d.CEM, and d.DOC) were estimated using a Gibbs sampling approach for threshold traits (Albert and Chib 1993), implemented in the thrgibbsf90 software (Tsuruta and Misztal 2006).

RESULTS

Table 1 summarises r_gs between GEBVs and their phenotypic target traits. Note that for d.WWM and d.CEM the correlation to the maternal genetic component of p.WW and p.CE, respectively, is given. In general, r_gs of 50KGB and 50K3 GEBVs were very similar and showed the same trend in response to changes of the GEBV set. For FULL sets, highest r_g of 0.69 was found for $\text{d.SC}^{\text{Full}}_{\text{50K3}}$, followed by 0.67 for $\text{d.BWD}^{\text{Full}}_{\text{50K3}}$. The exclusion of US training individuals (FULL→AU) led to a decrease in r_g of more than 0.1 only for d.BWD, d.WWD, d.YWD, d.MCW, d.SC and $\text{d.WWM}_{\text{50K3}}$. For all other GEBVs a decrease < 0.1 or even an increase (e.g. d.CIM, $\text{d.CWT}_{\text{50K3}}$, d.CEM) was be observed. Thus, r_gs of continuous reproductive and growth traits were affected most by this exclusion, whereas carcase and categorical traits were unaffected. When excluding additionally the progeny of training individuals (AU →AUS), r_gs of growth and reproductive traits decreased further (except

Table 1: Genetic correlation (accuracy)|*standard error* between GEBVs and their phenotypic target traits by estimation procedures and GEBV subsets

GEBV	phenotypic trait	50KGB[1]			50K3[2]		
		FULL[3]	AU[4]	AUS[5]	FULL	AU	AUS
d.BWD	p.BWD	0.65|*0.03*	0.45|*0.06*	0.46|*0.08*	0.67|*0.03*	0.44|*0.07*	0.35|*0.09*
d.WWD	p.WW	0.64|*0.03*	0.42|*0.06*	0.35|*0.09*	0.60|*0.03*	0.44|*0.06*	0.32|*0.10*
d.YWD	p.YWD	0.61|*0.03*	0.37|*0.06*	0.28|*0.10*	0.53|*0.04*	0.31|*0.07*	0.15|*0.10*
d.MCW	p.MCW	0.48|*0.05*	0.26|*0.08*	0.12|*0.11*	0.47|*0.05*	0.29|*0.08*	0.16|*0.12*
d.SC	p.SC	0.61|*0.03*	0.42|*0.07*	0.41|*0.10*	0.69|*0.03*	0.53|*0.07*	0.49|*0.10*
d.CWT	p.CWT	0.50|*0.12*	0.47|*0.14*	0.49|*0.18*	0.55|*0.12*	0.57|*0.15*	0.78|*0.16*
d.CIM	p.CIM	0.40|*0.13*	0.46|*0.15*	0.59|*0.17*	0.54|*0.14*	0.75|*0.14*	0.91|*0.16*
d.CEA	p.CEA	0.47|*0.13*	0.45|*0.16*	0.50|*0.20*	0.40|*0.16*	0.30|*0.19*	0.45|*0.26*
d.WWM	p.WW	0.35|*0.06*	0.30|*0.08*	0.26|*0.12*	0.36|*0.06*	0.24|*0.10*	0.20|*0.14*
d.CED	p.CE	-0.21|*0.11*	-0.18|*0.07*	0.04|*0.09*	-0.15|*0.11*	-0.11|*0.07*	0.17|*0.11*
d.CEM	p.CE	-0.24|*0.09*	-0.41|*0.06*	-0.38|*0.07*	-0.25|*0.10*	-0.47|*0.05*	-0.39|*0.09*
d.DOC	p.DOC	-0.23|*0.08*	-0.25|*0.09*	-0.13|*0.11*	-0.25|*0.09*	-0.27|*0.11*	-0.13|*0.11*

1: GEBV estimated by a GBLUP approach from trait-independent SNP genotypes obtained from an Illumina 50K Bead Chip, 2: GEBV estimated from 392 SNP individually chosen for each trait where genotypes were obtained from an Illumina 50K Bead Chip, 3: all genotype individuals, 4: individuals of Australian origin only, 5: individuals of Australian origin but no direct sons of US bulls.

$d.BWD_{50KGB}$), whereas r_gs of carcase traits increased (e.g. d.CIM, d.CEA and d.CWT).

Independently of the GEBV set size the vast majority of REML estimates of GEBV heritabilities (h^2) was one, and their standard errors increased as set size decreased (results not shown). Gibbs sampling h^2 estimates were never one, even for the FULL set regardless of the estimation procedure, and generally decreased with decreasing set size (from FULL to AUS). For continuously distributed traits the variance of the direct additive genetic effect (σ_a^2) was much larger for the phenotypic trait than for the related GEBV(e.g. 11.2 for p.CEA and 0.03 for $d.CEA^{Full}_{50K3}$, results not shown). In contrast, σ_a^2 of p.CE and p.DOC were generally smaller than those of their related GEBVs. Comparing both the estimation procedures, σ_a^2 of 50K3 GEBVs were always larger than those of 50KGB GEBVs (results not shown).

DISCUSSION

Using the AU set as a reference, the results given here (0.24 to 0.75 for continuous traits) are well in line with those already published (MacNeil *et al.* 2010; Northcutt 2011; Johnston *et al.* 2010). Blending Pfizer GEBVs of similar accuracies into Australian Angus BREEDPLAN estimated breeding values resulted in an increased overall accuracy of 1.4 % to 7.5 % dependent on the trait (Johnston *et al.* 2012). Thus, similar results can be achieved when blending Australian Angus estimated breeding values with Igenity GEBVs.

Results also show that a selection of 392 SNP individually chosen for each trait out of those present on the Illumina 50K Bovine Bead Chip performs as well as a GBLUP approach using all available SNP. Moreover, trends in r_gs and their standard errors of both prediction procedures are similar, and, apart from statistical significance, for the majority of traits r_gs from the 50K3 approach were slightly higher than from the 50KGB approach. Thus, if these 392 SNP track large haplotypes, it raises questions about the additional benefit of using 800K or full genome sequencing for accuracies of GEBVs. For growth and reproductive traits r_gs

decreased from FULL to AUS, which is in line with the theoretical expectation. Contrarily, especially for carcase traits r_gs did not generally decrease with an increasing genetic distance between the training and the validation set. This is especially the case for d.CIM_{50K3}, where r_g increased from 0.54 (FULL) to 0.91 (AUS), for d.CWT_{50K3} (0.55→0.78), and for d.CIM_{50KGB} (0.40→0.59). A possible reason for this observation is the decrease in subset sizes of GEBVs (AU (345) and AUS (188)) which possibly offset the effect of a decreasing relationship by sampling. However, a decrease in subset size occurred across GEBVs, thus also in those where r_gs decreased as expected. Compared to growth and reproductive traits, carcase traits are characterised by a generally low number of phenotypic observations. Excluding US animals when moving from FULL to AUS possibly increased the average genetic relationship between individuals with GEBVs and individuals with phenotypic observations for the 3434 p.CIM records much more than for the 234087 p.WW records. Thus, a possible positive effect of this increased relationship on r_gs might have superposed negative effects of a decreased GEBV subset size and increased genetic distance between training and validation set. However, as the average genetic relationship between GEBV sets and phenotypic trait sets was not analysed, further research in this area is necessary. Since sample sizes of GEBVs and also of phenotypic carcase traits are still limited, results need to be verified by larger number of phenotypic records and more individuals with both phenotypes and genotypes. Nevertheless, results indicate that blending Australian Angus estimated breeding values with Igenity GEBVs can improve overall accuracy especially for difficult to measure traits.

ACKNOWLEDGEMENTS

The authors thank Meat and Livestock Australia for financially supporting this work (Project B.BFG.0050). We acknowledge Igenity, Lincoln, NB, USA, for providing GEBVs and the Angus Australia Society for providing access to their data.

REFERENCES

Albert J. H. and Chib S. (1993) *J Am Stat Assoc* **88**:669 .

Johnston D. J., Jeyaruban M. G. and Graser H.-U. (2010) Evaluation of Pfizer Animal Genetics HD 50K MVP Calibration.

Johnston D. J., Tier B. and Graser H.-U. (2012) *Anim Prod Sci* **55**:100.

MacNeil M. D., Nkrumah J. D., Woodward B. W. and Northcutt S. L. (2010) *J Anim Sci* **88**(2):517.

Meyer K. (2007) *J Zhejiang Univ Sci B* **8**(11):815.

Northcutt S. L. (2011) Genomic Choices.

Tsuruta S. and Misztal I. (2006) in Proc. 8th. WCGALP, Belo Horizonte, Brazil, August, 2006, pp. 27 – 31.

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:215-219

WILL SEQUENCE SNP DATA IMPROVE THE ACCURACY OF GENOMIC PREDICTION IN THE PRESENCE OF LONG TERM SELECTION?

I.M. MacLeod[1,2], B.J. Hayes[2,3] and M.E. Goddard[1,2]

[1] Melbourne School of Land & Environment, University of Melbourne, Victoria 3010, Australia
[2] AgriBio, Department of Environment & Primary Industries, Victoria 3086, Australia
[3] Biosciences Research Centre, La Trobe University, Victoria 3086, Australia

SUMMARY

To date genomic prediction (GP) of breeding values in cattle generally exploits either ~50K or ~800K SNP chips. Now that whole genome sequence data is also available, it is important to evaluate its potential to improve the accuracy of GP. SNP chips include only more common SNP while sequence data includes rare and common SNP as well as all causal mutations (QTL). It is expected that sequence data will improve accuracy of GP particularly if QTL are rare because they have been under long-term negative selection. This study evaluates accuracy of GP using sequence data compared with the equivalent of ~800K or ~50K SNP densities. Accuracy of GP was tested in simulated populations (mimicking Holstein cattle) with and without long-term negative selection acting on QTL. GP was implemented with both BLUP (GBLUP) and Bayesian (BayesR) methods. There was not a very marked difference between GP accuracy in scenarios with neutral QTL or selected QTL because the recent low effective population size (N_e) of cattle decreased the proportion of rare causal mutations compared to expectations in larger N_e. Only the BayesR method was able to exploit an advantage from sequence data. We conclude that combining data from more than one breed in training (reference) populations and using Bayesian analyses, will take better advantage of sequence data for GP than using single breed and GBLUP analyses.

INTRODUCTION

Genomic prediction (GP) of breeding values is an efficient method of selecting livestock for traits that are difficult to measure, or traits not expressed in males (Meuwissen *et al.* 2001). To date GP in cattle generally exploits either ~50K or ~800K SNP chips, but soon whole genome sequence data (direct or imputed) could also be used to improve accuracy of GP. The advantage of sequence is that it contains the causal mutations. Furthermore, SNP chips include only common SNP and these may not be in high linkage disequilibrium with causal mutations if the latter are rare because they have been subject to long-term negative selection. In this case SNP chips will not be able to accurately estimate the QTL effects. It is therefore expected that sequence data will improve accuracy of GP, particularly if causal mutations have been under long-term negative selection.

Using a bovine-like neutral model to simulate data, Clark *et al.* (2011) demonstrated a 5-15% advantage for accuracy of GP using sequence compared to 50K SNP chip densities, but did not include a comparison with 800K SNP density. Druet *et al.* (2013) indirectly estimated the potential effect of long term negative selection on GP by simulating QTL effects on a subset of loci with low or very low minor allele frequencies (MAF). They demonstrated a 4-28% advantage in accuracy of GP using sequence data compared to 50K SNP densities, but did not test 800K SNP density. Although it can be argued that simulating QTL on rare mutations mimics the expected effect of long term negative selection, the approach may not reflect the true MAF distribution of loci actually subjected to long-term negative selection because demography also shapes the MAF distribution. For example, in populations with recent bottlenecks in effective population size (N_e), mutations with a deleterious effect on fitness are more likely to be lost, but may also sometimes rise to higher frequencies due to drift, compared to populations with large or expanding N_e. Using

simulations of bovine populations, we evaluate the accuracy of GP using sequence data, ~800K or ~50K SNP chip densities, with and without long term negative selection applied to QTL.

MATERIALS AND METHODS

We simulated sequence data with FREGENE (Chadeau-Hyam *et al.* 2008) using a demographic model demonstrated to mimic Holstein sequence data (Macleod *et al*. 2013), in which the effective population size (N_e) reduces from ancestrally very large to very small in recent times. For computational efficiency we simulated a genome size of 50Mb under the scaling argument demonstrated by Meuwissen & Goddard (2010): i.e. GP accuracy is proportional to the number of training individuals/Morgan (M) length of the genome. Therefore to achieve similar accuracies with a ~30M bovine genome, the training population size would need to increase by a factor of ~60. Simulations were either a neutral model (NEUT) or with long-term negative selection imposed on QTL (SEL). In the SEL model, 0.1% of new mutations were subject to an additive selection coefficient: $s = -2x10^{-4}$, and those still segregating at the end of the simulation were used as QTL. In both NEUT and SEL scenarios we simulated 20 replicates, each with 5000 individuals.

We created a "Medium Density" (MD) and "High Density" (HD) SNP panel for each replicate, by selecting a subset of 1000 and 10,000 SNP loci respectively: representing a density of 60K and 600K SNP across the whole bovine genome (the latter is equivalent to an 800K SNP panel because often after quality control in real data there are ~600 usable SNP). To mimic the ascertainment bias of commercial panels, SNP were only selected if MAF > 0.1 and SNP positions were then selected uniformly at random. We generated HD and MD SNP genotypes for all individuals in addition to the sequence data (SEQ). For each replicate, additive QTL effects were simulated from a normal distribution with two different QTL densities: number of QTL=50 or 15. In the NEUT populations, QTL were randomly selected from SNP loci, while in SEL scenarios the QTL were chosen from polymorphic loci subjected to selection. In 5 of the 20 replicate SEL populations, there were only 49, 47, 46, 46 and 41 selected loci still segregating, therefore for the scenario with QTL=50 the remaining QTL were drawn from neutral loci with MAF < 0.1. QTL effects were summed to give True Breeding Values (TBV_j) for each individual. Phenotypes were generated by adding a residual term to the TBV_j of each individual, drawn from a normal distribution to produce a trait heritability of 0.1. We randomly selected 3750 "training" individuals to calculate the genomic prediction equations (using genotypes and phenotypes). We used the remaining 1250 individuals from the same population (genotypes only) to validate the prediction equations (Gen=0, "validation" individuals). After both 10 and 15 further generations of random breeding, genotypes were again sampled for 2000 validation individuals (Gen=10 and Gen=15 validations).

We implemented both GBLUP and BayesR analyses to generate Genomic Estimated Breeding Values (GEBV). GBLUP was implemented in ASReml (Gilmour *et al.* 2005): $\mathbf{y} = \mu\mathbf{1} + \boldsymbol{Zg} + \boldsymbol{e}$, where μ is the population mean, $\mathbf{1}$ is a vector of 1s, $\boldsymbol{Z}$ is the incidence matrix for random individual effects. The $\boldsymbol{g}$ and $\boldsymbol{e}$ are vectors of GEBV and residuals, assumed normally distributed as $N(0, \boldsymbol{G}\sigma^2_g)$ and $N(0, \boldsymbol{I}\sigma^2_e)$, where $\boldsymbol{G}$ is the genomic relationship matrix (GRM) estimated either from MD, HD or SEQ genotypes (eg. Erbe *et al.* 2010). Our BayesR implementation (Erbe *et al.* 2012) omitted a polygenic effect because individuals were randomly bred with no close pedigree structure: $\mathbf{y} = \mu\mathbf{1} + \boldsymbol{Wu} + \boldsymbol{e}$, where μ is the mean, $\boldsymbol{e}$ is the vector of random residuals and $\boldsymbol{W}$ is the design matrix allocating records to the vector of marker effects, $\boldsymbol{u}$. The accuracy of GP was determined as the correlation between the $GEBV_j$ and the TBV_j in i=1…N validation individuals, averaged the across the 20 replicate simulations for each scenario.

RESULTS AND DISCUSSION

The marked reduction in recent effective population size (N_e) used in our simulation to mimic

the Holstein breed demography, resulted in a relatively flat derived allele frequency (DAF) distribution for neutral alleles compared to the expectation in a larger constant or expanding N_e. The recent reduction in N_e results in random drift very quickly purging low frequency loci as well as increasing linkage disequilibrium (LD) compared to larger N_e. Among neutral loci in our simulations, 19% had DAF < 0.1 while this figure increased to 31% for loci subjected to long term negative selection. This indicates that selection had a significant impact on allele frequency distribution while not being so strong as to immediately purge new mutations. The impact of the selection coefficient (s) is generally significant if: $|sN_e| >> 1$ and in our large ancestral bovine population $|sN_e| = 12$ which is similar to some estimates in humans (Keightley & Halligan 2009).

Fig 1A shows the results for the realised accuracy of GP when the number of QTL=50 (equivalent to ~3000 QTL affecting a trait genome wide) while results in Fig 1B are for QTL=15 (equivalent to 900 QTL genome wide). QTL densities were chosen to reflect realistic models based on recent mammalian estimates (eg. Kemper *et al*. 2012). In all scenarios there was an advantage for sequence (SEQ) over MD SNP (up to 11.8%) as for previous studies (Clark *et al.* 2011, Meuwissen & Goddard 2010, Druet *et al*. 2013), particularly with BayesR and an increasing number of generations separating training and validation populations (Gen=10 and 15). With GBLUP analyses there was generally no advantage for SEQ compared to HD SNP, except in the SEL scenario with QTL=15. With BayesR there was a modest advantage for SEQ over HD SNP (up to 3.6%), particularly in the Gen=15 validation and was consistently higher in SEL compared to NEUT scenarios. Furthermore, there was less decay in the BayesR accuracy compared to GBLUP when the number of generations separating training and validation individuals increased.

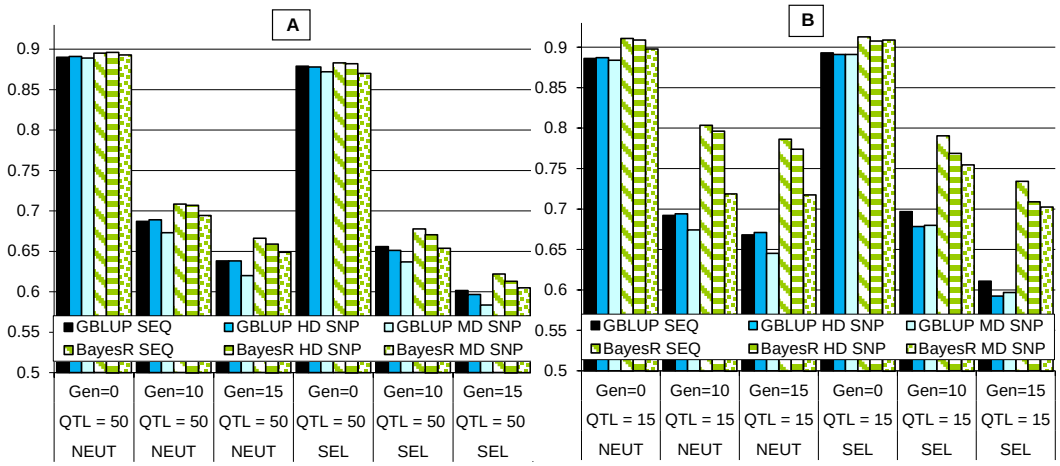


Figure 1A and B. **Genomic prediction accuracy in populations with QTL under a neutral (NEUT) or negative selection model (SEL), using GBLUP or BayesR analysis, with two contrasting QTL densities: number of QTL=50 (A) or QTL=15 (B). Zero, 10 or 15 generations separated training and validation individuals (Gen=0, 10 or 15).**

GBLUP assumes a quasi infinitesimal model with each sequence SNP assumed to contribute an additive effect sampled from a single normal distribution. BayesR method could be expected to perform better with sequence data because it sets a prior expectation that many SNP will have no effect, while the remaining effects will be sampled from a mixture of distributions, with many small effects and up to some rare large effects. However the recent reduction in N_e within *Bos taurus* cattle breeds has resulted in high (but variable) LD across relatively long chromosome

segments and therefore GBLUP will tend to "spread" the estimate of each QTL effect across a number of loci on chromosome segments in which SNP are in high LD. We estimated the number of "effectively independent chromosome segments" (M_e , see Goddard 2009) is ≈ 85 on our 50 Mb genome. Therefore, when the number of QTL=50, GBLUP works as well as BayesR in Gen=0, because nearly all segments contain a QTL and so the prior assumption that chromosome segment effects are normally distributed is approximately correct. Also, when animals are relatively closely related (Gen=0) there was no advantage for SEQ because HD and MD SNP are dense enough to predict the QTL effects given the low M_e.

Although BayesR analysis estimates an effect for each SNP with many set to zero, the method still has difficulty defining which SNP within a segment of high LD is the true QTL, and several SNP effects are estimated as contributing to part of the QTL effect, particularly with dense SNP. With QTL=15 the BayesR method showed an advantage over GBLUP even in Gen=0, and in all scenarios the advantage of BayesR becomes more pronounced in Gen=10 and 15. This implies that even with many SNP in high LD, BayesR is superior to GBLUP in accurately attributing SNP effects to a more precise chromosome region harbouring the real QTL. Recombination is therefore less likely to occur between the true QTL and the SNP to which BayesR has attributed part of the QTL effect and accuracy of GP is more persistent across generations. The decay in accuracy is more rapid with GBLUP than BayesR because more SNP effects over longer segments are contributing to predicting the individual QTL effects and therefore there is a much higher chance that recombination will disrupt the LD between QTL and SNP alleles.

Druet *et al.* (2013) tried to indirectly estimate the effect of negative selection on accuracy of GP by simulating QTL only on loci with MAF<0.1 compared to their neutral model allocating QTL randomly across all loci. They observed ~10% reduction in SEQ accuracy of GP with BayesR when QTL MAF<0.1. However, our simulation demonstrates that the MAF distribution of QTL subjected to long term negative selection is unlikely to be as extreme as assumed in Druet *et al.* (2013). There was a consistent reduction in the accuracy of GP due to the effect of selection, but only when there were 10 or more generations separating the training and validation populations. If a gamma distribution of QTL effects had been used in this study, the difference between BayesR and GBLUP accuracies might have been more pronounced, particularly when the number of QTL=15 because this is closer to BayesR assumed distribution of QTL effects. However no further differences in the results would be expected.

To gain more advantage from sequence, we conclude that training data should be combined from more than one breed to reduce the LD between more distant SNP (equivalent to an increase in the N_e). This will also require an increase in the size of training populations but should ensure better persistency of GP accuracy across generations with SEQ, provided that a reasonable proportion of QTL are segregating in both breeds. It is also likely to be more beneficial to use a Bayesian analyses and to select a subset of potentially more biologically active SNP from sequence data prior to analysis.

REFERENCES.

Chadeau-Hyam M., Hoggart C., O'Reilly P., Whittaker J., De Iorio M. and Balding D. (2008) *BMC Bioinformatics* **9**: 364.

Clark S., Hickey J. and van der Werf J. (2011) *Genet. Sel. Evol.* **43**: 18.

Druet T., MacLeod I.M. and Hayes B.J. (2013) *Heredity*. http://dx.doi.org/10.1038/hdy.2013.13

Erbe M., Hayes B.J., Matukumalli L.K, Goswami S., Bowman P.J., Reich C.M., Mason B.A. and Goddard M.E. (2012) *J. Dairy Sci.* **95**: 4114.

Goddard M.E. (2009) *Genetica* **136**: 245.

Gilmour A.R., Cullis B.R., Gogel B.J., Welham S.J. and Thompson R. (2005) *ASReml User Guide Release 2.0*. VSN International Ltd, Hemel Hempstead, HP1 1ES, UK.

Keightley P. & Halligan D. (2009) *Genetica* **136**: 359.
Kemper K.E., Visscher P.M. and Goddard M.E. *Genome Biology* 2012, 13:244.
Macleod I.M., Larkin D., Lewin H., Hayes B.J. and Goddard M.E. (2013) *Mol. Bio. Evol.* (Published ahead of print: doi:10.1093/molbev/mst125)
Meuwissen T.H.E., Hayes B.J. and Goddard M.E. (2001) *Genetics* **157**: 1819.
Meuwissen T.H.E. and Goddard M.E. (2010) *Genetics* **185:** 623.

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:220-223

ACCURACY OF GENOMIC PREDICTION FROM MULTI-BREED SHEEP REFERENCE POPULATION

N. Moghaddar[1, 2, 4], A.A. Swan [1, 3] and J.H.J. van der Werf[1, 2]

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

SUMMARY

Genomic estimated breeding values (GEBV) were calculated based on a combination of purebred and crossbred sheep for birth weight, weaning weight and post weaning weight using genomic best linear unbiased prediction (GBLUP). The genomic relationship matrix (G) was calculated based on population wide or breed of haplotype specific allele frequency using the 50k ovine Illumina SNP-chip. The accuracy of genomic prediction was estimated based on the correlation between genomic breeding value and an accurate breeding value based on progeny records. The result showed better genomic prediction accuracy for breeds with higher representation in the combined reference populations. Accuracies slightly decreased when the reference set contained a significant set of additional animals from another breed. This study showed no extra accuracy from across breed information using 50k SNP marker panel. The result showed a small but non-significant increase in accuracy when using breed specific allele frequencies in the calculation of G.

INTRODUCTION

Genomic selection can significantly increase the rate of genetic progress in quantitative traits by providing extra accuracy from exploiting Mendelian sampling and by reducing the generation interval (Meuwisen *et al*. 2001; Schaeffer 2006; Banks *et al*. 2009; Dalton 2009; van der Werf 2010). The size of the reference population has an important impact on the accuracy of genomic prediction (Goddard 2009). In the sheep industry data are often available from a mixture of breeds, multiple strains within a specific breed or from crossbreds. Combining populations of different pure and crossbred animals would be an advantage if it could be shown to increase the accuracy of genomic prediction, particularly for breeds which are not well represented in the combined reference population. The objective of this study is to assess the effect of a combined sheep reference population on accuracy of within breed genomic prediction using real data. The accuracy of genomic prediction was compared between GEBV prediction from purebred, crossbred and a combination of purebred and crossbred data which was extracted from a large multi breed/crossbred sheep reference population. Furthermore, two strategies in calculating the genomic relationship matrix (G) were compared to investigate the effect of accounting for different marker allele frequencies between breeds.

MATERIALS AND METHODS

Reference population and phenotypic data. The reference populations tested contained either purebred Merino sheep (M) or crossbreds of Border Leicester and Merino (BLxM), or a combination of both. Three population sizes (1000, 2000 and 3000) were used for the purebred Merino reference sets and these were compared with 3 sets where the purebred Merino populations were augmented with 1472 BLxM crossbreds. Data was extracted from the Sheep CRC Information Nucleus database (van der Werf *et al.* 2010). The traits investigated were birth weight

(BW), weaning weight (WW) and post weaning weight (PWW). Phenotypic means and standard deviations were 4.76 ± 1.02, 25.4 ± 5.78 and 40.2 ± 8.2 respectively.

Genotypes and validation population. Animals were genotyped using the 50K Ovine chip (Illumina Inc., SanDiego, CA, USA). This chip provided 48,559 SNP genotypes after applying quality control. The accuracy of GEBV was estimated as the correlation of GEBV and accurate EBV based on pedigree and phenotypes in an independent group of animals which had been genotyped for use as a validation population. The validation population comprised 175 Merino sires and 55 Border Leicester sires with average EBVs accuracies of 0.92 and 0.98, respectively. Comparison of correlation coefficient of two dependent samples was used as test statistics.

Statistical methods. Genomic best linear unbiased prediction (GBLUP) was used to calculate the GEBV using ASReml (Gilmour 2009). The following model was used for analysis of data: $y = Xb + Z_1g + Z_1Qq + Z_2m + e$ where y is a vector of phenotypes, b is a vector with fixed effects, g is the random additive genetic effect of the animal, q is a vector with random breed effects, m is a vector with maternal effects, and e is vector of random residual effects, X and Z_1 and Z_2 are incidence matrices and Q contains breed proportions as derived from a deep pedigree. g, q and e are considered normally distributed as $g \sim N(0, G\sigma_g^2)$, $q \sim N(0, Q\sigma_q^2)$, and $e \sim N(0, I\sigma_e^2)$, respectively, where G is the genomic relationship matrix. The fixed effects in the model were birth type, rearing type, gender, age at measurement (for weaning weight and post weaning weight) and contemporary group which was flock × birth year × management group. G was calculated using two approaches according to VanRaden (2008). In one approach G was calculated using the overall marker allele frequencies of the entire population (G1) while in the second approach the breed of haplotype specific marker allele frequencies were used (G2).

Table 1. Accuracy of genomic prediction from different reference populations for birth weight (BW) for Merino and Border Leicester (BL)

Reference population		Breed proportion (%)		GEBV accuracy[1]			
				G1		G2	
Type	Size	BL	Merino	BL	Merino	BL	Merino
(1) = Merino	1000	0.0	100	-0.03 [b]	0.38 [bc]	-0.03 [b]	0.38 [bc]
(2) = Merino	2000	0.0	100	-0.10 [ab]	0.42 [cd]	-0.10 [ab]	0.42 [cd]
(3) = Merino	3000	0.0	100	-0.16 [a]	0.47 [d]	-0.14 [a]	0.47 [d]
BLxMerino	1472	50.7	47.2	0.24 [c]	0.29 [a]	0.24 [c]	0.29 [a]
BLxMerino + (1)	2472	30.1	68.3	0.23 [c]	0.36 [b]	0.24 [c]	0.39 [bc]
BLxMerino + (2)	3472	21.4	77.3	0.17 [c]	0.39 [bc]	0.17 [c]	0.39 [bc]
BLxMerino + (3)	4472	16.6	82.4	0.18 [c]	0.42 [cd]	0.18 [c]	0.42 [cd]

G1: Genomic relationship matrix based on all SNP allele frequency. G2: Genomic relationship matrix based on haplotype SNP allele frequency. Different superscripts for accuracies indicate statistical differences.

RESULTS AND DISCUSSION

Tables 1, 2 and 3 show the GEBV accuracy for BW, WW and PWW according to the two methods to calculate G respectively. The results show that the GEBV accuracy for Merino sheep

increased as the size of the reference population increases, both for purebred and combined purebred and crossbred animals. However, the accuracies were higher when prediction was based on purebred Merinos, and accuracy slightly decreased if the BLxM batch of crossbred animals was added to a purebred Merino reference population.

The GEBV accuracy for BL validation animals in all three weight traits is higher when the genomic prediction is based on a population with a maximum proportion of BL animals. Adding additional purebred Merino animals to the BLxM reference population reduces the accuracy for BL validation animals. Differences were not statistically significant for BW but they were significant for WW and PWW ($P<0.05$).The results also show a slightly higher GEBV accuracy in some cases from using breed specific marker allele frequency (G2) compared to overall population marker allele frequency (G1) for construction of the genomic relationship matrix. However, most of these differences were not significant ($P<0.05$).

The results suggest that the genomic prediction accuracy within a specific breed is mainly determined by the effective number of haplotypes of that breed in the reference population. The accuracy of GEBV for Merinos increased based on prediction from a larger reference population. The rate of increase in accuracy as well as the level of accuracy in Merino was lower when prediction was based on a combination of crossbred and purebred Merinos compared to prediction from only purebred Merinos from a similar population size. This indicates neutral to some negative effect of adding BL haplotypes to the reference population. The accuracy of GEBV for BL when predicted from combined BLxM and purebred Merinos decreased with a decreasing proportion of BL haplotypes, indicating a negative effect of Merinos on accuracy of genomic prediction for BL animals. Genomic prediction from BLxM on their own provides some predictive power for Merinos because all progeny used had Merino dams.

Table 2. Accuracy of genomic prediction from different reference population for weaning weight (WW) for Merino and Border Leicester (BL)

Reference population		Breed proportion (%)		GEBV accuracy[1]			
				G1		G2	
Type	**Size**	**BL**	**Merino**	**BL**	**Merino**	**BL**	**Merino**
(1) = Merino	1000	0.0	100	-0.07 ^{b}	0.42 ^{b}	-0.06 ^{b}	0.42 ^{b}
(2) = Merino	2000	0.0	100	-0.13 ^{b}	0.49 ^{c}	-0.13 ^{b}	0.49 ^{c}
(3) = Merino	3000	0.0	100	-0.26 ^{a}	0.51 ^{c}	-0.22 ^{a}	0.51 ^{c}
BL*Merino	1547	50.0	47.6	0.32 ^{d}	0.31 ^{a}	0.32 ^{d}	0.31 ^{a}
BL*Merino + (1)	2547	30.3	67.6	0.22 ^{c}	0.43 ^{b}	0.24 ^{c}	0.41 ^{b}
BL*Merino + (2)	3547	22.1	76.8	0.16 ^{c}	0.46 ^{b}	0.18 ^{c}	0.45 ^{b}
BL*Merino + (3)	4547	17.0	82.3	0.17 ^{c}	0.47 ^{bc}	0.18 ^{c}	0.44 ^{b}

[1] As defined in Table 1

The degree of relationship between validation and reference population animals affects the accuracy of genomic prediction (Habier *et al.* 2007) and therefore genomic relationships between reference and validation populations were explored. There was on average a low to moderate genomic relationship between Merino validation animals and the reference populations while it

was almost close to zero between BL validation sires and the purebred Merino reference population.

This study showed that the accuracy of GEBV prediction from a multi breed reference population depends highly on breed representation in the reference population, both through numbers and proportion. Daetwyler *et al.* (2010) showed that across breed information does not contribute to genomic prediction accuracy using the 50k marker density. This study showed neutral to negative effect of adding information from animals of a different breed. Applying denser SNP marker panels could potentially lead to better prediction from across breed information. More investigations with larger validation population and also with denser genetic markers are required.

Table 3. Accuracy of genomic prediction from different reference population for post weaning weight (PWW) for Merino and Border Leicester (BL)

Reference population		Breed proportion (%)		GEBV accuracy			
				G1		G2	
Type	**Size**	**BL**	**Merino**	**BL**	**Merino**	**BL**	**Merino**
(1) = Merino	1000	0.0	100	-0.02 [a]	0.53 [b]	0.00 [a]	0.53 [b]
(2) = Merino	2000	0.0	100	-0.04 [a]	0.57 [bc]	-0.04 [a]	0.57 [bc]
(3) = Merino	3000	0.0	100	-0.08 [a]	0.59 [c]	-0.07 [a]	0.59 [c]
BL*Merino	1514	50.7	47.2	0.49 [c]	0.45 [a]	0.49 [b]	0.45 [a]
BL*Merino + (1)	2514	30.5	68.2	0.42 [bc]	0.56 [bc]	0.47 [b]	0.57 [bc]
BL*Merino + (2)	3514	21.8	77.2	0.37 [b]	0.54 [bc]	0.42 [b]	0.57 [bc]
BL*Merino + (3)	4514	17.0	82.3	0.36 [b]	0.56 [bc]	0.41 [b]	0.57 [bc]

G1: Genomic relationship matrix based on overall SNP allele frequency. G2: Genomic relationship matrix based on each population SNP allele frequency.

ACKNOWLEDGEMENTS

The authors are thankful to Daniel Brown (AGBU/UNE) for consultations and providing industry sires ASBVs, Klint Gore for performing quality control and database management and project executive staff at the various research herds. Australian Wool Innovation and Meat and Livestock Australia made the data available from the SGP.

REFERENCES

Banks R.G. and van der Werf J.H.J. (2009) *Proc. Assoc. Advmt. Anim. Breed. Genet.* **18**:430.

Daetwyler H.D., *et al.* (2010) *Animal Production Science* **50:**1004.

Dalton R. (2009). *Nature* **457**: 369.

Gilmour A.R., *et al.* (2009) ASReml 3. ww.vsni.co.uk

Goddard M.E. (2009) *Genetica* **136:**245.

Habier D., *et al.* (2007) *Genetics* **177**:2389.

MeuwissenT., *et al.* (2001) *Genetics **157**:1819*

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

Van der Werf J.H.J. (2010) *Anim.Prod.Science* **50:** 998.

VanRaden P.M. (2008) *J. Dairy Sci* **91:** 4414.

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:224-228

USING MALE PERFORMANCE TO IMPROVE GENOMIC SELECTION FOR FEMALE FERTILITY IN BRAHMAN CATTLE

Y. D. Zhang[1], D. J. Johnston[1], S. Bolormaa[2], A. Reverter[3], M.R.S. Fortes[4] and B. Tier[1]

[1] Animal Genetics and Breeding Unit[*], UNE, Armidale, NSW; [2] Department of Primary Industries, Victoria; [3] CSIRO Animal, Food and Health Sciences, St. Lucia, Queensland; [4] Queensland Alliance for Agriculture and Food Innovation

SUMMARY

Genomic selection offers the opportunity to improve female fertility in the Northern Australian beef herd. However, genomic predictions for a number of female fertility traits - derived solely from the small number of female fertility records collected thus far - are only of modest accuracy. In this study measures of Brahman male reproduction were used jointly with the female records, increasing the accuracy of genomic predictions for Brahman female fertility. Scrotal circumference measured at 18 months was found to be the most useful to increase accuracies of cow GEBVs for a range of traits upto 22%.

INTRODUCTION

Improving cow fertility has the potential to increase the profitability of Northern Australian beef cattle enterprises. Reducing the age at which heifers reach puberty, and/or increasing the probability of post-partum reconception in subsequent matings can both lead to improved calving rates. Johnston *et al.* (2010) found age at puberty and post-partum anoestrous interval (PPAI, defined as the time from calving to cycle) were moderately to highly heritable in Brahman cows. Early reproduction (measured as the number of calves in the first two opportunities) and lifetime reproduction (number of calves in the first six opportunities) were shown to be lowly heritable in tropical genotypes (Brahman and Tropical Composite) (Johnston *et al.* 2013a). With development of genomic markers, genomic selection could play an important role in genetic improvement. Zhang *et al.* (2013) demonstrated the usefulness of genomic selection for various measures of female fertility; however, accuracies of genomic breeding values, derived from a data set of limited size, were low. Amongst the reproduction traits measured in tropical beef bulls, scrotal circumferences at different ages were found to be highly heritable (Corbet *et al.* 2013) and correlated with female traits (Johnston *et al.* 2013b). This study examines whether the accuracy of genomic breeding values of reproduction traits of Brahman cows could be increased by using scrotal circumference information from their male relatives.

MATERIALS AND METHODS

Animal and measurements. The Brahman bulls and cows used in this study were part of the 'Northern Breeding Project' resource population, bred by the Cooperative Research Centre for Beef Genetic Technologies (Beef CRC) in the tropical regions of Northern Australia (Burrow *et al.* 2003; Barwick *et al.* 2009). A total of 1035 females were phenotyped. The first postpartum anoestrous interval (PPAI) records were observed on the 635 cows that calved at their first opportunity. The cows were progeny of 54 sires (Barwick *et al.* 2009). Age at puberty (AP) was defined as the age when the first c*orpus luteum* (CL) was observed using regular ultrasound scanning. Also, up to 6 calving occurrences were recorded for cows. These observations were used to determine the following fertility traits: 1) PPAI1 - the first PPAI, 2) CR12 - calves born in the

[*]AGBU is a joint venture of The NSW Department of Primary Industries and University of New England.

first two opportunities, 3) WR12 - number of calves weaned in the first two opportunities, 4) LP – lactating-pregnancy status (a binary trait scored as "1" for cows both lactating and pregnant, otherwise "0"), 5) ACR – average calving rate in the first six opportunities and 6) AWR - average weaning rate in the first six opportunities. A comprehensive description of the bull data was provided by Corbet *et al.* (2013). Bulls were born between 2004 and 2010 and were progeny of the cows measured above. Scrotal circumference (SC) measured at ages of 12 (SC12) and 18 months (SC18) of 1142 bulls born from 2004 to 2008 were used in this study.

Genotypes. The SNP genotype data used in this study was a subset of Beef CRC genomic dataset. Details on genotyping, editing and imputation of the Beef CRC genomic data set has been described by Bolormaa *et al.* (2013). Briefly, 49, 821 and 126 cows were genotyped on the Illumina BovineSNP 7K, 50K and 700K SNP platforms (www.illumina.com/agriculture), respectively. The bulls were genotyped with the 50K platform. Genotypes with poor GenCall scores, very low minor allele frequencies and significant deviation from Hardy-Weinberg equilibrium were deleted. Missing genotypes for animals genotyped with the less dense chips were imputed to 700K using BEAGLE (Browning and Browning, 2009). Thus genotypes of 729,068 SNP for 996 Brahman cows and 1118 Brahman bulls were available for subsequent analyses.

Statistical methods. Genetic parameters were estimated for all traits of cows and bulls, with all phenotypic records using pedigree based REML (Wombat, Meyer 2007). Models for all cow traits (Barwick *et al.* 2009, Johnston *et al.* 2009, Johnston *et al.* 2010, Johnston *et al.* 2013a) and for scrotal size of bulls (Corbet *et al.* 2013) were described previously and used in this analysis.
GBLUP Genomic estimated breeding values (GEBVs) for each trait were estimated for animals with genotypes only using Wombat (Meyer 2007). The GBLUP model was as $\mathbf{y = Xb + Zg + e}$, where the phenotype (**y**) is a function of systematic effects (**b**), breeding values (**g**) and residuals (**e**), with incidence matrices (**X, Z**) assigning observations to effects. Covariances among breeding values were modeled with $\mathbf{G}\sigma_g^2$ – where **G** is the genomic relationship matrix (Yang *et al.* 2010) and σ_g^2 the genetic variance – and among residuals with $\mathbf{I}\sigma_e^2$ – where **I** is an identity matrix and σ_e^2 is the residual variance. SNP with very low minor allele frequencies (<0.005) were excluded when calculating **G**

.

Cross validation. A 5-fold internal cross-validation procedure was carried out for each female trait. Genotyped cows were divided into 5 approximately equal groups, with 4 subsets used as a training set to predict the 5th subset as the test set. Animals were grouped so that complete paternal half-sib families were in the same subset**.** Cross validation was carried out using univariate and multi-variate analyses with bull scrotal circumferences (SC12 and SC18) as the other trait(s). Each of the cow traits were analysed jointly with either SC12 or SC18 of bulls in bivariate analyses and with both SC12 and SC18 in trivariate analyses.

Accuracy. GEBVs for test animals were correlated with their phenotypes adjusted for systematic effects. Accuracies were calculated as r/h where r is the correlation coefficient between GEBVs and phenotypic values and h is the square root of the heritability of the trait (estimated using all phenotypes). The average of 5 accuracies from the cross validations is presented as the accuracy for genomic prediction of each trait.

RESULTS AND DISCUSSION

The average accuracies and their standard errors of GEBVs for reproduction traits of cows in univariate and bivariate analyses with SC12 or with SC18 of bulls are shown in Table 1. Univariate analyses showed high accuracies of GEBVs for AP and AWR but low values for

PPAI1 and ACR. Accuracies of GEBVs from the bivariate analyses with SC18 were higher than the corresponding univariate values, the highest being for PPAI1 which increased from 0.18 to 0.22. Increases in accuracies of GEBVs were found for those traits with relatively low accuracies from univariate analyses.

However, results from bivariate analyses with SC12 were mixed. Accuracies for AP and PPAI1 were higher than the corresponding univariate results,notable reductions in accuracies were observed for remaining traits. Both SC12 and SC18 were expected to contribute similarly to the accuracies of cow fertility traits as SC12 and SC18 have a genetic correlation of 0.95 (Corbet *et al.* 2013). The genetic correlations between AP or PPAI1 and SC12 were similar to those with SC18, but they were low and not significantly different from zero for others traits with SC12. These mixed accuracies may be related to the low genetic correlations between SC12 and cow traits in genotyped data. Most of the genetic correlation coefficients were associated with large standard errors. These results were in line with results by Johnston *et al.* (2013b). The heritability estimate for SC12 (0.65) was lower than that for SC18 (0.75) (Corbet *et al.* 2013).

Table 1. Accuracies of genomic breeding values (standard errors) of cows in univariate and bivariate analyses with SC12 or with SC18 of bulls.

Trait	Univariate	Bivariate with SC12		Bivariate with SC18		$h^{2\#}$
		Accuracy	Change*	Accuracy	Change*	
AP	0.33 (0.06)	0.38 (0.07)	+16%	0.35 (0.09)	+6%	0.56
PPAI1	0.18 (0.05)	0.19 (0.05)	+6%	0.22 (0.06)	+22%	0.51
CR12	0.25 (0.09)	0.21 (0.11)	-16%	0.28 (0.09)	+12%	0.15
WR12	0.24 (0.07)	0.18 (0.08)	-25%	0.25 (0.07)	+4%	0.21
LP	0.20 (0.05)	0.19 (0.04)	-5%	0.21 (0.05)	+5%	0.39
ACR	0.16 (0.06)	0.10 (0.07)	-37%	0.18 (0.09)	+13%	0.16
AWR	0.39 (0.06)	0.32 (0.05)	-18%	0.40 (0.08)	+3%	0.13

*the percentages of change are based on average accuracies from corresponding univariate analyses. #h^2 from analysis of complete phenotypic data.

Table 2. Average accuracies (standard errors) of GEBV for reproduction traits of cows in trivariate analyses with SC12 and SC18 of bulls.

Trait	Accuracy	Change*
AP	0.37 (0.09)	+12%
PPAI1	0.20 (0.06)	+11%
CR12	0.24 (0.09))	-4%
WR12	0.24 (0.07)	0%
LP	0.21 (0.05)	+5%
ACR	0.13 (0.10)	-19%
AWR	0.33 (0.10)	-15%

*the percentage of changes are based on average accuracies from corresponding univariate analyses.

The average accuracies and their standard errors of GEBVs for cow traits with bull scrotal sizes in the trivariate analyses are shown in Table 2. Change in accuracies from trivariate analyses appeared to be within the ranges observed from bivariate analyses with SC12 and with SC18 (Table 1). Use of both SC12 and SC18 enhanced the accuracies for GEBV of AP and PPAI1 up to 12%. The changes for accuracies of early life time and life time reproduction traits were very small or negative.

These results suggest that the inclusion of scrotal circumference measures from male relatives can enhance the accuracy of GEBVs for female fertility in Northern Australian Brahman cattle. However, their use is limited because the genetic correlations between scrotal measure (SC18) and the female fertility traits ranged from low (0.18) to medium (0.49). More training data is required to increase the accuracies cow GEBVs.

CONCLUSIONS

This study shows that use of reproduction phenotypes and genotypes of bulls can improve the accuracies of genomic selection for traits measured in cows. Incorporating scrotal circumference of bulls can improve accuracies of GEBV for AP and PPAI, up to 22%. Scrotal circumference measured at 18 months was found to be most useful. Results suggest that the use other source of information such as bull fertility measures and increasing quality of phenotypes and records of training population can enhance accuracy of genomic selection.

ACKNOWLEDGEMENTS

We would like to acknowledge colleagues from CSIRO, DAFFQ, UQ and AGBU for their contribution to data collection and preliminary editing. The significant efforts of other project scientists and support staff located throughout the former Beef CRC's network are also acknowledged.

REFERENCES

Corbet N.J., Burns B.M., Johnston D.J., Wolcott M.L., Corbet D.H., Venus B.K., Li Y., McGowan M.R. and Holroyd R.G. (2013) *Anim. Prod. Sci.* **53**:101.

Bolormaa S., Pryce J.E., Kemper K., Savin K., Hayes B.J., Barendse W., Zhang Y., Reich C., Mason B., Bunch R.J., Harrison B.E., Reverter A., Herd R.M., Tier B. and Goddard M.E. (2013) *J. Anim. Sci.*91:3088.

Barwick S.A., Wolcott M.L., Johnston D.J., Burrow H.M., Sullivan M.T. (2009) *Anim. Prod. Sci.* **49**:351.

Browning B.L. and Browning S.R. (2009) *Am. J. Hum.Genet.* **84**: 210.

Burrow H.M., Johnston D.J., Barwick S.A., Holroyd R.G., Barendse W., Thompson J.M., Griffith G.R. and Sullivan M. (2003) *Proc. Assoc. Advmt. Anim. Breed. Genet.* **15**:359.

Johnston D.J., Barwick S.A., Corbet N.J., Fordyce G., Holroyd R.G., Williams P.J. and Burrow H.M. (2009) *Anim. Prod. Sci.* **49**: 399.

Johnston D.J., Barwick S.A., Fordyce G. and Holroyd R.G. (2010) *9th World Congr. Genet. Appl. Livestock Prod 923*.

Johnston D.J., Barwick S.A., Fordyce G., Holroyd R.G., Williams P.J. and Corbet N.J. (2013a) *Anim. Prod. Sci. (in press)*.

Johnston D.J., Corbet N.J., Barwick S.A., Wolcott M.L. and Holroyd R.G. (2013b) *Anim. Prod. Sci. (in press)*.

Meyer K. (2007) *J. of Zhejiang Uni., SCI. B* **8**: 815.

Taylor W.J. and Rudder T.H. (1986) *Proc. Aust. Soci. Anim. Prod.* **16**: 379.

Yang J., Benyamin B., McEvoy N.P., Gordon S., Henders A.K., Nyholt D.R., Madden P.A., Heath A.C., Martin N.G., Montgomery G.W., Goddard M.E. and Visscher P.M. (2010) *Nat. Genet.* **42**: 565-569.

Zhang Y.D., Johnston D.J., Bolormaa S., Hawken R.J. and Tier B. (2013) *Anim. Prod. Sci. (in press)*.

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:229-232

IMPLICATIONS OF GENETIC ARCHITECTURE ON THE EFFICACY OF GENOMIC SELECTION

F.S. Hely[1], S.A. Clark[2] and P.R. Amer[1]

[1]AbacusBio Limited, P O Box 5585, Dunedin 9058, New Zealand
[2]School of Environmental and Rural Science, University of New England, Armidale, NSW, 2351, Australia

SUMMARY

A simulation model is described which has been constructed to address the issue of how true underlying genomic architecture might impact on the efficacy of genomic selection. A current specific focus of the model is on how epistatic genetic architectures might impact on the added value expected from increasing the density of SNP markers. Results to date suggest that genomic selection has greater superiority over BLUP genetic prediction under the additive genetic architecture simulated relative to an epistatic architecture with similar heritability. While we expect marker density to improve accuracy under GBLUP with some additive genetic architectures, our simulation results suggest that this may not happen with comparable (in terms of narrow sense trait heritability) genetic architectures with epistatic gene action contributing to both additive and non-additive genetic variance.

INTRODUCTION

The underlying genetic architecture of economically important traits in sheep remains unclear. There is a reasonable body of biological evidence (Gianola and de los Campos 2008) that suggests interacting genetic loci (i.e. epistatic loci) are a significant source of genetic variation. It is yet to be determined how single step genomic selection will perform when epistatic effects among loci contribute significantly to underlying additive genetic variation. The genomic best linear unbiased prediction (GBLUP) method of genomic selection assumes each SNP marker has an equal effect on trait variance and uses information from the genomic relationships between candidates to estimate the merit of genotyped candidates as opposed to alternative Bayes methods which use the effects of minor and major genes weighted differently. In this study, simulation work was under taken to model the application of single step genomic selection methodology to the New Zealand sheep industry using a combination of low and high density SNP panels. A set of QTL were simulated, and the accuracy of prediction using both conventional BLUP genetic evaluation and the single step GBLUP genetic evaluation was compared with and without epistatic genetic effects simulated for a single trait in a population resembling a major NZ dual purpose sheep breed.

METHODOLOGY

Population and SNP data were simulated using the QMSim software developed by Sargolzaei and Schenkel (2009). The parameters used in the simulations are shown in Table 1. These parameters were chosen to try and generate a population with similar characteristics to the major New Zealand dual purpose sheep breeds. The QMSim software uses a two stage method for simulating a population; a historical phase and a recent population phase. The historical phase uses random mating over a large number of generations to create linkage disequilibrium and drift in a base population. The recent population phase is used to create the desired population structure for analysis, no mutation occurs and the allele effects are fixed at the end of the historical phase.

Table 1: Parameter estimates for the population simulated using QMSim

Parameter	Value
Effective population size for the historical phase	4000
Number of females per male in the historical phase	20
Number of generations for the recent population	60
Number of females per male in the recent population	50
Litter size in the recent population	50% single, 50% twins
Proportion of male progeny in the recent population	0.5
Replacement ratio for sires/dams	1.5 yrs/3 yrs
Number of chromosomes	26
Marker and QTL mutation rates	2.5×10^{-5}

Once the QMSim data were generated, epistatic and purely additive true breeding values were simulated for all individuals with marker data available. The additive true breeding values (TBV_{add}) were calculated using the sum of the allele effects provided by QMSim for 100 QTL segregating at the end of the historical phase. These QTL had additive effects which were sampled from a normal distribution. The epistatic true breeding value (TBV_{epi}) was calculated in a similar way. For *n* pairs of loci with epistatic effects simulated between them *n* 9x9 matrices of epistatic effects for all possible combinations of genotypes were simulated. For a given pair of loci *A* and *B* each with two alleles (*a* and *A*, *b* and *B* respectively) a matrix was created as below:

	bb	*bB*	*BB*
aa	e_{ab}	0	e_{aB}
aA	0	0	0
AA	e_{Ab}	0	e_{BB}

Thus, if an individual had the combination of genotypes *aa* and *BB* then $TBV_{epi} = TBV_{epi} + e_{aB}$. The epistatic effects *e* were drawn from a normal distribution. In order to compare genomic breeding values based on additive versus epistatic true breeding values, it was necessary to scale the variance of the true breeding values so that the additive genetic variance estimated by ASReml (Gilmour *et al.* 1999) was the same for both the epistatic and additive genetic models. i.e.

$$TBV^* = TBV \times \frac{\sqrt{h^2}}{\sigma_{TBV_N}}$$

where TBV* is the rescaled TBV, h^2 is the desired trait heritability, σ_{TBV_N} is the additive genetic standard deviation (narrow sense) estimated by ASReml. Phenotypes were then simulated as

$$PHEN = TBV^* + (1 - h^2 \times \frac{\sigma^2_{TBV_B}}{\sigma^2_{TBV_N}}) \times \delta$$

where σ_{TBV_B} is the standard deviation of the original TBVs in the broad genetic sense and δ is a random normal deviate with mean of 0 and standard deviation of 1. In this way, the two different

architectures are constructed in such a way that they would appear to be identical when undertaking variance component estimation using conventional quantitative genetic analysis.

A genomic best linear unbiased prediction GBLUP evaluation was run on the phenotypic values for both the additive and epistatic traits using the BLUPF90 family of programs (Misztal *et al.* 2002) with a SNP marker file. A traditional BLUP evaluation was also run using ASReml and the estimated breeding values from both evaluations were combined with the TBV and phenotypic data. Accuracies of genomic predictions were computed as the correlation between the additive and epistatic TBVs and their corresponding genomic estimated breeding values (GEBVs).

SNP panel densities from 10,000 to 100,000 were simulated, with the accuracies, measured as the correlation of the TBVs with the GEBVs and BLUP EBVs, for the different panel densities compared. The TBVs were scaled to give an additive genetic variance of 0.3.The number of QTL used to generate the additive and epistatic TBV remained constant at 100 for all scenarios. From QMSim, the marker data were retained for individuals generated in generations 57 to 60. For a training and validation trial, the individuals born in generation 57 had phenotypic data and all other individuals had a missing phenotype. Correlations between estimated breeding values and true breeding values are reported for animals from generation 60.

For all scenarios 20 replicates were run, where replication was performed by using the same base population markers and pedigree from QMSim for the 20 replicates, but with a new true genetic values for each replicate. Within each replicate, the GBLUP, Bayes Lasso and pedigree BLUP methods are applied to the exact same trait data with the same model.

RESULTS AND DISCUSSION

The accuracy of genomic selection (as indicated by correlations between predicted breeding values and true breeding value) exceeded the accuracy of BLUP genetic predictions for animals in the validation population which did not have their own phenotypic records (Table 2). BLUP genetic predictions appeared slightly more accurate under the additive model than under the epistatic model although the difference was not statistically significant. In contrast, genomic prediction was much more accurate under the additive model than under the epistatic model. Increasing the SNP density from 5k to 100k did not have any meaningful impact on the results with these genomic architectures and population structures.

Table 2: Correlations and the standard errors between true and estimated breeding values using GBLUP (TBV-GEBV) and traditional BLUP (TBV-EBV) for additive and epistatic traits, along with the heritability as estimated by ASReml with the standard error (simulated heritability was 0.3 for all scenarios).

Panel Size	Additive			Epistatic		
	TBV- GEBV	TBV-EBV	herit	TGV- GEBV	TGV-EBV	herit
5k	0.65 (0.004)	0.36 (0.002)	0.38 (0.015)	0.6 (0.009)	0.28 (0.013)	0.3 (0.012)
10k	0.69 (0.005)	0.43 (0.004)	0.34 (0.012)	0.60 (0.01)	0.26 (0.012)	0.33 (0.012)
20k	0.75 (0.003)	0.45 (0.003)	0.36 (0.007)	0.62 (0.01)	0.31 (0.01)	0.3 (0.009)
50k	0.74 (0.002)	0.39 (0.005)	0.32 (0.007)	0.64 (0.009)	0.30 (0.014)	0.3 (0.012)
100k	0.75 (0.004)	0.49 (0.003)	0.30 (0.011)	0.60 (0.01)	0.33 (0.012)	0.28 (0.012)

We hypothesise that with further exploration of population structures and genomic architectures, we will find situations where increasing marker density will increase the accuracy of

genomic predictions under the additive genetic architecture, but they will be less beneficial under the epistatic genetic architecture. This is because similarity among relatives due to sharing equivalent epistatic gene combinations breaks down much more quickly over successive meiosis than similarity due to inheritance of similar additive genetic effects. It is acknowledged that some patterns within the results appear inconsistent with the relative small sizes of standard errors. We believe that this may be due to replication being undertaken with the same set of SNPs.

CONCLUSION

If our hypothesis is confirmed through further work, then new approaches other than GBLUP, Bayes predictions, and single step genetic evaluation may be required to capture the full benefits from increased marker density when traits whose observed narrow sense heritability is driven by epistatic effects. Alternatively, the failure of Bayes methods, and increased marker density to meaningfully improve the accuracy of genomic selection in many practical situations tested to date, could be further evidence that epistasis is an important contributor to observed heritability in livestock populations. The alternative theory of many genes with very small effects has led to considerable, but so far fruitless, efforts to use increasingly dense marker chips to improve genomic selection both within and across breeds beyond what can be achieved with moderate density chips (e.g. 50k).

ACKNOWLEDGEMENTS

We gratefully acknowledge funding for this work by Ovita Limited.

REFERENCES

Gianola, D. and de los Campos, G. (2008) *Genetics Research* **90**: 525.

Gilmour A.R., Cullis B.R., Welham S.J. and Thompson R. (1999*) NSW Agriculture Bulletin* No. **3**.

Misztal I., Tsuruta S., Strabel T., Auvray, B., Druet T. and Lee D. (2002) *Proc.7th WCGALP* 28.

Sargolzaei, M. and Schenkel F.S. (2009) *Bioinformatics* **25**: 680.

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:233-236

THE IMPACT OF MEASURING ADULT FLEECE TRAITS WITH GENOMIC SELECTION ON ECONOMIC GAIN IN MERINO SELECTION INDEXES

A.A. Swan and D.J. Brown

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

SUMMARY

Stochastic simulation of a Merino sheep breeding program showed that measurement of adult fleece weight and fibre diameter, the two key adult production traits in Merino sheep, increased economic gain compared to measuring yearling expressions of the traits alone. Comparing three different selection indexes, gain increased in fleece weight by up to $1.10 per ewe per year over 10 years of selection, depending on the importance of the trait in the selection index. For fibre diameter the increase in gain was lower, to a maximum of $0.70 per ewe per year, because genetic correlations between yearling and adult performance are higher for fibre diameter. There was little benefit in multiple adult measurements of these traits, and since the Australian sheep industry's evaluation system already accommodates one adult measurement, most of the gains possible can be realised by breeders. Genomic selection of young rams resulted in further increases in gain when combined with adult measurements, particularly for fleece weight.

INTRODUCTION

There are perceptions among Merino breeders in Australia that the industry's genetic evaluation service MERINOSELECT (Brown *et al.* 2007) places too much emphasis on the performance of young animals at the expense of lifetime productivity. These perceptions are related to the fact that although estimated breeding values (EBVs) are available for lifetime productivity traits and these are included in selection indexes, only small numbers of animals are measured at ages beyond one year of age.

A companion paper in these proceedings (Brown *et al*. 2013) estimates genetic parameters for lifetime wool production of the two key economic traits, fleece weight and fibre diameter. In this paper we use those parameter estimates to quantify the impact on economic gain of including varying amounts of information on lifetime production in selection indexes, both with and without genomic selection.

MATERIALS AND METHODS

Selection indexes. Predictions of economic gain were made from stochastic simulations of breeding programs with three Australian industry standard selection indexes with varying amounts of emphasis on fleece weight and fibre diameter: Index1 (M3.5) with high emphasis on increasing fleece weight while maintaining diameter, Index2 (M7SS) with balanced emphasis on increasing fleece weight and reducing fibre diameter, and Index3 (M14SS) with high emphasis on reducing fibre diameter while maintaining fleece weight. All three of these indexes separate wool traits into yearling and adult expressions, but treat all adult expressions as single traits. The indexes were modified to treat adult fleece weight or fibre diameter as separate traits between two and five years of age by multiplying the economic value for either trait by the proportion of wool harvested in each age class. Assuming equal fleece weights in each age class, these proportions were 0.28, 0.25, 0.24, and 0.23 for two, three, four and five year old ewes respectively. Other assumptions in deriving these economic values were that the price of wool is the same for all age classes, and there is no adult wether flock.

* AGBU is a joint venture of NSW Dept. of Primary Industry and the University of New England

Breeding program simulations. Breeding programs were simulated with 300 ewes mated annually to 10 rams. Age at first lambing was two years for both ewes and rams, and both sexes were given a maximum of eight mating opportunities (during the selection phase animals were culled on genetic merit). A realistic model of flock dynamics was used which included assumptions of mortality rates across ages in rams and ewes, and fertility, litter size, and lamb survival across ages in ewes. The average number of lambs weaned per ewe joined across ewe ages was 0.83 in the base flock.

True breeding values were simulated as the sum of mid-parent breeding values and Mendelian sampling random deviates drawn from the multivariate normal distribution defined by the genetic covariance matrix of all index traits, appropriately adjusted for inbreeding. Genetic covariances for adult fleece weight and fibre diameter between two and five years of age were as estimated by Brown *et al.* (2013). All other covariances were derived from the genetic covariance matrix used in MERINOSELECT. In addition to fleece weight and fibre diameter, the traits simulated included fibre diameter coefficient of variation, staple strength, body weight, and reproduction rate as defined in the selection indexes.

Phenotypes were simulated by summing true breeding values with random deviates sampled from the multivariate distribution based on a residual covariance matrix constructed by combining parameters estimated by Brown *et al.* with MERINOSELECT parameters as above.

Ten years of random selection were carried out to stabilise the flock, followed by fifteen years of selection on estimated index values. These were constructed from estimated breeding values calculated from multi-trait animal model BLUP analyses with varying amounts of phenotypic information: Y, including yearling measurements of clean fleece weight, fibre diameter, CV of fibre diameter, and body weight measured on both sexes; A2, adding adult clean fleece weight measured on ewes at two years to Y; A3 adding clean fleece weight on ewes at three years to A2; A4 adding clean fleece weight on ewes at four years to A3; and A5 adding clean fleece weight on ewes at five years to A4.

Genomic selection was added to the breeding program by including GBV “phenotypes” for all young ram selection candidates to BLUP analyses as an additional trait, as described by Swan *et al.* (2011). With this method, the genomic information contributes to increased accuracy of traits in the index during BLUP analyses via the genetic correlations between GBV and index traits. Modification of index values constructed from EBVs is not necessary. The accuracy of the GBV as a predictor of the target trait (either adult fleece weight or fibre diameter) was assumed to be 0.5.

The five measurement scenarios described above form the basis for comparisons in this study. They were run both with and without genomic selection, and repeated separately for adult fleece weight and fibre diameter at each age i.e., A2 adds adult fibre diameter at 2 years of age to Y etc.

Selection was by truncation on estimated index value across age classes, allowing the development of optimal age structures of males and females. The BLUP analyses were performed “annually” including all phenotypes available at the time. This means that animals were regularly selected before they had adult trait measurements. The ability of this method to match the timing of trait expressions with selection decisions is the reason why stochastic simulation was used in this study.

Equilibrium economic gains. One hundred replicates of each scenario were simulated, and mean true breeding values saved for all traits by year of birth. Annual rates of gain for each trait at equilibrium were then calculated as the slope of the regression of mean true breeding value on year of birth for the last ten years of the breeding program. Economic gains were then calculated by multiplying economic values by trait gains. Individual economic gains for each age class were summed to calculate total adult gains, weighted by the proportion of wool harvested in each age class as described above.

RESULTS

Economic and genetic gains for total adult fleece weight over ten years of selection are shown in Tables 1 and 2 respectively. Adding adult measurements of fleece weight increased economic gain by \$1.10 per ewe per year for Index1 (a 41% increase), \$0.50 for Index2 (a 33% increase), and by \$0.50 for Index3, compared to the industry standard of measuring only at the yearling stage. Most of the gain was captured with a single adult measurement (A2), although there was some additional benefit in recording at later ages, in particular for Index2. Genomic selection with yearling only measurement increased economic gain by up to \$0.40 for Index1, \$0.30 for Index2 and \$0.20 for Index3, lower than adding adult measurements. The highest economic gains were realised when combining adult measurements with genomic selection (up to \$1.50 per ewe for A3 with Index1). Genetic gains in Table 2 demonstrate the basis for economic gains, with similar increasing patterns of gain.

Table 1. Economic gain (\$/ewe/year) over 10 years for adult clean fleece weight under different selection scenarios with (+) and without (-) genomic selection (GS)

Index	Emphasis	GS	Y	A2	A3	A4	A5
1	(FW↑ FD↔)	-	2.70	3.80	3.60	3.70	3.60
		+	3.10	3.70	4.20	4.10	4.00
2	(FW↑ FD↓)	-	1.50	2.00	2.20	2.30	2.30
		+	1.80	2.60	2.70	2.60	2.70
3	(FW↔ FD↓)	-	-0.30	0.20	0.00	0.10	0.50
		+	-0.10	0.60	0.60	0.50	0.40

Table 2. Genetic gain (%) over 10 years for adult clean fleece weight under different selection scenarios with (+) and without (-) genomic selection (GS)

Index	Emphasis	GS	Y	A2	A3	A4	A5
1	(FW↑ FD↔)	-	9.4	12.3	12.3	12.2	12.6
		+	10.5	12.5	14.1	13.8	13.3
2	(FW↑ FD↓)	-	5.1	6.9	8.0	8.4	8.3
		+	6.4	9.1	9.7	9.1	9.6
3	(FW↔ FD↓)	-	-0.9	0.7	0.3	0.5	1.4
		+	-0.2	2.2	1.9	1.7	1.5

Economic and genetic gains for adult fibre diameter are shown in Tables 3 and 4 respectively. Adding adult measurements of fibre diameter had less impact on economic gain than fleece weight in absolute terms, with a maximum increase relative to yearling only measurement of \$0.70 per ewe per year when fibre diameter was most important in Index3 (a 14% increase).

Gain increased by \$0.40 per ewe per year for Index2 (a 50% increase), but was unchanged for Index1. Likewise, genomic selection was only beneficial for Index3, in which yearling only measurement with genomic selection was as effective as the adult measurement strategies. In the majority of cases, there was no benefit in measuring adult fibre diameter more than once, with the exception of Index3 with genomic selection.

Table 3: Economic gain ($/ewe/year) over 10 years for adult fibre diameter under different selection scenarios with (+) and without (-) genomic selection (GS)

Index	Emphasis	GS	Y	A2	A3	A4	A5
1	(FW↑ FD↔)	-	-0.10	0.00	-0.10	0.00	0.00
		+	-0.10	-0.10	-0.10	0.00	0.00
2	(FW↑ FD↓)	-	0.80	1.20	1.20	1.20	1.10
		+	0.80	1.20	1.10	1.40	1.40
3	(FW↔ FD↓)	-	5.00	5.70	5.70	5.70	5.60
		+	5.70	5.60	6.20	6.30	6.40

Table 4: Genetic gain (microns) over 10 years for adult fibre diameter under different selection scenarios with (+) and without (-) genomic selection (GS)

Index	Emphasis	GS	Y	A2	A3	A4	A5
1	(FW↑ FD↔)	-	0.1	0.0	0.1	0.0	0.0
		+	0.1	0.1	0.1	0.0	0.0
2	(FW↑ FD↓)	-	-0.4	-0.6	-0.6	-0.6	-0.6
		+	-0.4	-0.6	-0.6	-0.7	-0.7
3	(FW↔ FD↓)	-	-1.2	-1.4	-1.4	-1.4	-1.3
		+	-1.4	-1.3	-1.5	-1.5	-1.5

DISCUSSION

For the genetic correlations estimated by Brown *et al.* (2013), adding adult wool measurements to the breeding program increases economic and genetic gain, more so for fleece weight than fibre diameter. This is because the genetic correlations between ages are lower for fleece weight than fibre diameter. However, there was little benefit in measuring more than one adult expression, especially for fibre diameter. Consequently, because the Merino genetic evaluation system already accommodates one adult expression of these traits, breeders can already capture most of the benefits possible. Genetic gains would be increased throughout MERINOSELECT if more breeders recorded these traits.

A limitation on the genetic gain which can be made in adult wool traits is that measurement takes place after the most intense selection point (selection of young rams), and only on females selected for breeding. Genetic gains shown in Tables 2 and 4 are in fact lower than gains in equivalent yearling traits (results not shown), despite the fact that the adult measurements have higher heritabilities and phenotypic variances (Brown *et al.* 2013). This means that these traits are candidates for genomic selection, and the results of this study confirm that genomic selection for adult wool traits has benefits even when the traits are measured in the breeding program. These results support the findings of Van der Werf (2009) that the main benefit of genomic selection in Merinos is increased genetic gain in adult wool traits.

ACKNOWLEDGEMENTS

This work was funded by MLA Project B.SGN.0127.

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.

Brown D.J., Swan A.A. and Gill J.S. (2013) *Proc. Assoc. Advmt. Anim. Breed. Genet.* **20**: this publ

Swan A.A., Johnston D.J., Brown D.J., Tier B. and Graser H-U. (2012) *Anim. Prod. Sci.* **52:** 126.

Van der Werf J.H.J. (2009) *Proc. Assoc. Advmt. Anim. Breed. Genet.* **18:** 38.

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:237-240

GENETIC RELATIONSHIPS BETWEEN LAMB SURVIVAL AND MEAT TRAITS

F.D. Brien[1,2], D.L. Rutley[1,2], S.I. Mortimer[1,3] and J.E. Hocking-Edwards[1,4]

[1]Sheep CRC and [2]University of Adelaide, Roseworthy, SA, 5371
[3]NSW Department of Primary Industries, Trangie Agricultural Research Centre, Trangie, NSW, 2823
[4]South Australian Research and Development Institute**,** Struan Research Centre, Naracoorte, SA, 5271

SUMMARY

Correlations between survival traits (expressed by the lamb) and meat traits were estimated from analyses of four years of data (2007-2010) from the Sheep CRC's Information Nucleus, with records from 20,498 lambs, up to 8,596 dams and 377 sires. Tissue depth at the GR site and eye muscle depth had positive genetic correlations with lamb survival of 0.34±0.05 and 0.17±0.07, respectively, while the genetic correlations of lamb survival with lean meat yield and shear force were unfavourable (-0.33±0.06 and 0.27±0.07, respectively). Selection programs that enhance lean meat yield and reduce tissue depth at the GR site and increase tenderness need to consider the possibility of small correlated genetic losses in lamb survival, although appropriate index selection should be able to manage this risk, as the correlations were low. Conversely, genetic increases in tissue depth at the GR site may be correlated with small improvements in lamb survival.

INTRODUCTION

As poor lamb survival is a major contributor to sheep reproductive inefficiency (Alexander 1984), renewed attention is being given to its improvement through breeding. Under Australian conditions of extensive grazing systems, survival of lambs to marking or weaning age can vary considerably and is often less than 80% of lambs born, with losses considerably higher for those born as multiples (Kleemann and Walker 2005).

Lamb survival and net reproduction rate in sheep in general may be affected by correlated changes following selection on other production and quality traits. Little information on these relationships is available; what exists more relates to relationships between growth and some carcass traits with overall ewe reproduction traits (such as the number of lambs born and weaned per ewe joined) and the component traits of fertility and litter size (Safari, Fogarty and Gilmour 2005; Safari *et al.* 2007; Safari *et al.* 2008) rather than with lamb survival expressed as a trait of the lamb. The one exception is a report of positive genetic correlations between ewe body condition scores during pregnancy and ewe rearing ability (Everett-Hincks and Cullen 2009).

The results in this paper give the first estimates of genetic correlations between lamb survival and related traits (birth weight, crown rump length, rectal temperature and time taken to bleat) and a number of meat production and quality traits under study by the Sheep CRC.

MATERIALS AND METHODS

The design of the Information Nucleus (IN) has been described in detail by Fogarty *et al.* (2007). The IN program established base flocks in late 2006 at 8 sites around Australia. Annual artificial insemination matings of the IN base ewes occurred at all 8 sites from 2007 to 2011 (except at the Trangie Research Centre in 2007). The data studied here consisted of complete records of 20,498 observations from eight flocks collected from 2007 to 2010. The records included full pedigree data back to genetic groups, sex of lamb, type of birth (single, twin or multiple), age of dam (two to eight years), sire breed (one of 18 breeds), dam breed (Merino or

crossbred), birth weight, survival at weaning and birth day (day of year). The pedigree included 64,869 identities.

Records collected. The measurements/scores collected at all IN sites that are most relevant to lamb survival and reproduction traits are described by Brien *et al.* (2010) for data collected from 2007 to 2009. Only 4 lamb traits previously reported by Brien *et al.* (2010) to be correlated with lamb survival to weaning (birth weight, time taken to bleat, rectal temperature and crown rump length) have been included in this study. For meat production and quality traits recorded, see Mortimer *et al.* (2010). The number of animals, dams and sires represented in the data set for each trait and the abbreviation, units, mean and standard deviation for each trait, are given in Table 1.

Statistical Analysis. Bivariate analyses were conducted with ASREML (Gilmour *et al.* 2009) on the lamb survival and meat production and quality data from the IN collected from 2007 to 2010. Lamb survival, although a binary trait, was assumed to be distributed normally for these analyses and has been treated as a trait of the lamb. In general, the bivariate analytical models fitted to the data were those used in the analyses described by Brien *et al.* (2010) and Mortimer *et al.* (2010), except that a maternal variance term could not be included.

Table 1. Summary of the data.

Trait	Abbreviation	Animals	Dams	Sires	Mean	SD
Lamb survival to weaning	LSW	20498	8596	377	0.79	0.41
Birth weight (kg)	BWT	20084	8589	377	4.7	1.1
Time taken to bleat (s)	BLT	12931	6561	298	9.0	17.7
Rectal temperature (°C)	RT	14528	6981	299	39.1	1.1
Crown-rump length (cm)	CRL	15646	7174	300	45.7	5.0
Pre-slaughter weight (kg)	PSWT	8734	5276	364	50.5	6.6
Shear force, aged 5 days (N)	SHEARF5	5572	3713	274	26.9	9.7
Intramuscular fat (%)	IMF	5735	3815	279	4.2	1.0
Tissue depth GR site (mm)	HGRFAT	8681	5286	364	13.2	5.4
Carcass weight (kg)	HCWT	8694	5256	363	23.1	3.8
Dressing percentage (%)	DP	8608	5217	363	45.6	3.7
Carcass fat depth 5^{th} rib (mm)	CFAT5	7585	4934	363	7.1	3.5
Eye muscle depth (mm)	CEMD	7657	4979	363	30.0	4.0
Eye muscle area (cm^2)	CEMA	7654	4979	363	14.7	2.5
Lean meat yield (%)	LMY	6147	4049	362	58.0	3.1

RESULTS

Phenotypic correlations. Phenotypic correlation estimates are shown in Table 2. Phenotypic correlations with lamb survival to weaning (LSW) are not reported as lambs must survive to slaughter age to be measured for meat traits.

BWT. All correlations were either in the low (-0.2 to -0.4 or +0.2 to +0.4) or the negligible range (-0.2 to +0.2). Of all the correlations, that with pre-slaughter weight was the highest, at 0.33. The next highest were those with carcass weight (0.26) and fat at the GR site (-0.26). Remaining correlations were below 0.15. The non-zero and positive correlations with pre-slaughter weight and carcass weight were expected, given previous estimates of similar scale for correlations between weights at birth, weaning and hogget age (Safari *et al.* 2007).

CRL, RT and BLT. All estimates were in the negligible range, largely 0.07 or closer to zero. The exceptions were correlations between CRL and HGRFAT (-0.16) and HCWT (0.17).

Table 2. Estimated phenotypic (r_P) and genetic correlations (r_g) between lamb survival to weaning, key survival indicator traits and meat traits. SE in parentheses.

Trait	LSW	BWT		CRL		RT		BLT	
	r_g	r_P	r_g	r_P	r_g	r_P	r_g	r_P	r_g
PSWT	0.12	0.33	0.50	0.02	0.41	0.02	-0.15	0.02	0.10
	(0.05)	(0.01)	(0.04)	(0.01)	(0.05)	(0.02)	(0.07)	(0.02)	(0.07)
HCWT	0.21	0.26	0.39	0.17	0.35	0.04	-0.08	0.02	0.05
	(0.07)	(0.01)	(0.04)	(0.02)	(0.05)	(0.02)	(0.07)	(0.02)	(0.07)
DP	0.22	-0.04	-0.04	-0.01	0.02	0.06	0.08	0.01	-0.06
	(0.06)	(0.01)	(0.05)	(0.02)	(0.06)	(0.02)	(0.07)	(0.02)	(0.07)
LMY	-0.33	0.14	0.38	0.07	0.24	-0.05	0.06	-0.02	-0.14
	(0.06)	(0.02)	(0.05)	(0.02)	(0.06)	(0.02)	(0.08)	(0.02)	(0.08)
HGRFAT	0.34	-0.26	-0.43	-0.16	-0.25	0.04	0.13	-0.01	-0.07
	(0.05)	(0.01)	(0.04)	(0.02)	(0.05)	(0.02)	(0.07)	(0.02)	(0.07)
CFAT5	0.00	-0.14	-0.47	-0.04	-0.18	0.02	0.09	0.02	0.13
	(0.08)	(0.01)	(0.06)	(0.02)	(0.07)	(0.02)	(0.09)	(0.02)	(0.09)
CEMD	0.17	-0.04	-0.01	-0.05	-0.09	0.02	-0.04	-0.01	0.07
	(0.07)	(0.02)	(0.06)	(0.02)	(0.07)	(0.02)	(0.08)	(0.02)	(0.08)
CEMA	0.04	0.01	0.04	-0.02	-0.03	0.03	0.02	-0.01	0.02
	(0.06)	(0.02)	(0.05)	(0.02)	(0.06)	(0.02)	(0.08)	(0.02)	(0.08)
SHEARF5	0.27	0.06	0.16	0.06	0.12	0.02	-0.06	0.02	0.13
	(0.07)	(0.02)	(0.06)	(0.02)	(0.07)	(0.02)	(0.09)	(0.02)	(0.09)
IMF	0.09	-0.10	-0.17	-0.04	-0.07	-0.01	-0.03	0.00	0.07
	(0.06)	(0.02)	(0.05)	(0.02)	(0.06)	(0.02)	(0.08)	(0.02)	(0.08)

Genetic correlations. Genetic correlation estimates are shown in Table 2.

LSW. All estimates were either in the low range or are close to zero. There were low genetic correlations with HGRFAT (0.34), LMY (-0.33), SHEARF5 (0.27), DP (0.22) and HCWT (0.21). Correlations with CEMD, PSWT, IMF, CEMA and CFAT5 were negligible or close to zero.

BWT. The correlation estimates in, or close to, the moderate range were those with PSWT (0.50), CFAT5 (-0.47), HGRFAT (-0.43), HCWT (0.39) and LMY (0.38). In the negligible range were correlations with IMF and SHEARF5. Correlations with DP, CEMD and CEMA were near zero.

CRL. The only genetic correlation in the moderate range was that with PSWT (0.43), although that with HCWT (0.35) was not much less. These positive correlations are expected, given the strong genetic correlation between CRL and BWT of 0.72 (Brien and Rutley, unpublished). HGRFAT (-0.25) and LMY (0.24) had low genetic correlations with CRL. The remaining correlations were in the negligible range, although that with CFAT5 (0.18) bordered on the low range.

RT and BLT. Correlation estimates were mostly in the negligible range and below ±0.10. The exceptions were RT with PSWT and HGRFAT (-0.15 and 0.13, respectively) and BLT with PSWT, SHEARF5, CFAT5 and LMY (0.10, 0.13, 0.13 and -0.14 respectively).

DISCUSSION

Tissue depth at the GR site (0.34, positive) and LMY (0.33, negative and unfavourable) had the strongest estimated genetic correlations with LSW of all traits analysed. The genetic correlation of CEMD with LSW, at 0.17, although significantly greater than zero, is overshadowed by the correlation with fat at the GR site and the unfavourable correlation with LMY. Thus, any selection program that increases LMY, reduces fat (particularly at the GR site) and increases meat

tenderness through selection for lower shear force values will need to take account of the possibility of a genetic reduction in lamb survival. Notwithstanding, the estimated zero or near zero genetic correlations with other fat measurements, such as carcase fat depth at the 5th rib (r_g = 0.00) and intramuscular fat (r_g =0.09) do not indicate sizeable unfavourable consequences for lamb survival if those fat depots are genetically decreased. Genetic increases in slaughter weights, carcase weights, dressing percentages and eye muscle depth, traits likely to be part of breeding objectives for dual purpose and specialised sheep meat production, should all be associated with small genetic increases in LSW.

In earlier work that did not examine lamb survival, Safari *et al.* (2008) concluded that there was no antagonism between reproduction traits and carcass and meat quality indicator traits, with potential to concurrently improve reproduction, carcass and meat quality traits in Merino sheep. Whilst in general agreement, our findings suggest that if sustained selection is practiced for increased LMY and reductions in carcase fatness and shear force, lamb survival may eventually be compromised unless some selection emphasis is dedicated to the trait via appropriately weighted index selection. Conversely, if increasing carcass fatness is used as a selection criterion to enhance reproduction rate and mothering ability (e.g. for dual purpose Merino production systems) a small genetic improvement in lamb survival may be one of the benefits.

ACKNOWLEDGEMENTS

The IN and associated research programs are supported by the Australian Government's Cooperative Research Centres Program, Meat & 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" pp. 199–209, editors D.R. Lindsay and D.T. Pearce, Australian Academy of Sciences: Canberra.

Brien F.D., Hebart M.L., Smith D.H., Hocking-Edwards J.E., Greeff J.C., Hart K.W., Refshauge G., Bird-Gardiner T.L., Gaunt G., Behrendt R., Robertson M.W., Hinch G.N., Geenty K.G.and van der Werf J.H.J. (2010) *Anim. Prod. Sci.* **50:** 1017.

Everett-Hincks J.M. and Cullen N.G. (2009) J. Anim. Sci. **87:** 2753.

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. Gene.* **17:** 29.

Gilmour A.R., Gogel B.J., Cullis B.R. and Thompson R. (2009). 'ASReml User Guide. Release 3' (VSN International Ltd: Hemel Hempstead, UK) *Available at http://www.vsni.co.uk.*

Kleemann D.O. and Walker S.K. (2005) *Theriogenology* **63:** 2075.

Mortimer S.I., van der Werf J.H.J., Jacob R.H., Pethick D.W., Pearce K., Warner R.D., Geesink G.H., Hocking-Edwards J.E., Gardner G.E., Ponnampalum E.N., Kitessa S.M., Ball, A.J. and Hopkins D.L. (2010) *Anim. Prod. Sci.* **50:** 1135.

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. (2007) *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.

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:241-244

GENETIC PARAMETERS FOR BODY WEIGHT, CARCASS AND WOOL TRAITS IN DOHNE MERINO

L. Li[1], D.J. Brown[1] and J.S. Gill[2]

[1]Animal Genetic and Breeding Unit[*], University of New England, Armidale, NSW 2351
[2]Meat and Livestock Australia, University of New England, PO Box U254, Armidale, NSW 2351

SUMMARY

The Dohne Merino was introduced to Australia as a dual-purpose (wool and meat) breed at the end of 1990s with very limited genetic parameters available. Up to 373,639 records per trait were used to estimate genetic parameters for eight traits in this study. Heritability estimates (± s.e.) were 0.34 ± 0.02 for weaning weight (Wwt), 0.27 ± 0.01 for yearling weight (Ywt), 0.13 ± 0.01 for yearling fat depth (Yfat), 0.19 ± 0.05 for yearling eye muscle depth (Yemd), 0.37 ± 0.01 for yearling greasy fleece weight (Ygfw), 0.27 ± 0.01 for yearling clean fleece weight (Ycfw), 0.46 ± 0.01 for yearling fibre diameter (Yfd) and 0.28 ± 0.01 for yearling fibre diameter coefficient of variation (Ydcv), respectively. Significant maternal and maternal environmental effects (± s.e.) were found, being highest for Wwt (0.12 ± 0.01 and 0.05 ± 0.01, respectively) and of smaller magnitude for Ywt, Ygfw and Ycfw (ranging from 0.02 to 0.04). Negative correlations between direct and maternal genetic effects was found for Wwt, Ywt, Ygfw and Ycfw, ranging from -0.41 to -0.75. The genetic and phenotypic correlations between Ygfw and Ycfw were high (0.79 ± 0.01 and 0.89 ± 0.01, respectively) and moderate positive genetic correlations were found between Wwt and Ywt, Wwt and Ygfw, Yfat and Yemd, ranging from 0.26 to 0.54. These values were within the range of estimates found in the literature for Merino sheep.

INTRODUCTION

With the increasing interested in both wool and meat production in the Australian sheep industry, more dual-purpose (wool and meat) sheep breeds have been introduced to Australia (Brown and Fozi 2005). The Dohne Merino, originating from a cross between German Mutton Merino rams and South African Merino ewes in the 1930s (Cloete *et al.* 2001), is such a breed, which was introduced to Australia at the end of 1990s (Casey 2002). The Dohne Merino has been proved an adaptable dual-purpose breed, with easy-care and an ability to thrive under diverse environmental conditions (van Wyk *et al.* 2008). Many records are now available in the Sheep Genetics (SG) database (Brown *et al.* 2007). However, very few genetic parameters have been published for the Dohne Merino. Accurate estimates of variances and covariances are essential for the multiple trait genetic evaluation system used by SG to predict breeding values and further index development. The objective of this study was to estimate genetic parameters for 2 body weight, 2 carcass and 4 wool traits recorded in the Dohne Merino.

MATERIALS AND METHODS

Performance records were extracted from SG database. The traits analysed were weaning weight (Wwt), yearling weight (Ywt), yearling fat depth (Yfat), yearling eye muscle depth (Yemd), yearling greasy fleece weight (Ygfw), yearling clean fleece weight (Ycfw), yearling fibre diameter (Yfd) and yearling fibre diameter coefficient of variation (Ydcv). The minimum and maximum numbers of records were 111,304 for Yemd and 373,639 for Ywt which contained 154 and 130 Australian and South African flocks, respectively. The pedigree was built using all available ancestors in the SG database. A summary of the data for each trait is shown in Table 1.

[1]AGBU is a joint venture of NSW Department of Primary Industries and the University of New England

Variance and covariance components were estimated using an animal model in ASREML (Gilmour *et al.* 2009). Contemporary groups described animal's breed, flock, year, sex and management group and were fitted as fixed effects in the model for each trait. Additional fixed effects included birth type (1 to 4), age of animal at recording (covariate) as well as age of dam and body weight (fitted as quadratic polynomial) depended on the specific trait (Table 2). Rearing type (1 to 4) was only significant for Wwt, however the solutions were not biologically sensible, so it was not included in the model. Random effects including additive genetic effects of individual animal, maternal genetic effects and maternal environmental effects were evaluated by log likelihood ratio tests in univariate analyses. The maternal genetic and maternal environmental effects did not significantly improve the fit of the models for Yfat, Yemd, Yfd and Ydcv and were therefore not included in the models for these traits. A complete set of bivariate analyses was then performed for each trait combination.

Table 1. Summary statistics of the phenotypic data for weaning weight (Wwt), yearling weight (Ywt), yearling fat depth (Yfat), yearling eye muscle depth (Yemd), yearling greasy fleece weight (Ygfw), yearling clean fleece weight (Ycfw), yearling fibre diameter (Yfd) and yearling fibre diameter coefficient of variation (Ydcv)

Traits	Animals with data	Total Pedigree	Sires	Dams	No. CG	Mean	SD	Min	Max
Wwt (kg)	149,001	154,914	2,100	38,044	4,590	25.0	3.79	7.4	46.6
Ywt (kg)	373,639	421,675	6,019	121,133	6,636	45.0	4.70	19.3	82.7
Yfat (mm)	111,472	122,417	1,931	32,860	2,731	2.5	0.58	0.5	7.0
Yemd (mm)	111,304	122,156	1,931	32,736	2,733	27.5	4.86	10.0	48.0
Ygfw (kg)	123,838	131,093	1,971	33,696	3,395	3.5	0.55	0.8	9.0
Ycfw (kg)	279,441	330,702	5,377	101,268	4,383	2.5	0.38	0.1	7.0
Yfd (micron)	370,278	418,685	6,023	120,378	6,565	17.5	1.07	11.1	27.4
Ydcv (%)	123,772	130,690	1,971	33,502	3,378	17.8	2.55	10.2	32.8

RESULTS AND DISCUSSION

Solutions and levels of significance of the fixed effects for each trait are presented in Table 2. Relative to single born animals, twin born lambs were 2.81 and 1.84 kg lighter for Wwt and Ywt and produced 0.19 and 0.02 kg lighter greasy and clean fleece, with 0.14 micron and 0.09% higher fibre diameter and variation in fibre diameter respectively. Triple and quadruple lambs had similar trends as twin lambs with solutions of slightly higher magnitude. Old ewes and animals had significantly increased weaning and yearling weights and yearling greasy fleece weight, with the exception of the non-significant effect of animal age on yearling clean fleece weight. Yearling fat and eye muscle depth increased with heavier body weight. These results are similar to those observed in Australian Merino sheep (Huisman *et al*. 2008). No significant effect of rearing type on Ywt was found in this study, unlike the result found in Australian Merino sheep by Huisman *et al*. (2008), where animals reared as a single were heavier as yearlings than other rearing types.

The phenotypic variances were found similar or slightly lower for 4 wool traits, lower for 2 body weights and higher for 2 carcass traits compared to the estimates of 13.50 (Wwt), 21.80 (Ywt), 0.29 (Yfat) and 4.37 (Yemd) reported by Huisman *et al*. (2008) in Australian Merino sheep. Moderate direct heritabilities were estimated for Wwt (0.34), Ywt (0.27), Ygfw (0.37), Ycfw (0.27) and Ydcv (0.28) (Table 3). These were comparable to estimates of 0.30 (Wwt), 0.30 (Ywt) and 0.22 (Ycfw) reported by Olivier and Cloete (2011) in the South African Dohne Merino. Direct

heritabilities of 0.13 and 0.19 were estimated for Yfat and Yemd, respectively. The highest direct heritability estimate (0.46) was found for Yfd, which was very similar to the estimates reported by Cloete *et al.* (2001) and Olivier and Cloete (2011) in the South African Dohne Merino. These values were within the range found in the Merino sheep (Safari and Fogarty 2003).

Maternal genetic and maternal environmental effects were significant for Wwt, Ywt, Ygfw and Ycfw, with small estimates, which ranged from 0.02 to 0.03, except for Wwt (0.12 and 0.05). The estimates for Wwt compared well with the weighted mean of 0.10 and 0.07 presented by Safari *et al.* (2005). However higher estimates have been reported for Merino sheep for all traits (Huisman *et al*. 2008) and in Dohne Merino for Wwt, Ywt and clean fleece weight (Cloete *et al.* 2001).

Table 2. Solutions of fixed effects including birth type (BT), dam age, animal age (Age), body weight (Wt) for each trait with standard errors in subscript (excluding CG)

Traits	BT2 [A]	BT3 [A]	BT4[A]	Dam Age	Dam Age^2	Age	Wt	Wt^2
Wwt (kg)	**-2.81**$_{0.03}$	**-3.76**$_{0.09}$	**-3.61**$_{0.36}$	**0.62**$_{0.03}$	**-0.06**$_{0.004}$	**0.13**$_{0.002}$		
Ywt (kg)	**-1.84**$_{0.02}$	**-2.72**$_{0.05}$	**-2.23**$_{0.17}$	**0.96**$_{0.02}$	**-0.09**$_{0.003}$	**0.09**$_{0.001}$		
Yfat (mm)							**0.087**$_{0.001}$	**-0.0004**$_{0.00001}$
Yemd (mm)							**0.456**$_{0.005}$	**-0.0017**$_{0.00005}$
Ygfw (kg)	**-0.19**$_{0.01}$	**-0.27**$_{0.01}$	**-0.29**$_{0.05}$	**0.07**$_{0.01}$	**-0.01**$_{0.001}$	**0.01**$_{0.0002}$		
Ycfw (kg)	**-0.02**$_{0.00}$	**-0.10**$_{0.01}$	**-0.08**$_{0.01}$	**0.08**$_{0.01}$	**-0.01**$_{0.001}$	0.0004$_{0.0001}$		
Yfd (micron)	**0.14**$_{0.01}$	**0.23**$_{0.01}$	**0.21**$_{0.04}$	-0.01$_{0.01}$	**0.00**$_{0.001}$			
Ydcv (%)	**0.09**$_{0.01}$	**0.18**$_{0.05}$	**0.17**$_{0.20}$	**0.05**$_{0.02}$	-0.003$_{0.002}$			

[A]The solutions for birth type are relative to a single born lambs. Estimates in bold are significant (P<0.05).

Table 3. Estimates of phenotypic variance (σ^2_p), direct (h^2) and maternal (m^2) heritability, maternal environmental effect (pe^2) as a proportion of phenotypic variance, correlation between direct and maternal genetic effects (r_{DM}) as well as genetic (above diagonal) and phenotypic (below diagonal) correlations with standard errors in subscript

	Wwt	Ywt	Yfat	Yemd	Ygfw	Ycfw	Yfd	Ydcv
σ^2_p	13.50$_{0.08}$	21.80$_{0.06}$	0.29$_{0.00}$	4.37$_{0.02}$	0.32$_{0.00}$	0.15$_{0.00}$	1.27$_{0.00}$	4.52$_{0.02}$
h^2	0.34$_{0.02}$	0.27$_{0.01}$	0.13$_{0.01}$	0.19$_{0.01}$	0.37$_{0.01}$	0.27$_{0.01}$	0.46$_{0.00}$	0.28$_{0.01}$
m^2	0.12$_{0.01}$	0.03$_{0.00}$	-	-	0.03$_{0.00}$	0.02$_{0.00}$	-	-
pe^2	0.05$_{0.00}$	0.03$_{0.00}$	-	-	0.04$_{0.00}$	0.03$_{0.00}$	-	-
r_{DM}	-0.69$_{0.02}$	-0.41$_{0.03}$	-	-	-0.75$_{0.04}$	-0.65$_{0.03}$	-	-
Wwt		0.54$_{0.02}$	-0.13$_{0.04}$	-0.11$_{0.03}$	0.33$_{0.03}$	0.10$_{0.03}$	0.04$_{0.02}$	-0.16$_{0.02}$
Ywt	0.60$_{0.00}$		0.03$_{0.04}$	-0.02$_{0.03}$	0.04$_{0.02}$	0.01$_{0.02}$	0.16$_{0.01}$	-0.19$_{0.02}$
Yfat	-0.05$_{0.00}$	0.09$_{0.00}$		0.44$_{0.03}$	-0.10$_{0.04}$	-0.05$_{0.06}$	0.12$_{0.02}$	-0.09$_{0.03}$
Yemd	0.00$_{0.00}$	0.02$_{0.01}$	0.26$_{0.00}$		-0.10$_{0.03}$	-0.16$_{0.05}$	0.05$_{0.02}$	-0.10$_{0.03}$
Ygfw	0.31$_{0.01}$	0.37$_{0.00}$	-0.02$_{0.00}$	0.00$_{0.00}$		0.79$_{0.01}$	0.15$_{0.02}$	0.20$_{0.02}$
Ycfw	0.30$_{0.01}$	0.33$_{0.00}$	0.00$_{0.01}$	-0.01$_{0.01}$	0.89$_{0.001}$		0.16$_{0.01}$	0.20$_{0.03}$
Yfd	0.08$_{0.00}$	0.16$_{0.00}$	0.08$_{0.00}$	0.05$_{0.00}$	0.19$_{0.00}$	0.19$_{0.00}$		-0.10$_{0.02}$
Ydcv	-0.10$_{0.00}$	-0.13$_{0.00}$	-0.05$_{0.00}$	-0.06$_{0.00}$	0.02$_{0.00}$	-0.01$_{0.01}$	-0.10$_{0.00}$	

The genetic correlation between direct and maternal genetic effects was highly negative (ranging from -0.41 to -0.75) for Wwt, Ywt, Ygfw and Ycfw in this study. It was different to the weighted mean of 0.34 for dual purpose sheep reported by Safari *et al.* (2005), but similar to

estimates published in the studies of Huisman *et al.* (2008) and Cloete *et al.* (2001). These high estimates were considered to be inflated by the data structure as described by Clement *et al*.(2001). It is noteworthy that lower direct heritabilities were estimated for Wwt (0.19), Ywt (0.22), Ygfw (0.24) and Ycfw (0.20) when covariance between direct and maternal genetic effects was fixed at zero in the models.

Highly positive genetic and phenotypic correlations between Ygfw and Ycfw (0.79 and 0.89, respectively) and moderate positive genetic correlations between Wwt and Ywt, Wwt and Ygfw, Yfat and Yemd, ranging from 0.26 to 0.54, were found in this study. Very few estimates have been reported for the correlations between these traits in the Dohne Merino. van Wyk *et al.* (2008) obtained similar genetic and phenotypic correlations between mean fibre diameter and yearling weight (0.13 and 0.13) and clean fleece weight (0.16 and 0.18) in this breed. Higher genetic correlations between Wwt and Ywt (0.83), Ycfw (0.32) along with Yfd (0.12) were reported in the same breed by Olivier and Cloete (2011). Compared to the estimates in Australian Merino sheep, most of estimates are in agreement with those found by Huisman and Brown (2008 and 2009) with some exceptions including much higher genetic and phenotypic correlations between Ywt with Yfat (0.29 and 0.47), Yemd (0.85 and 0.83) and Ygfw (0.32 and 0.32) along with higher genetic correlation between Ygfw and Yfd (0.42), Ycfw and Yfd (0.42) in Australian Merino sheep.

CONCLUSIONS

Accurate genetic parameters were estimated for 2 body weight (Wwt and Ywt), 2 carcass (Yfat and Yemd) and 4 wool (Ygfw, Ycfw, Yfd and Ydcv) traits with large amounts of phenotypes available from SG database for the Dohne Merino. Most of these estimates were similar to other Merino breeds. These genetic parameters will be used to review those being used in the SG evaluation system and further index development for the Dohne Merino.

ACKNOWLEDGMENTS

This research was funded by Meat and Livestock Australia (MLA) and Australian Wool Innovation (AWI). The authors would like to thank the breeders from the Australian Dohne Breeders Association who collected the pedigree and performance data used in this study.

REFERENCES

Brown D. J. and Fozi M. A. (2005) *Proc. Assoc. Advmt. Anim. Breed. Genet.* **16:** 310.

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.

Casey A. E. (2002) *Wool Tech. Sheep Breed.* **50:** 390.

Clement V., Bibe B., Verrier E., Elsen J.-M., Manfredi E., Bouix J. and Hanocq E. (2001) *Gent. Sel. Evol.* **33:** 369.

Cloete S. W. P., Scholtz Anna J. and Aucamp B.B. (1998) *S. Afr. J. Anim. Sci.* **28:** 185.

Cloete S. W. P., Schoeman S. J., Coetzee J. and Morris J. D. (2001) *Aust. J. Exp. Agric.* **41:** 145.

Gilmour A.R., Gogel B.J., Cullis B.R. and Thompson R. (2009) ASReml user guide. Release 3.0. VSN International Ltd., Hemel Hempstead, UK. 398 pp.

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. (2008) *Aust. J. Exp. Agric.* **48:** 1186.

Huisman A.E. and Brown D.J. (2009) *Anim. Prod. Sci.* **49: 2**89.

Olivier J.J. and Cloete S.W.P. (2011) *Proc. Assoc. Advmt. Anim. Breed. Genet.* **19:** 167.

Safari A. and Fogarty N. M. (2003) *Technical Bulletin* **49,** NSW Agriculture, Orange, Australia.

Safari E., Fogarty N. M., and Gilmour A.R. (2005) *Liv. Prod. Sci.* **92:** 271.

van Wyk J.B., Swanepoel J.W., Cloete S.W.P., Olivier J.J. and Delport G.J. (2008) *S. Afr. J. Anim. Sci.* **38:** 31.

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:245-248

ESTIMATES OF HERITABILITY FOR COLOUR CIE a* MEASUREMENTS AT FOUR TIME POINTS FOR CHILL AGED LAMB

M.J. Bixley[1], J.C. McEwan[1], W.E. Bain[1], E. Young[1], N.J. McLean[2] and P.L. Johnson[1]

[1]AgResearch, Invermay Agricultural Centre, PB 50034, Mosgiel, New Zealand
[2]AbacusBio Limited, PO Box 5585, Dunedin 9058, New Zealand

SUMMARY

This paper investigates genetic control of redness for 8 week chill aged lamb. Heritability of CIE a* values (Commission Internationale de l' Eclairage, 1976, a measure of redness) has been estimated from 18,913 carcasses of crossbred lambs born 2003-2010. Colour was recorded at 24, 48, 96 and 168 hours post display wrapping. Heritability estimates for the combined dataset were 0.55±0.03, 0.57±0.03, 0.58±0.03 and 0.29±0.03 respectively for the 4 time points, indicating that the colour of chill aged lamb loins is under moderate genetic control.

INTRODUCTION

More than 90% of New Zealand lamb is exported as chilled product. Predicting and controlling shelf life of this product is of crucial importance. Colour stability of lamb meat entering the fresh retail market is a primary factor in determining retail shelf life. The colour of the meat when on retail display is a major selection criterion for purchasers (Killinger *et al.* 2004; Savell *et al.* 1989) with the majority of consumers preferring bright red coloured meat which they associate with freshness (Killinger *et al.* 2004).

Data from a number of industry progeny tests, spanning multiple years, were available to investigate genetic variation in colour stability of New Zealand lamb**.** This paper investigates the role of genetics at 4 time points of long term chilled pasture-fed lamb. In addition, we investigate the role that genetics plays expanding on work completed by Campbell *et al.* (2004, unpublished) and Johnson *et al.* (2007, unpublished).

MATERIALS AND METHODS

Eight week colour stability data was available from the *M. longissimus dorsi* (loin) on 18,913, 2003-2010 born lambs from 1075 sires. Lambs were sourced from a number of industry progeny tests, with the majority of lambs sourced from the Beef + Lamb New Zealand Central Progeny Test (McLean *et al.* 2006) and the Rissington Breedline Primera Progeny Test (Johnson *et al.* 2007, McLean *et al.* 2009). Mixtures of terminal and dual-purpose sires were used to generate progeny. Lambs were slaughtered in commercial plants with the carcasses electrically stimulated. One day post slaughter the carcasses were processed into primal cuts. The boneless loins were vacuum packed and stored at -1°C for 8 weeks.

At 8 weeks, pH was measured on the loin. Three 2cm thick slices (avoiding the ends) of the loin were placed on small plastic trays and wrapped using semi permeable cling film to replicate supermarket conditions (no gas flushing). These were stored at 4°C. Meat colour was measured using a Minolta Chromameter (Konica Minolta Sensing, Inc., Osaka Japan) with the CIE L* a* b* system (which measures relative lightness, redness and yellowness respectively) at 24, 48, 96 and 168 hours post display wrapping, using. One measurement was taken from each of the three slices with the average of these values recorded in the Sheep Improvement Limited (SIL) database. The primary measure of interest was CIE a* which is the objective measure of redness/brownness most correlated with consumers subjective measures of colour acceptability (Moore & Young, 1991). From the 2007 born animals onwards, a change in facility (new processing room and chillers for storage and display) in which the measurements were made was the only alteration in protocol.

Variance components were estimated using restricted maximum likelihood (REML) procedures fitting an animal model in ASREML3 (Gilmour *et al.* 2009). Univariate analyses were used to estimate the heritability for each trait. Genetic variation of CIE a* measured 24, 48, 96 and 168 hours after processing for lamb loins that were chill aged was estimated. Birth rearing rank, age of dam and contemporary group were fitted as fixed effects, live weight at 6 months of age and pH as covariates, and animal as a random effect. An interaction between sex, birth flock and contemporary group was also fitted as a fixed effect. For the analysis of data spanning multiple years, birth year was fitted as a fixed effect.

Table 1. Summary of colour CIE a* of chill aged lamb loin records collected on 2003-2010 born lambs

			24 Hour		48 Hour		96 Hour		168 Hour	
Year	Records	Sires	Mean	SD	Mean	SD	Mean	SD	Mean	SD
2003	1155	118	22.76	1.89	20.91	1.82	18.23	1.87	14.69	2.09
2004[a]	2241	117	19.98	1.45	18.04	1.38	15.35	1.37	NA[b]	NA
2005	2216	106	21.98	1.73	20.84	1.71	18.79	1.71	16.82	1.83
2006	2170	135	22.96	1.71	22.05	1.62	20.68	1.56	18.77	1.75
2007	1062	66	19.53	1.91	18.89	1.62	17.36	1.46	15.77	1.38
2008	2629	157	17.91	1.39	16.41	1.38	14.03	1.65	12.52	1.54
2009	3187	164	16.47	1.78	14.4	1.66	12.52	1.92	10.02	1.81
2010	4253	235	16.21	1.97	14.56	1.73	12.08	1.69	10.24	2.08
All Years	18913	1075	18.97	3.14	17.45	3.32	15.26	3.49	15.66	3.65

[a]2004 data affected post slaughter by inadequate refrigeration, [b]NA=Not Available

RESULTS AND DISCUSSION

Narrow sense heritability estimated for each birth year from 2003-2010 and for the combined data are shown in Figure 1. The greatest variation in estimated heritability was at 96 hours. At this time point the greatest estimates of heritability 0.55±0.06 and 0.58±0.03 were observed in the 2009 and 'all years' data respectively. Two thirds of heritability estimates were greater than 0.2 with the exception of years 2004 and 2007. Heritability was significant ($p<0.05$) for all values except for 2007 at 48, 96 and 168 hours. This is likely to be a result of reduced sampling with only 66 sires in 2007 (Table 1)

Current literature contains very few estimates of heritability for meat quality traits in sheep, as noted in Hopkins *et al.* (2011) in their update of the extensive review of Safari *et al.* (2003). Heritability of CIE a* in Merino has been reported as 0.10±0.03 (Greef *et al.* 2008) and has also been reported to not be significantly different from zero (Fogarty *et al.* 2003, Cloete *et al.* 2008). In Scottish Black Face it is reported as 0.45±0.19 (Karamichou *et al.* 2006). A heritability of 0.19 was reported for a Central Progeny Test (CPT) subset of the data used in this analysis (Payne *et al.* 2009). A Rissington Breedline Progeny Test subset reported an estimated heritability for CIE a* at 168 hours among Suffolk, White Suffolk and Poll Dorset to be 0.26±0.04 (McLean *et al.*, 2009). Mortimer *et al.* (2011) present results for a* from 3328 animals of various breeds in the Information Nucleus program of the CRC for Sheep Industry Innovation. One, 2 and 3 day values are reported as 0.18±0.04, 0.23±0.04 and 0.20±0.04 respectively. Other than merino, the reported values are in the same range as those reported in this paper.

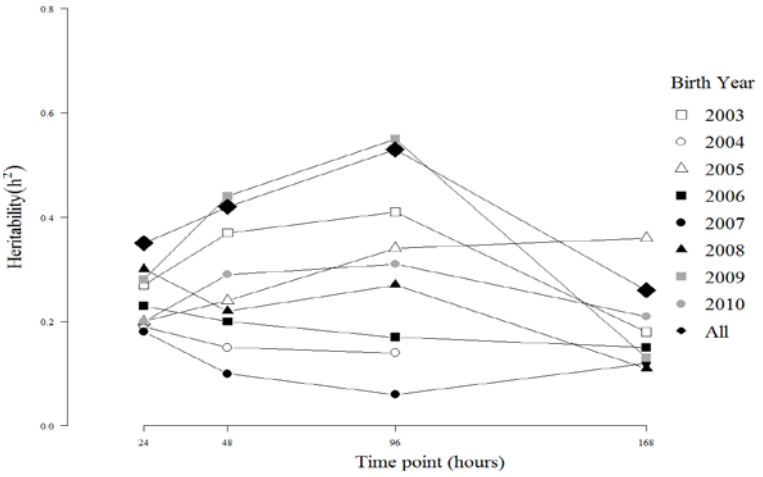


Figure 1. Heritability estimates of colour CIE a* of chill aged lamb loin at 24, 48, 96 and 168 hours after further processing by Birth Year and for the combined data.

Data in the initial year of collection was for a greater number of time points, every 24 hours until 90% of samples fell below CIE a* = 16 (Figure 2. dashed line). Subsequent data collection was terminated at 168 hours where 75% of the samples had fallen below the value of 16. In 2005, 2006 and 2007 less than 75% of the samples fell below this threshold.

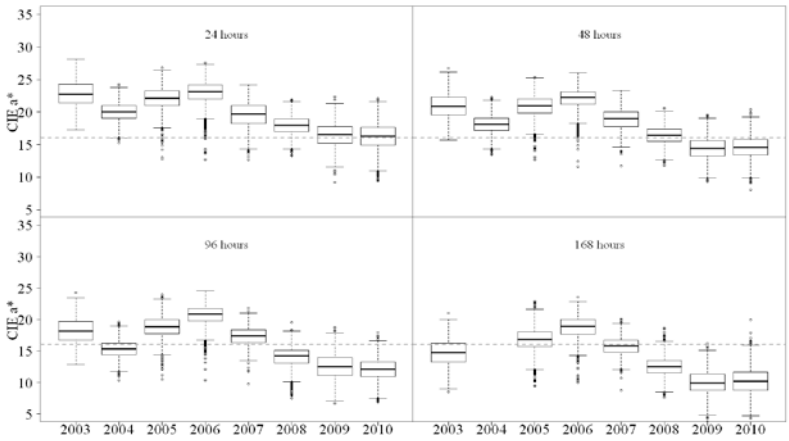


Figure 2. Variation of CIE a* values at 4 time points by each year of data collection. Dashed horizontal line is CIE a* = 16.

From 2008 (birth year 2007) new facilities were used for data collection and there are marked differences in the datasets from this year on, see Figure 2. Almost half of the samples are now below the original lower limit of 16 by 24 hours, making analysis of time to 16 somewhat difficult or redundant. In the 2004 born cohort there is also a notable decrease in initial colour values. This was attributed to the storage temperature not being maintained at -1°C early post slaughter. This has an effect in the All Years mean values in Table 1. Excluding 2004 would see an increase in

values. However, even within a facility there is significant variation between years and between flocks within year (significant in the linear model analysis). This would suggest that un-known pre-slaughter factors and storage facilities are also important in determining the colour stability of chill aged lamb loins.

Ultimately the objective of the colour analysis was to determine whether or not redness of chill aged lamb loins is under genetic control, and as such can be used as a selection criterion to enhance the shelf life of the product. The heritability estimates do suggest that the redness of chill aged lamb loins is under moderate genetic control. There is interest by some overseas supermarkets in genetic selection for colour stability of redness, however, financial rewards are not currently offered to commercial producers. Despite the lack of financial rewards some breeders are using breeding values generated from this data set to exclude outliers with poor colour stability breeding values from their breeding programme in the aim of producing premium product for future market differentiation. Further analysis will investigate the genetic correlations between colour measurements and other meat quality and production traits.

ACKNOWLEDGEMENTS

The authors would like to recognise the significant and on-going contribution to data collection made by the Central Progeny Test and Rissington Breedline (now Focus Genetics) Primera Progeny Test, the staff involved in the management of the programs and those that collect and manage the traits measured. The field team at AgResearch, Invermay, and the staff on the killing chains at various processing plants, who had their daily routine disrupted, are also thanked. This work was funded by Ovita and Beef and Lamb New Zealand.

REFERENCES

Cloete S.W.P., Cloete J.J.E. and Hoffman L.C. (2008) *Proc. 54th Int. Cong. Meat. Sci. Technol.*

Gilmour A.R., Gogel B.J, Cullis B.R. and Thompson R. (2009) 'ASReml User Guide Release 3.0' VSN Int. Ltd, Hemel Hampstead, UK.

Fogarty N.M., Safari E., Taylor P.J. and Murray W. (2003) *Aust. J. Agric. Res*. **54**:715.

Greef 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

Hopkins D.L., Fogarty N.M. and Mortimer S.I. (2011) *Small Rum. Res.* **101**:160.

Killinger K.M., Calkins C.R., Umberger W.J., Feuz D.M. and Eskridge K.M. (2004) *J. Anim. Sci.* **82**:3288.

Karamichou E., Richardson R.I., Nute G.R., McLean K.A. and Bishop S.C. (2006) *Anim. Sci.* **82**:151.

McLean N.J., Jopson N.B., Campbell A.W., Knowler K., Behrent M., Cruickshank G., Logan C.M., Muir P.D. and McEwan J.C. (2006) *Proc. N.Z. Soc. Anim. Prod*. **66:**368.

McLean N.J., Johnson P.L., Bain W.E., Greer G.J. and Dodds K.G. (2009) *Proc. N.Z. Soc. Anim. Prod.* **69**:220.

Mortimer S.I., Jacob R.H., Warner R.D., Pearce K.L., Pethick D.W., van der Werf J.H.J., Ball A.J. and Hopkins D.L. (2011) *Proc. Assoc. Advmt. Anim. Breed. Genet.* **19**:107.

Payne G.M., Campbell A.W., Jopson N.B., McEwan J.C., Logan C.M. and Muir P.D. (2009) *Proc. N.Z. Soc. Anim. Prod*. **69**:210.

R Development Core Team (2011) 'R: A language and environment for statistical computing.' R Foundation for Statistical Computing, Vienna, Austria. ISBN3-900051-07-0, http://www.R-project.org/.

Safari A. and Fogarty N.M. (2003) *Technical Bulletin 49*, NSW Agriculture, Orange, Australia.

Savell J.W., Cross H.R., Francis J.J., Wise J.W., Hale D.S., Wilkes D.L. and Smith G.C. (1989) *J. Anim. Sci.* **12:**251.

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:249-252

GENETIC ASSOCIATIONS OF EARLY GROWTH AND ULTRASOUND SCANNED TRAITS IN SEVERAL BEEF BREEDS

M. G. Jeyaruban and D. J. Johnston

Animal Genetics and Breeding Unit[1], University of New England, Armidale, NSW 2351

SUMMARY

Genetic association of early growth traits with ultrasound scanned traits in heifers and bulls of Angus, Charolais, Hereford, Simmental and Santa Gertrudis were examined. Early growth traits such as birth weight (BW), 200 day weight (200D) and 400 day weight (400D) along with ultrasound scanned rump (P8), intramuscular fat (IMF) and eye muscle area (EMA) in heifers and bulls, were considered in this analysis. Estimated genetic correlations between BW and scanned fat traits were negative and ranged from -0.48 to -0.04. The 200D and 400D had positive genetic correlations with the scanned fat of heifers of all breeds, and also for the scanned fat of Angus and Herefords bulls. Breed influence on the genetic correlation of early growth traits with scanned traits was evident for all traits. The magnitude of the estimated genetic correlations of early growth traits with scanned fat traits may not be adequate to cause significant correlated changes in these traits. Therefore, selection objectives, combining early growth and ultrasound scanned traits, is required if it is necessary to change these traits in the five breeds studied.

INTRODUCTION

BREEDPLAN, the national Australian genetic evaluation, has been used for nearly 30 years to improve the genetic potential of Australian seedstock herds (Graser *et al.* 2005). Estimated breeding values are calculated for more than twenty five economically important traits in various beef breeds by using a multi-trait genetic evaluation system (Johnston 2007). Genetic correlations between traits are breed specific. Therefore, knowledge of the genetic relationships between traits are important in the multi-trait genetic evaluation system for the accurate prediction of correlated responses in genetically correlated traits and the identification of early indicator traits for traits expressed later in life or which are difficult to measure. Currently growth traits are easy and cheap to record and are widely recorded in most of the beef breeds in Australia. Therefore, it is very important that their relationships with body composition traits (fat and eye muscle area) are well estimated for all breeds. The objective of this paper was to study the genetic association of early growth and scanned traits in commonly used beef breeds in Australia to understand the biological relationship of lean to fat content in these beef breeds.

MATERIALS AND METHODS

Data used for this study were submitted by breeders to their breed societies for use in BREEDPLAN. The breeds included Angus (ANG), Charolais (CHA), Hereford (HER), Simmental (SIM) and Santa Gertrudis (SAN). Early growth traits considered were birth weight (BW), 200-day weight (200D) and 400-day weight (400D) with age at recording ranging from 80 to 300 days for 200D and 301 to 500 days for 400D. Real-time ultrasound scan measurements included fat depth at the P8 (rump) site for bulls (BP8) and heifers (HP8), intramuscular fat in bulls (BIMF) and heifers (HIMF) and longissimus muscle area in bulls (BEMA) and heifers (HEMA), with age at recording for all traits ranging from 300 to 800 days. For ANG and HER, data from animals born after 2004 were used to reduce computational requirements. Genetic parameters for growth and scanned traits were estimated using a univariate linear animal model as given below,

[1] AGBU is a joint venture of NSW Department of Primary Industries and University of New England

$Y_{ijkl} = cg_i + \beta_1 age_j(season_l) + \beta_2 age_j^2(season_l) + \beta_3 hf_j(season_l) + \beta_4 age_k(sex_k) + a_k + e_{ijkl}$
where Y_{ijkl} is the trait measured in animal k in a fixed contemporary group i (cg_i), age_j is the age of dam j at calving deviated from five years of age nested within season (Autumn and Spring), hf_j (heifer factor) is the additional dam age function fitted to improve the fit for offspring of dams less than 2.5 years old, age_k is the age of animal k nested within sex of animal k, β_1, β_2 and β_3 are the regression coefficients for linear, quadratic and heifer factor effects of dam age, β_4 is the regression coefficient for age of animal, a_k is the random genetic effect of animal k and e_{ijkl} is the random error associated with each observation. For 200D, 400D and scanned traits, the age was deviated from 200, 400 and 500 days, respectively, as currently implemented in BREEDPLAN. Additional random maternal genetic and random permanent maternal environment effects of dam j are fitted to BW, 200D and 400D to account for the maternal influence on these traits. Contemporary groups were as defined by Graser *et al*. (2005). Bivariate animal models were used to estimate genetic correlations of BW, 200D and 400D with scanned traits for each breed separately. Complete pedigree information going back three generations was used. Estimates of (co)variance components in the univariate and bivariate analyses were obtained using WOMBAT (Meyer 2007).

RESULTS AND DISCUSSION

The number of records used for each trait by breed is presented in Table 1. The ANG had the highest number of records for all the traits, while CHA had the least number of records, except for BW and BIMF. The number of animals recorded for growth traits were higher than the number of animals recorded for the scanned traits. Table 2 gives the estimated heritability (h^2) for age adjusted growth and ultrasound scanned traits. The BW was moderately heritable in all breeds with the estimated h^2 ranged from 0.24 to 0.40. Low and low to moderate h^2 were estimated for 200D and 400D, respectively. The estimated h^2 ranged from 0.12 to 0.17 for 200D and 0.19 to 0.26 for 400D. Ultrasound scanned traits on heifers were moderately heritable and were generally higher than the estimates for bulls, except for HIMF of SIM. Breed variations in h^2 of HP8 (0.35 to 0.49), HIMF (0.23 to 0.49), HEMA (0.24 to 0.42), BP8 (0.21 to 0.28), BIMF (0.17 to 0.40) and BEMA (0.22 to 0.34) were observed. The estimated h^2 of the scanned traits of ANG and HER were very similar. Estimated h^2 for growth and ultrasound scanned traits were within the range published in the literature (Koots *et al*. 2005).

Table 1. Number of growth and ultrasound scanned traits records used from Angus (ANG), Charolais (CHA), Hereford (HER), Simmental (SIM) and Santa Gertrudis (SAN)

Breed		BW	200D	400D	HP8	HIMF	HEMA	BP8	BIMF	BEMA
ANG	N[A]	308938	273546	186377	73865	70752	74338	76265	73044	77243
HER	N[A]	155733	177749	114847	30105	27144	30339	39064	32256	39299
SIM	N	136541	121287	77103	6926	2254	6980	11425	2782	11566
CHA	N	48100	58554	33746	2740	1518	2767	4115	3163	9646
SAN	N	3832	111982	60910	11618	3966	11780	19820	6318	20023

[A] Using data from animals born after 2004 to reduce computational requirements

Table 2. Estimated heritabilities (direct) for early growth and ultrasound scanned traits (standard error in parenthesis) using univariate animal model evaluation

Breed	BW	200D	400D	HP8	HIMF	HEMA	BP8	BIMF	BEMA
ANG	0.32	0.12	0.22	0.43	0.28	0.26	0.28	0.17	0.24
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
HER	0.32	0.12	0.19	0.36	0.28	0.24	0.26	0.21	0.22
	(0.02)	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.01)
SIM	0.26	0.14	0.26	0.40	0.23	0.42	0.21	0.29	0.26
	(0.02)	(0.01)	(0.01)	(0.03)	(0.03)	(0.06)	(0.03)	(0.03)	(0.06)
CHA	0.24	0.17	0.26	0.49	0.39	0.35	0.27	0.27	0.34
	(0.03)	(0.01)	(0.02)	(0.06)	(0.09)	(0.06)	(0.03)	(0.06)	(0.03)
SAN	0.40	0.17	0.22	0.35	0.49	0.30	0.27	0.40	0.27
	(0.07)	(0.01)	(0.02)	(0.03)	(0.06)	(0.03)	(0.02)	(0.05)	(0.02)

Table 3. Estimated genetic correlations between early growth and ultrasound scanned traits (standard error in parenthesis)

Breed	Trait	HP8	HIMF	HEMA	BP8	BIMF	BEMA
ANG	BW	-0.26(0.02)	-0.22(0.02)	0.25(0.02)	-0.29(0.03)	-0.16(0.03)	0.13(0.03)
	200D	0.21(0.03)	0.34(0.03)	0.79(0.02)	0.01(0.04)	0.26 (0.04)	0.67(0.03)
	400D	0.21(0.05)	0.19(0.03)	0.60(0.02)	0.08(0.03)	0.19 (0.04)	0.57(0.02)
HER	BW	-0.40(0.03)	-0.26(0.04)	0.22(0.04)	-0.34(0.04)	-0.15(0.05)	0.24(0.04)
	200D	0.05(0.06)	0.16(0.06)	0.74(0.04)	0.08(0.06)	0.05(0.07)	0.72(0.04)
	400D	0.01(0.04)	0.07(0.05)	0.66(0.03)	0.11(0.05)	0.12(0.05)	0.65(0.03)
SIM	BW	-0.19(0.06)	-0.38(0.10)	0.31(0.07)	-0.21(0.07)	-0.02(0.15)	0.36(0.07)
	200D	0.05(0.08)	0.01(0.14)	0.63(0.07)	-0.24(0.08)	0.04(0.18)	0.47(0.07)
	400D	0.20(0.06)	0.01(0.12)	0.60(0.06)	-0.11(0.07)	0.04(0.17)	0.45(0.06)
CHA	BW	-0.24(0.08)	-0.48(0.14)	0.25(0.10)	-0.23(0.10)	-0.28(0.13)	0.13(0.07)
	200D	0.03(0.10)	0.25(0.14)	0.52(0.09)	0.20(0.10)	-0.07(0.12)	0.48(0.06)
	400D	0.07(0.10)	0.31(0.15)	0.45(0.09)	-0.01(0.11)	-0.01(0.13)	0.53(0.06)
SAN	BW	-0.33(0.10)	-0.13(0.09)	0.17(0.11)	-0.17(0.13)	-0.04(0.11)	0.16(0.14)
	200D	0.15(0.05)	0.08(0.08)	0.52(0.04)	-0.01(0.04)	-0.01(0.07)	0.47(0.04)
	400D	0.21(0.05)	0.16(0.09)	0.59(0.04)	0.04(0.05)	0.18(0.07)	0.49(0.04)

Table 3 presents the estimated genetic correlations between age adjusted early growth and ultrasound scanned traits. Ultrasound scanned fat traits (P8 and IMF) on heifers and bulls had low to moderate negative correlation with BW. The genetic correlations of BW with scanned fat traits in heifers ranged from -0.40 (HER) to -0.19 (SIM) for HP8 and -0.48 (CHA) to -0.13 (SAN) for HIMF. The genetic correlations of BW with scanned fat traits in bulls ranged from -0.34 (HER) to -0.17 (SAN) for BP8 and -0.28 (CHA) to -0.02 (SIM) for BIMF. However, BW had low to moderate positive correlations with HEMA and BEMA. The breed influence was evident on the genetic correlations of BW with HEMA (0.17 to 0.31) and BEMA (0.13 to 0.36).

The 200D had low to moderate positive genetic correlations with scanned fat traits of heifers. The genetic correlations of 200D with the scanned fat traits of bulls were also generally positive, with the exception of BP8 in SIM, BIMF in CHA and BP8 and BIMF in SAN. The 400D had low

to moderate positive correlations with scanned fat traits of heifers. The genetic correlation observed between 400D and scanned fat traits in bulls were low and ranged from -0.11 to 0.11 for BP8 and -0.01 to 0.19 for BIMF. The 200D and 400D were moderately to highly correlated with HEMA and BEMA in all breeds and the correlations ranged from 0.52 to 0.79 for HEMA and 0.47 to 0.72 for BEMA.

Low to moderate correlations of BW and scanned fat traits in heifers and bulls indicated that selection for lower BW would result in slightly higher fat depths in all breeds. Selecting for higher 200D and 400D is expected to increase the fat in heifers of all breeds, and also for the scanned fat of Angus and Herefords bulls. However, the magnitude of the estimated correlations of age adjusted early growth traits with age-adjusted scanned fat traits indicated that the expected changes in the fat content of heifers and bulls would not alter the body composition significantly in any of the five breeds studied.

CONCLUSIONS

There was variation for the estimated heritabilities of early growth and ultrasound scanned traits of the five breeds. For all breeds, correlations indicate that genetically high BW is associated with reduced fatness and increased EMA when considered on an age-constant basis. Similarly, genetically heavier animals at 200D and 400D have larger EMA. However, there is some variation in the magnitude of estimates between breeds and genders. In contrast, there appears no consistent genetic relationship between 200D and 400D weight traits with fatness when considered on an age constant basis corrected to 500 days of age. All of these relationships might change if scan traits are corrected based on weight-constant basis. Further research is required to validate this claim. Results demonstrated the importance of combining early growth and scanned traits in selection objectives to alter the body composition to fulfil different market requirements.

ACKNOWLEDGEMENT

The authors would like to thank Meat and Livestock Australia for their financial support through B.BFG.0050 and the Angus, Charolais, Hereford, Simmental and Santa Gertrudis Breeders Associations for providing data for this study. The authors also would like to acknowledge the contribution of Kim Bunter for her comments on the manuscript.

REFERENCES

Graser H-U., Tier B., Johnston D.J. and Barwick S.A. (2005) *Aust. J. Exp. Agric*. **45**:913.
Johnston D.J. (2007) *Proc. Assoc. Advmt. Anim. Breed. Genet.* **17**:8
Koots K.R., Gibson J.P., Smith C. and Wilton J.W. (1994). *Anim. Breed. Abst.* **62**:309
Meyer K (2007) *J. Zhej. Uni. Sci.* **8**:815.

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:253-256

SELECTION OPPORTUNITIES FROM USING ABATTOIR CARCASS DATA

E. Wall, M. Coffey and T. Pritchard

SRUC, Edinburgh, EH25 9RG, Scotland

SUMMARY

Genetic improvement of UK beef carcass traits currently uses predictor traits (weights and ultrasound measurements) taken on live performance recorded pedigree selection candidates. These phenotypes are very low in numbers and have only a moderate-high correlation with the goal traits in crossbred slaughter populations. However, carcass phenotypes taken from abattoir records are available in large quantities and are a key target goal trait for many terminal pedigree beef breeders. The extra information from abattoir data will improve the accuracy of future genetic evaluations on carcass traits. This study uses carcass information from UK abattoirs and information from the national cattle tracing system to create files for the genetic evaluation of carcass traits. A combination of sources merged together might contain sufficient information which could then be used to produce Estimated Breeding Values (EBVs) for carcass traits. The major breeds present in the carcass population (with over 100,000 animals, including crosses) were Limousin, Aberdeen Angus, Holstein Friesian, Charolais, Hereford, Simmental, and Belgian Blue, and these accounted for 92% of the animals (2.4 million records in the merged dataset, 2001-2012). Genetic analyses were performed on a subset of the data for animals with a Charolais sire (17,125 records after editing). Heritability estimates for carcass weight, conformation and fat class were 0.31, 0.24, and 0.14 respectively. The results of this feasibility study indicate that genetic analysis for carcass traits is realistic, particularly for breeds which make up a major part of the carcass population and have sufficient information on the sire. This, in turn, suggests that improving carcass traits through genetic selection is entirely possible, thereby warranting more detailed investigation of their genetic background, particularly their relationship with other traits of importance and within, between and across breeds.

INTRODUCTION

Currently, genetic improvement of beef carcass traits in the UK makes use of predictor traits, weight and ultrasound measurements, taken on live animals (Amer *et al*. 1998). One of the UK beef breeding goals, "Beef Value" (a function of weight and carcass merit, in terminal sire and dual purpose breeds), has been shown to be effective in bringing about genetic change. For example in the period 1999-2003, there were substantial annual increases in Beef Value £0.69/yr (Amer *et al*. 2007). These are mainly recorded on pedigree animals in low numbers and this is where the majority of the genetic change is expressed. However, carcass phenotypes taken from abattoir records are mostly measured outside the pedigree sector and are available in large quantities. The extra information from abattoir data may improve the accuracy of future genetic evaluations on carcass traits. However, abattoir data alone would not be informative enough for genetic evaluations i.e. without pedigree and management information. A combination of sources merged together might contain sufficient information which could then be used to produce Estimated Breeding Values (EBVs) for carcass traits. These sources include abattoirs, Beef and Sheep Company (BASCO), breed societies, milk recording organisations and British Cattle Movement Service (BCMS). The overall aim of the study 'Carcass Trait Evaluations,' was to produce a consolidated dataset of carcass traits and pedigree for beef and dairy cattle by combining all sources of information. Data description was undertaken in this study to reveal the suitability of its use for genetic evaluations in the future.

MATERIALS AND METHOD

Phenotypic data. The sources of data that were combined for these analyses are described in Figure 1. The carcass data was obtained from three UK based meat processing/slaughter companies from 2001-2012 [NOTE: not all companies had data for all years]. These carcass records were joined to the BCMS database based on UK eartag. Initially, 3 million individual carcass abattoir records (from 3 abattoirs) and approximately 48 million BCMS animal records were made available for this project. Using intelligent string matching, 82% of the individual carcass records could be matched to BCMS animal records, using UK eartag identity, resulting in a dataset of ~ 2.4 million abattoir records for further investigation. In addition to the carcass measures, the abattoir data provided information on date of birth, kill date, breed and sex. However, in some cases these were incomplete, thus data merging with BCMS data was used to fill in some of the gaps. The three additional traits available from abattoir records were carcass weight, conformation and fat class .

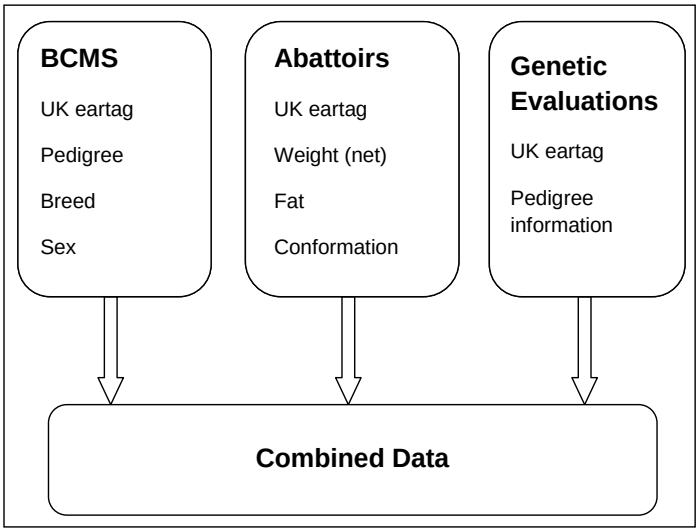


Figure 1. Compiling data from the available information sources

Pedigree data. A pedigree file was created based upon BCMS records and by matching to other national data sources to provide as much pedigree information as possible. This resulted in a pedigree file containing over 50 million animals going back a maximum of 13 generations. The pedigree data contained over 2.3 million additional records obtained from milk recording organisations (MROs) and BASCO.

Editing the data for genetic analysis. A subset of the abattoir/BCMS matched data was extracted for all slaughter animals with Charolais as a sire breed (the third most numerous sire breed). Various edits were then carried out to create a file for genetic parameter estimation, resulting in a considerable reduction in the size of the dataset (as expected), with 28% of the animals remaining. The pedigree was extracted for 6 generations and consisted of 43,069 animals.

Data analysis. The complete combined dataset were described and results are presented below. Genetic analyses were conducted for the traits: carcass weight, conformation, carcass fat grades using ASReml (Gilmour *et al*. 2009). In each case the model accounted for the fixed effects of sex, dam breed, birth herd, birth-year-season, location of death, finishing herd-year-season; linear and

quadratic covariates of age at slaughter; and, the random effect of the animal. A year consisted of two seasons for the definition year-season due to relatively small contemporary group size.

RESULTS AND DISCUSSION

Description of BCMS and abattoir data. In BCMS data the five breeds (including crosses) with most animals registered each year were Holstein Friesian, Limousin, Charolais, Simmental and Aberdeen Angus. The major breeds present in the slaughter population were Limousin, Aberdeen Angus, Holstein Friesian, Charolais, Hereford, Simmental, and Belgian Blue, and these accounted for 92% of the animals present in the matched abattoir/BCMS dataset. Records on dam breed emphasise that dairy cows are a major component of beef production with Holstein-Friesian being the most common dam breed, accounting for 46% of the slaughter population. The majority of offspring from dairy cows were crossbreds and these were most commonly sired by Aberdeen Angus or Hereford bulls. Similar results of breed distribution in the UK are shown by Todd *et al*., (2011)

Although sire identity it is not compulsory to record it was known for approximately 23% of animal records in BCMS, with the level of recording generally higher in more recent years (e.g., 11% in 2001 and 23% in 2011). When the BCMS records were merged with other pedigree sources (e.g., MRO data, breed society data) 25% of slaughter animals had a sire record.

Table 1. Overall distribution of carcasses for fat and conformation class in animals slaughtered from 3 to 36 months of age

Fat class	%	Conf. class	%
1	0.6	E	0.6
2	6.8	U	14.8
3	28.5	R	41.8
4L	50.9	O	41.4
4H	11.8	P	1.4
5L	1.3		
5H	0.1		

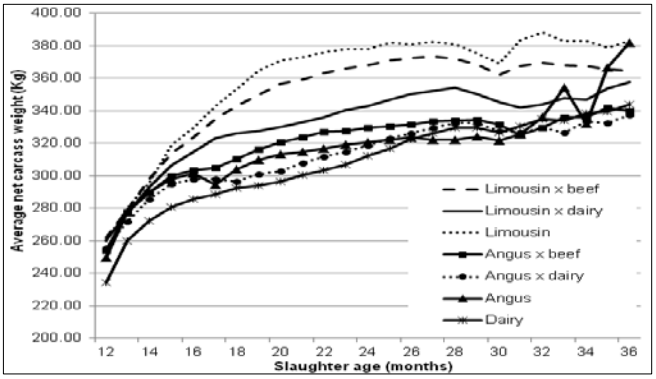


Figure 2 Average carcass weight of male animals grouped by slaughter age for different breed types of cattle

Just over 1.9 million animals were slaughtered between 3 and 36 months of age. The average carcass weight of males and females was 335 kg and 298 kg respectively. Average conformation and fat were equivalent to –R and +3/-4, respectively (Table 1). Overall, the mean number of days to slaughter and average carcass daily weight gain were 743 days and 0.45 kg, respectively. Carcass traits differ between breed types as illustrated in Figure 2.

Table 2. Genetic parameter estimates of carcass traits in Charolais using an animal model

	Animal variance	Residual variance	Phenotypic variance (s.e.)	Heritability (s.e.)
Carcass wt.	196.80	447.07	643.9 (9.89)	0.31 (0.04)
Conformation	2.29	7.20	9.48 (0.14)	0.24 (0.04)
Fat	1.72	10.48	12.21 (0.17)	0.14 (0.03)

Genetic Parameters. Heritability estimates (and standard errors) for carcass weight, conformation, and fat class were 0.31 (0.04), 0.24 (0.04), and 0.14 (0.03) respectively, using an animal model (Table 2). Genetic correlations between carcass weight and conformation, carcass weight and fat, and conformation and fat were 0.38 (0.09), -0.54 (0.12), and -0.67 (0.11). These estimates are in line with other studies (e.g., Gregory *et al.*1995, Hickey *et al*. 2007).

CONCLUSIONS

There is a wealth of data recorded in the UK, some of it being compulsory, which could have uses other than its original purpose (for instance BCMS), and when combined with other data sources, provide added value. The creation of a pedigree file has allowed us to link dairy and beef genetics as well as opening up the opportunity to perform genetic analysis for traits in the commercial populations with multiple breeds and crosses represented. As well as carcass traits, the merging of pedigree information also provides opportunities for other traits that are affecting both dairy and beef populations.

The results indicate that genetic analysis for carcass traits is realistic, particularly for breeds which make up a major part of the cattle population and where sire identity is recorded. The use of carcass trait evaluations should reduce the current knowledge gap between the pedigree breeders and the commercial beef producer. Providing clearer signals to pedigree breeders on where improvements need to be made should lead to benefits being filtered down into commercial beef production with cattle that perform more efficiently and hit market specifications. To move the cattle industry forward, the various parts of the food chain need to work together and share information, which in part this study has demonstrated.

ACKNOWLEDGEMENTS

We would like to acknowledge the abattoirs (ABP, Stoddarts, Dovecote Park), BCMS, and EGENES for supplying data. Funding was received from English Beef and Lamb Executive (EBLEX), DairyCo and Hybu Cig Cymru (HCC).

REFERENCES

Amer P. R., Crump R. and Simm G. (1998) *Animal Science* **67**: 445.

Amer P. R., Nieuwhof G. J., Pollott G. E., Roughsedge T., Conington J., and Simm G. (2007) *Animal* **1**: 1414.

Gilmour A. R., Gogel B. J., Cullis B. R., Thompson R., and Butler D. (2009). ASReml user guide release 3.0. VSN International Ltd, Hemel Hempstead, UK.

Gregory K. E., Cundiff L. V., and Koch R. M. (1995) *Journal of Animal Science* **73**: 1920.

Hickey J. M., Keane M. G., Kenny D. A., Cromie A. R., and Veerkamp R. F. (2007) *Journal of Animal Science* **85**: 314.

Todd D. L., Woolliams J. A., and Roughsedge T. (2011) *Animal* **5:** 1874.

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:257-260

COMPARISON OF MEASURES OF RELATEDNESS USING PEDIGREE OR GENOMIC DATA IN A MULTI-BREED SHEEP POPULATION

B. Auvray and K.G. Dodds

AgResearch Ltd, Invermay Agricultural Centre, Mosgiel 9053, New Zealand

SUMMARY

Numerator relationship matrices (NRM) between individuals based on SNP genotypes, estimated using a method proposed by VanRaden (2008), and combined with modifications which rescale the NRM and which account for population substructure were compared with **A**, the NRM derived from pedigree. Getting matrices closely resembling the **A** matrix may be desirable because, in a crossbred or multi-breed context, the elements of **A** (particularly off-diagonal elements between breeds) are closer to the true average identity by descent between individuals. On a crossbred sheep data set of 7,855 individuals genotyped for 47,084 SNP, the NRM where population stratification was not accounted for performed poorly (overall mean absolute difference (MAD) from pedigree relatedness = 2.8%, MAD in Texel, the most differentiated breed in the dataset in term of allele frequencies, = 24.6%) while NRM corrected for population structure performed better (overall MAD = 1.2%, MAD in Texel = 2.9%). The impact of the rescaling was marginal, as it only reduced the overall and per breed MAD from pedigree by up to 0.1%.

INTRODUCTION

Many methods to estimate numerator relationship matrices (NRM) from SNP to use e.g. in genome enabled prediction (GEP, Meuwissen *et al.* 2001) have been proposed, as by VanRaden (2008), or, in a crossbred context, by Harris and Johnson (2010). Here we suggest modifications to account for population structure and to rescale these NRM and we compare them with the pedigree NRM **A** (Henderson 1976) in a NZ sheep data set. Getting NRM closely resembling **A** may be desirable because, in a crossbred or multi-breed context, the elements of **A** (particularly off-diagonal elements between breeds) are often closer to the true average identity by descent between individuals. Nevertheless increased similarity with **A** does not equal increased accuracy in GEP.

MATERIALS AND METHODS

Data. A data set of $n = 7{,}855$ animals, sourced from industry and research flocks, with Sheep Improvement Ltd (SIL, http://www.sil.co.nz) pedigree records and Illumina OvineSNP50 BeadChip (http://www.illumina.com) genotypes was used for this analysis. The animals were mostly sires born between 1986 and 2010 of pure and composite recorded breeds of Romney (R), Coopworth (C), Perendale (P) and Texel (T).

Genotypes were cleaned (Dodds *et al.* 2009), which included filtering SNP on call rate, quality score (from the Illumina scoring algorithm), monomorphism, and extreme departure from Hardy-Weinberg Equilibrium (HWE). Any SNP that were not retained as part of the ovine HapMap project (http://www.sheephapmap.org) or were denoted or appeared to be X-linked (including pseudo-autosomal) were removed, leaving $k = 47{,}084$ SNP of the initial 53,903 SNP on the chip.

The pedigree was extracted as deep as possible (up to 23 generations) from the SIL database and consisted of 41,087 animals (including the 7,855 genotyped animals) born between 1969 and 2010. No effort was made to correct the recorded pedigree using the SNP genotypes.

Breed groups. Animals were assigned to 6 groups according to the following definitions. 'Pure' bred R, C, P and T were defined as being ≥ 75% of that breed. Two groups of composites were defined for those animals not achieving this purity definition. cRCP have ≥ 50% of R, C and P combined, and < 25% T. cRCPT have ≥ 50% of R, C and P combined, and ≥ 25% T. These

definitions were applied after decomposing recently developed 'breeds' into their R, C, P and T components, as far as possible, by estimating their breed proportions using the methodology presented in Dodds *et al.* (2012). Table 1 shows the number of animals per breed group.

Table 1. Number of animals per breed group

Breed group	Total	R	C	P	cRCP	T	cRCPT
Number of animals	7,855	4,270	1,697	551	777	317	243

Statistical tools and notation. Here the sum of the elements of matrix **X** (or vector **x**) is denoted $\sum x_{ij}$ (or $\sum x_i$), omitting index and bounds of summation when appropriate. The mean of the elements of **X** (and similarly for vectors) is denoted $\bar{\mathbf{X}}$ and $\bar{\mathbf{X}} = \sum x_{ij}/N$, where *N* is the number of elements of **X**. The difference between NRM was assessed using the mean absolute difference (MAD), calculated as $MAD(\mathbf{X},\mathbf{Y}) = \sum|x_{ij} - y_{ij}|/N$ for matrices **X** and **Y** (and likewise for 2 vectors). Data manipulation and analysis was done in R (R Development Core Team 2012).

Measures of relatedness between individuals. A was computed for all 41,087 animals and only the sub-matrix corresponding to the 7,855 animals genotyped was kept.

Genomic NRM (**G**) were obtained using the methods described below. Care was taken to use only methods producing (in the worst case semi) positive definite **G**. First, from VanRaden (2008):

$$\mathbf{G}_a = \mathbf{Z}_a\mathbf{Z}_a'/(2\sum p_j(1-p_j))$$

where $\mathbf{Z}_a = \mathbf{M} - 2\mathbf{P}$, with **M** being the $n \times k$ matrix of SNP genotypes m_{ij} scored as 0, 1 or 2 for animal *i* with respectively a BB, AB or AA call for SNP *j*, and **P** a matrix of allele frequencies (AF), whose column $\boldsymbol{p} = \mathbf{1}p$, with **1** a vector of 1's of size *n* and p the frequency of the 'A' allele for a SNP, calculated on the entire population. A second matrix $\mathbf{G}_b$ was created as:

$$\mathbf{G}_b = \mathbf{Z}_b^*\mathbf{Z}_b^{*\prime}$$

where $\mathbf{Z}_\mathrm{b} = \mathbf{M} - 2\mathbf{P}_\mathrm{b}$ and $\mathbf{P}_\mathrm{b} = \boldsymbol{\Lambda}\mathbf{P}_{\boldsymbol{\lambda}}$ is a $n \times k$ matrix of AF pertaining to each animal, $\boldsymbol{\Lambda}$ being a $n \times l$ matrix whose element λ_{ij} is the proportion of breed *j* for animal *i* for a total of *l* breeds (fixed *a priori*), and $\mathbf{P}_{\boldsymbol{\lambda}}$ a matrix of AF estimates for each SNP and each breed. $\mathbf{Z}_\mathrm{b}^*$ is a rescaled version of $\mathbf{Z}_\mathrm{b}$ so that each element $z^*_{\mathrm{b}ij} = z_{\mathrm{b}ij}/\sqrt{\mathbf{2}\sum_{j=1}^{k} p_{b\,ij}\left(1-p_{b\,ij}\right)}$, with $p_{b\,ij}$ being the element of $\mathbf{P}_\mathrm{b}$ relating to animal *i* and SNP *j*. This method of calculation tries to account for AF differences between breeds when estimating **G** in multi-breed populations. This topic has been discussed extensively in Harris and Johnson (2010). Next, 2 variations of the 2 methods above were devised where we tried to rescale **G** so that $\bar{\mathbf{G}} = \bar{\mathbf{A}}$. The first variation is a convex combination of **G** and a constant β:

$$\mathbf{G}_x^* = \pi\mathbf{G}_x + (1-\pi)\mathbf{1}\mathbf{1}'\beta$$

where $x = a$ or b, $\pi \in [0,1[$ and $\beta = (\bar{\mathbf{A}} - \pi\bar{\mathbf{G}})/(1-\pi)$. The second variation rescales **G** by using adjusted AF. The expected contribution $E[c_{ijk}]$ of 1 SNP *k* in HWE to element g_{ij} of **G** relating to animals *i* and *j* is 0 if they are unrelated and come from the same population. This can be tested by noting that $E[c_{ijk}] = E[\mathbf{v}\mathbf{v}'] = \sum(\mathbf{u}\mathbf{u}' \circ \mathbf{v}\mathbf{v}') = 0$, where $\mathbf{u} = (1-p^2, 2p(1-p), p^2)$ the vector of genotype probabilities for a bi-allelic marker under HWE, $\mathbf{v} = (-2p, 1-2p, 2-2p)$ the vector of centred genotypes and the operator $\circ$ denotes the Hadamard (entrywise) product. If we adjust the AF used when calculating **G** by adding a constant δ, the expected contribution of 1 SNP *k* in HWE is now $E\left[c_{\delta\,ijk}\right] = \sum(\mathbf{u}\mathbf{u}' \circ \mathbf{v}_\delta\mathbf{v}_\delta') = 4\delta^2$, where $\mathbf{v}_\delta = (-2(p+\delta), 1-2(p+\delta), 2-2(p+\delta))$. For $\mathbf{G}_a$, we can now choose δ_a as a root of the quadratic equation:

$$2k\delta_a^2 - (\bar{\mathbf{A}} - \bar{\mathbf{G}}_a)\sum_{j=1}^{k}(p_j + \delta_a)\left(1 - p_j - \delta_a\right) = 0$$

to construct $\mathbf{G}_a^{**}$ as $\mathbf{G}_a$, but replacing $\mathbf{P}$ by $\mathbf{P}^* = \mathbf{P} + \mathbf{1}\mathbf{1}'\delta_a$. Similarly for $\mathbf{G}_b$, δ_b satisfies equation:

$$2k\delta_b^2 - (\bar{\mathbf{A}} - \bar{\mathbf{G}}_a)\Sigma = 0$$

where $\Sigma = (\sum_{i=1}^{n}\sum_{\boldsymbol{j=1}}^{\boldsymbol{k}} 2(p_{b_{ij}} + \delta_b)\left(1 - p_{b_{ij}} - \delta_b\right))/n$ with $p_{b_{ij}}$ the element of $\mathbf{P}_\mathrm{b}$ relating to animal *i* and SNP *j*, and $\mathbf{G}_b^{**}$ can be constructed as $\mathbf{G}_b$ but replacing $\mathbf{P}_b$ by $\mathbf{P}_b^* = \mathbf{P}_b + \mathbf{1}\mathbf{1}'\delta_b$.

RESULTS AND DISCUSSION

Mean pedigree relatedness between groups. Overall, $\bar{\mathbf{A}} = 0.0048$. Table 2 reports the mean inbreeding coefficient ($\bar{\mathbf{f}} = \overline{diag(\mathbf{A}) - \mathbf{1}}$) and $\bar{\mathbf{A}}$ within (ignoring the diagonal) and between breed groups. Relatedness within groups ranged from 0.008 (R) to 0.028 (C and cRCPT). Relatedness between groups ranged from < 5 × 10^{-4} to 0.015 (C × cRCP and C × cRCPT).

Table 2. $\bar{\mathbf{A}}$ (%) and $\bar{\mathbf{f}}$ (%) overall between breed groups

Breed group	$\bar{\mathbf{f}}$	R	C	P	cRCP	T	cRCPT	Total
R	2.0	0.8	0.0	0.0	0.3	0.0	0.1	
C	2.4		2.8	0.0	1.5	0.1	1.5	
P	1.1			1.6	0.1	0.0	0.0	
cRCP	1.1				1.5	0.2	1.2	
T	1.2					2.2	1.0	
cRCPT	0.9						2.8	
Total	1.8							0.5

Mean genomic relatedness and comparison with pedigree relatedness. Overall, $\bar{\mathbf{G}}_a = 0.0000$ and $\bar{\mathbf{G}}_b = 0.0018$. The roots δ_a for the AF adjustment were $\delta_a = (-0.0211, 0.0212)$. The roots δ_b were $\delta_b = (-0.0166, 0.0166)$. Fixing $\pi = 0.99$ so that the diagonal of $\mathbf{G}_x^*$ are not shrunk down excessively, using any of the 2 methods of rescaling $\mathbf{G}$ and any estimate of δ_x lead to virtually the same matrix, as $MAD(\mathbf{G}_x^*, \mathbf{G}_x^{**})$ ranged from 3.6 × 10^{-3} to 8.4 × 10^{-3}. It is worth noting that $\pi = 0.99$ is not the value of π minimising $MAD(\mathbf{G}_x^*, \mathbf{A})$. These are $\pi_a = 0.1106$ and $\pi_b = 0.2419$, that produce $\mathbf{G}_a^*$ and $\mathbf{G}_b^*$ that are unreasonably shrunk (ideally, diagonal elements should be kept ≥ 1 as much as possible), because $\mathbf{A}$ is very sparse. A potential improvement would be to minimise $MAD(\mathbf{G}_x^*, \mathbf{A})$ only for elements of $\mathbf{A}$ reaching a certain threshold. Table 3 reports $\bar{\mathbf{f}}$, $\bar{\mathbf{G}}$ and $MAD(\mathbf{G}, \mathbf{A})$ within and between groups respectively for $\mathbf{G}_a$, $\mathbf{G}_b$, $\mathbf{G}_a^*$ and $\mathbf{G}_b^*$. $MAD(\mathbf{G}_a, \mathbf{A})$ within Texel (24.6%) and between groups with Texel or cRCPT was very high. Using $\mathbf{G}_a^*$ instead of $\mathbf{G}_a$ slightly increased $MAD(\mathbf{G}, \mathbf{A})$ within breed (up to 0.4%), but somewhat reduced $MAD(\mathbf{G}, \mathbf{A})$ overall (0.1%) and between breeds (up to -0.5%). Using $\mathbf{G}_b$ reduced overall and per breed $MAD(\mathbf{G}, \mathbf{A})$ and $MAD(\mathbf{f_G}, \mathbf{f_A})$ dramatically compared to $\mathbf{G}_a$, especially for Texel (2.9%). $\mathbf{G}_b^*$ lead to a slight decrease in $MAD(\mathbf{G}, \mathbf{A})$ over $\mathbf{G}_b$ (0.1%). The values of $MAD(\mathbf{f_G}, \mathbf{f_A})$ and $MAD(\mathbf{G}, \mathbf{A})$ within breed obtained with $\mathbf{G}_a$ (and $\mathbf{G}_a^*$) were very highly correlated with $MAD(\mathbf{p}_{\text{total}}, \mathbf{p}_{\text{breed}})$, the MAD between AF calculated overall and per breed, with correlations of respectively 0.935 and 0.958. Together with the extremely high $MAD(\mathbf{G}_a, \mathbf{A})$ in Texel, this suggested that $\mathbf{G}_a$ (and hence $\mathbf{G}_a^*$) is not well suited to predict $\mathbf{A}$ in a crossbred situation. $\mathbf{G}_b$ and $\mathbf{G}_b^*$ on the other hand predicted $\mathbf{A}$ reasonably well. The impact of rescaling the matrices was marginal.

Table 3. Within breed group $\bar{\mathbf{f}}$, $MAD(\mathbf{f_G}, \mathbf{f_A})$, $\bar{\mathbf{G}}$ and $MAD(\mathbf{G}, \mathbf{A})$, and between group $\bar{\mathbf{G}}$ (above diagonal) and $MAD(\mathbf{G}, \mathbf{A})$ (below diagonal) using different G, all in %

G	Breed group	$\bar{\mathbf{f}}$	$MAD(\mathbf{f_G}, \mathbf{f_A})$	$\bar{\mathbf{G}}$	$MAD(\mathbf{G}, \mathbf{A})$	R	C	P	cRCP	T	cRCPT
$\mathbf{G}_a$	R	1.5	2.4	2.7	2.4		-3.2	-0.7	-1.7	-4.8	-3.3
	C	4.7	2.9	6.7	4.1	3.3		-1.4	2.4	0.5	3.0
	P	6.9	6.1	8.5	7.0	1.1	1.5		0.0	-0.2	-0.7
	cRCP	2.3	2.8	1.6	1.6	2.4	2.1	1.4		1.6	1.9
	T	21.6	20.4	26.7	24.6	5.1	1.6	1.0	2.8		9.9
	cRCPT	4.0	3.3	6.3	3.7	3.5	2.3	1.3	1.8	8.9	
	Total	4.5	4.4	0.0	2.8						
$\mathbf{G}_a^*$	R	1.0	2.5	3.2	2.7		-2.7	-0.2	-1.2	-4.3	-2.7
	C	4.1	2.6	7.1	4.5	2.8		-0.9	2.8	1.0	3.4
	P	6.3	5.5	8.9	7.4	1.0	1.2		0.5	0.3	-0.2
	cRCP	1.8	2.6	2.1	1.6	2.1	2.2	1.4		2.1	2.4
	T	20.9	19.7	27.0	24.8	4.6	1.7	0.9	2.9		10.3
	cRCPT	3.4	2.9	6.7	4.1	3.0	2.5	1.1	1.9	9.3	
	Total	4.0	4.3	0.5	2.7						
$\mathbf{G}_b$	R	3.1	3.1	0.0	2.0		0.0	0.0	0.1	0.0	0.0
	C	2.1	2.6	1.4	2.2	0.8		0.0	1.2	-0.2	1.1
	P	-0.1	3.1	-0.2	2.6	0.8	0.7		0.0	0.0	0.0
	cRCP	1.4	2.4	1.6	1.3	1.0	1.3	0.8		-0.4	1.1
	T	0.1	3.6	1.8	2.9	0.7	0.8	0.6	1.2		0.2
	cRCPT	-1.3	3.1	2.3	1.5	0.8	1.3	0.7	1.2	1.6	
	Total	2.6	3.3	0.2	1.2						
$\mathbf{G}_b^*$	R	2.3	3.0	0.3	1.9		0.3	0.3	0.4	0.3	0.3
	C	1.4	2.8	1.7	2.1	0.8		0.3	1.5	0.2	1.4
	P	-0.8	3.3	0.1	2.4	0.8	0.7		0.3	0.3	0.3
	cRCP	0.6	2.5	1.9	1.3	1.0	1.3	0.8		-0.1	1.3
	T	-0.6	3.9	2.1	2.8	0.7	0.8	0.6	1.1		0.5
	cRCPT	-2.0	3.6	2.5	1.4	0.8	1.3	0.7	1.1	1.5	
	Total	1.9	3.3	0.5	1.2						

ACKNOWLEDGMENTS

Financial assistance was provided by Ovita Ltd (consortium between AgResearch Ltd and Beef + Lamb NZ). Data was made available by SIL.

REFERENCES

Dodds K. G., Auvray B., Pickering N. and McEwan J. C. (2009) *Proc. Assoc. Advmt. Anim. Breed. Genet.* **18**: 296.

Dodds K. G., Auvray B., Newman S.-A. N. and McEwan J. C. (2012) *Australasian Applied Statistics Conference (GenStat & ASReml), Queenstown, New Zealand.* http://www.aasc2012.com/file/Dodds-2012-AASC-talk.pdf. (Accessed 24 January 2013.)

Harris B. L. and Johnson D. L. (2010) *J. Dairy Sci.* **93**: 1243.

Henderson C. R. (1976) *Biometrics* **32**: 69.

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

R Development Core Team (2012) 'R: A language and environment for statistical computing' R Foundation for Statistical Computing, Vienna, Austria.

VanRaden P. M. (2008) *J. Dairy Sci.* **91**: 4414.

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:261-265

COMPARISONS OF IDENTICAL BY STATE AND IDENTICAL BY DESCENT RELATIONSHIP MATRICES DERIVED FROM SNP MARKERS IN GENOMIC EVALUATION

S.A. Clark[1,2], B.P Kinghorn[1] and J.H.J. van der Werf[1,2]

[1] University of New England, Armidale, NSW 2351
[2] CRC for Sheep Industry Innovation, University of New England, Armidale, NSW 2351

SUMMARY

In animal populations, family members inherit alleles through common ancestors and these shared regions are referred to as identical by descent (IBD). Furthermore, animals may also share alleles due to a random association with no known common inheritance pattern. This identity by state (IBS) also covers known relationships, such that regions that are IBD are also IBS, however regions that are IBS may not always be IBD. In the genetic evaluation of livestock, IBD and IBS information can be used to build the genomic relationship matrix (GRM) and breeding values can be predicted using genomic best linear unbiased prediction (gBLUP).

This study compares a number of different methods to construct the GRM, using IBD and IBS information. Each method was evaluated using a reference dataset of 1781 Merino sheep and validated using 164 progeny tested sires that had accurate breeding values. Estimates of variance components were also compared. There was no significant difference between the accuracy achieved by the IBS and IBD methods. However the accuracy of the EBVs decreased as a greater restriction was applied to whether a region was IBD or not IBD. Furthermore, estimates of variance components were substantially different for IBD and IBS methods.

INTRODUCTION

In animal populations there is often a high resemblance between the phenotypes of family members due to genes inherited from common ancestors (Fisher 1918). This theory has been widely discussed in the field of quantitative genetics and is currently used for the prediction of merit in livestock and detection of disease in humans (Henderson 1975; Donnelly 1983). In livestock genetic evaluation, best linear unbiased prediction (BLUP) (Henderson 1975) uses this concept to form the co-variances among the phenotypes of known relatives through the use of a numerator relationship matrix (NRM). Included in this matrix are coefficients of relationships which are the expected proportion of alleles that individuals share in common, identical by descent (IBD) based on pedigree information. Theories and methods using the same principles have also been described for the estimation of variance components.

Marker information has already been included in mixed model analyses (BLUP) using a relationship matrix derived from these markers, called the genomic relationship matrix (GRM) (Visscher *et al*. 2006; VanRaden 2008). This matrix can potentially describe the underlying covariance structure among individuals more fully than a matrix based on pedigree information alone, because the GRM uses estimates of realised relationships rather than expected relationships (Hayes *et al*. 2009). Popular methods for forming the genomic relationship matrix have been described by VanRaden (2008) and Yang *et al*. (2010). These methods use identical by state (IBS) information which is scaled by the allele frequencies to build the GRM, as shared rare alleles are more likely to be IBD than common alleles. However, these methods do not explicitly differentiate between IBD and IBS information. In contrast, there are very few methods that explicitly define IBD and often these methods only perform as well as IBS methods. However, many of these IBD methods have only been used in simulation and therefore constrained by the model used to simulate variation (Calus *et al.* 2008; Hickey *et al.* 2013).

In human quantitative genetics there has been a large focus placed on IBD information (Thompson 2008). Often, genotype probabilities that accommodate the probabilities of cross over events are used for determining IBD between individuals (Donnelly 1983) i.e. the more distant the relationship between two individuals the higher the probability that many crossovers have occurred. Many methods and programs have been described for the estimation of IBD and have been used for the detection of regions of the genome that are IBD, e.g. PLINK (Purcell *et al*. 2007) and fastIBD (Browning and Browning 2011).

The aim of this study was to compare the use of IBD and IBS genomic relationship information to predict genomic breeding values using real data. The differences between each GRM were investigated, together with their effects on the estimation of breeding values and variance components, and the accuracies of resulting estimates of breeding value (EBVs).

METHODS

The data used in this study consisted of phenotypic and genotypic records from the Australian Sheep Cooperative Research Centre (CRC) information nucleus flock (INF). This dataset consisted of a reference dataset consisted of phenotypic and genotypic records from 1781 merino animals and a validation dataset of 164 merino sires with accurate Australian Sheep Breeding Values (ASBV's). Definitions of ASBV's can be obtained from Sheep Genetics Australia. Phenotypic information on the trait scanned eye muscle depth (SEMD) was analysed. To observe the effect of relatedness, the validation population was split into three groups based on their pedigree relationship to the animals in the reference dataset (Clark *et al*. 2012). The three groups consisted of; 50 closely related animals (Close), with a maximum relationship of greater than 0.25; the 54 distantly related animals (0.01-0.249) (Dist); and 60 unrelated animals that shared zero pedigree relationship.

All animals in each dataset were genotyped using the Illumina 50K ovine SNP chip. All SNP in this dataset underwent a number of genotyping quality control measures (see Daetwyler *et al*. (2010)). The following fixed effects were fitted in the analysis of SEMD: Sex, birth type, rearing type, age of dam, contemporary group (birth year • birth month) (site • management group), age-at-trait recording and live weight at scanning.

As in Daetwyler *et al*. (2010) we assumed the gBLUP model;

$$y = Xb + Zg + e$$

where y is a vector of phenotypes, X is a design matrix relating the fixed effects (as described above) to each animal, b is a vector of fixed effects, Z is a design matrix allocating records to breeding values, g is a vector additive genetic effects for animals in the reference dataset and the validation dataset and e is a vector of random normal deviates σ_e^2. Furthermore V(g) = Gσ_g^2 where G is the genomic relationship matrix, and σ_g^2 is the genetic variance for this model. The GRM (G) was formed using two IBS methods defined by (G_V) VanRaden (2008) and (G_Y) Yang *et al*. (2010) and five IBD methods were also evaluated. Two fastIBD matrices (Browning and Browning 2011) were formed. $G_{Fast(h)}$ was based on the stringent threshold for IBD used in human genetics and $G_{Fast(R)}$ used a relaxed threshold on whether a region was IBD or not. Three probability of IBD methods (Kinghorn 2012) were also used: G_{Prob} used an IBD probability estimate for each individual loci that was based on IBD information from adjacent marker information. This method was extended such that regions were identified as IBD if animals shared haplotypes of 10 (G_{Prob10}) and 50 (G_{Prob50}) markers with an IBD probability of greater than 0.98 and if regions were shorter than the given length they were assumed to be IBS and did not contribute to the estimate of relationship.

RESULTS

The IBS and unrestricted IBD methods (G_{Prob} and $G_{Fast(R)}$) were the most accurate methods to predict breeding value (Table 1). These results are similar to simulation studies by Hickey *et al.* (2013) and Calus *et al.* (2008) were there was little difference between the IBS and best IBD methods. However, accuracy was reduced when a restriction was placed on whether a region was IBD or not, by either increasing the length of the IBD segment as in G_{Prob10} and G_{Prob50} or by increasing the significance threshold as in $G_{Fast(h)}$. The highly restricted fastIBD ($G_{Fast(h)}$) method was the least accurate method (Table 1).

Table 1 The EBV accuracy (correlation between gEBV and ASBV) and regression of gEBV on ASBV, estimated using the alternative ways to define the genomic relationship matrix

	IBS*		IBD[i]				
	G_V	**G_Y**	**G_{Prob}**	**G_{Prob10}**	**G_{Prob50}**	**$G_{Fast(R)}$**	**$G_{Fast(h)}$**
Accuracy							
All Animals	0.456	0.451	0.453	0.413	0.340	0.465	0.239
Unrelated	0.224	0.206	0.226	0.226	0.172	0.281	0.137
Distantly related	0.450	0.499	0.478	0.394	0.334	0.434	0.216
Closely related	0.640	0.643	0.650	0.622	0.555	0.668	0.413
Regression							
All Animals	0.882	0.873	0.914	1.033	1.249	1.011	0.834

* IBS methods were constructed using methods by VanRaden (**G_V**) and Yang *et. al.* (**G_Y**)

i IBD methods were constructed using: 1) IBD probabilities (**G_{Prob}**) with different haplotype lengths (**G_{Prob10}** and **G_{Prob50}**) and 2) the FastIBD module of the Beagle software (**G_{Fast}**) with either a relaxed (**$G_{Fast(R)}$**) or strict (**$G_{Fast(h)}$**) constraint on whether a region was IBD or not.

When animals were unrelated or distantly related to the reference population, accuracy was reduced for both IBD and IBS methods. Accuracy decreased in all cases when the IBD segment length increased. Furthermore, when fast IBD was highly restricted ($G_{Fast(h)}$), its ability to predict breeding value in unrelated animals was also reduced. A reduction in accuracy was observed because, as the restriction on whether a region was IBD or not increased, some useful information about rare, short haplotypes was lost. Interestingly, in unrelated animals, the $G_{Fast(R)}$ tended to be the most accurate method (although not significantly better than G_V, G_Y or G_{Prob}).

Table 1 also shows the regression of GEBV on ASBV for each of the different GRMs. It shows that the IBS and G_{Prob} methods had a regression coefficient less than one, showing the GEBVs are over dispersed. In contrast, the $G_{Fast(R)}$ and $G_{Prob(10)}$ methods had a regression coefficient close to 1 showing that both sets of EBV's are on a similar scale to the progeny tested ASBV's. The IBS methods: G_V and G_Y are very similar and resulted in a 0.999 correlation between the breeding values estimated using these methods. The $G_{Fast(R)}$ and G_{Prob} methods were only slightly different with a correlation between breeding values of 0.96 and 0.94 respectively with the IBS methods. Finally, $G_{Prob(10)}$ used partially different information as the breeding values estimated from this method were only 0.88 correlated with G_V. Although the methods appear to be very similar, given the high correlation between breeding values, the variance components (Table 2) estimated from each method were different.

The G_V and G_Y methods by VanRaden (2008) and Yang *et al.* (2010) resulted in similar variance component estimates. The ProbIBD methods (G_{Prob} and G_{Prob10}) also resulted in higher estimates of genetic variance. In contrast, the fastIBD method ($G_{Fast(R)}$) resulted in a substantially lower estimate of genetic variance and therefore heritability. This implicitly shows that the scale of the various GRM's (which relates to the methods used to construct each GRM) can have a large impact on variance component estimation.

Table 2 Variance components estimated using various methods to define the genomic relationship matrix

	Pedigree	IBS		IBD				
		G_V	G_Y	G_{Prob}	G_{Prob10}	G_{Prob50}	$G_{Fast(R)}$	$G_{Fast(H)}$
Va	1.12	1.288	1.305	1.883	1.545	1.636	0.715	0.904
Ve	3.03	3.021	3.015	2.935	2.778	2.646	3.635	3.387
Vtotal	4.15	4.309	4.320	4.818	4.323	4.282	4.350	4.291
h^2	0.269	0.299	0.302	0.391	0.357	0.382	0.16	0.211

[*] IBS methods were constructed using methods by VanRaden ($\mathbf{G_V}$) and Yang *et. al.* ($\mathbf{G_Y}$)

[i] IBD methods were constructed using: 1) IBD probabilities ($\mathbf{G_{Prob}}$) with different haplotype lengths ($\mathbf{G_{Prob10}}$ and $\mathbf{G_{Prob50}}$) and 2) the FastIBD module of the Beagle software ($\mathbf{G_{Fast}}$) with either a relaxed ($\mathbf{G_{Fast(R)}}$) or strict ($\mathbf{G_{Fast(h)}}$) constraint on whether a region was IBD or not.

CONCLUSION

This study shows that IBD probabilities and information from the fastIBD module of Beagle can be used to predict breeding value in real data. Furthermore, this study has shown that some IBD relationship matrices will perform as well as IBS based methods for genomic evaluation, even in unrelated animals. However, accuracy will reduce when breeding values are estimated using IBD methods that place a large restriction on whether a region is IBD or not. The variance components estimated from each GRM is impacted by the scale of the relationship matrix. The scale is impacted by the definition of the relationship information (IBD or IBS) and the allele frequencies that are used to scale the GRM.

REFERENCES

Browning B.L. and Browning S.R. (2011). *Am. J. Hum. Genet.* **88**: 173.

Calus M.P.L., Meuwissen T.H.E., de Roos A.P.W. and Veerkamp R.F. (2008) *Genetics*. **178**:553.

Clark S.A., Hickey J.M., Daetwyler H.D. and van der Werf J.H.J. (2012). *Genet. Sel. Evol.* **44**: 4.

Daetwyler H.D., Hickey J.M., Henshall J.M. and Dominik S. (2010). *Anim. Prod. Sci.* **50**: 1004.

Donnelly K.P. (1983). *Theor. Pop. Biol.* **23**: 34.

Fisher R.A. (1918). *Trans. R. Soc. Edin.* **52**: 399.

Hayes B.J., Visscher P.M. and Goddard M.E. (2009). *Genet. Res.* **91**: 47.

Henderson C.R. (1975). *Biometrics* **31**: 423.

Hickey J.M., Kinghorn B.P., Teir B. Clark S.A. and van der Werf J.H.J (2013) *J Anim Breed Genet.* **130**: 259.

Kinghorn B.P. (2012). *Proc. Inter. Conf. Quant. Genet.* **32**:80.

Purcell S., Neale B., Todd-Brown K., Thomas L. and Ferreira M.A. (2007). *Am. J. Hum. Genet.* **81**:559.

Thompson E.A. (2008). *Theor. Pop. Biol*. **73**: 369.

VanRaden P.M. (2008). *J. Dairy Sci.* **91**: 4414.

Visscher P.M., Medland S.E., Ferreira M.A., Morley K.I. and Zhu, G. (2006). *PLoS Genet*. **2**:41.

Yang J., Benyamin B., McEvoy B.P., Gordon S.D., and Henders A.K. (2010). *Nat.Genet.* **42**:565.

Table 2 Variance components estimated using various methods to define the genomic relationship matrix

	Pedigree	IBS		IBD				
		G_V	G_Y	G_{Prob}	G_{Prob10}	G_{Prob50}	$G_{Fast(R)}$	$G_{Fast(H)}$
Va	1.12	1.288	1.305	1.883	1.545	1.636	0.715	0.904
Ve	3.03	3.021	3.015	2.935	2.778	2.646	3.635	3.387
Vtotal	4.15	4.309	4.320	4.818	4.323	4.282	4.350	4.291
h^2	0.269	0.299	0.302	0.391	0.357	0.382	0.16	0.211

[*] IBS methods were constructed using methods by VanRaden ($\mathbf{G_V}$) and Yang *et. al.* ($\mathbf{G_Y}$)

[i] IBD methods were constructed using: 1) IBD probabilities ($\mathbf{G_{Prob}}$) with different haplotype lengths ($\mathbf{G_{Prob10}}$ and $\mathbf{G_{Prob50}}$) and 2) the FastIBD module of the Beagle software ($\mathbf{G_{Fast}}$) with either a relaxed ($\mathbf{G_{Fast(R)}}$) or strict ($\mathbf{G_{Fast(h)}}$) constraint on whether a region was IBD or not.

CONCLUSION

This study shows that IBD probabilities and information from the fastIBD module of Beagle can be used to predict breeding value in real data. Furthermore, this study has shown that some IBD relationship matrices will perform as well as IBS based methods for genomic evaluation, even in unrelated animals. However, accuracy will reduce when breeding values are estimated using IBD methods that place a large restriction on whether a region is IBD or not. The variance components estimated from each GRM is impacted by the scale of the relationship matrix. The scale is impacted by the definition of the relationship information (IBD or IBS) and the allele frequencies that are used to scale the GRM.

REFERENCES

Browning B.L. and Browning S.R. (2011). *Am. J. Hum. Genet.* **88**: 173.

Calus M.P.L., Meuwissen T.H.E., de Roos A.P.W. and Veerkamp R.F. (2008) *Genetics*. **178**:553.

Clark S.A., Hickey J.M., Daetwyler H.D. and van der Werf J.H.J. (2012). *Genet. Sel. Evol.* **44**: 4.

Daetwyler H.D., Hickey J.M., Henshall J.M. and Dominik S. (2010). *Anim. Prod. Sci.* **50**: 1004.

Donnelly K.P. (1983). *Theor. Pop. Biol.* **23**: 34.

Fisher R.A. (1918). *Trans. R. Soc. Edin.* **52**: 399.

Hayes B.J., Visscher P.M. and Goddard M.E. (2009). *Genet. Res.* **91**: 47.

Henderson C.R. (1975). *Biometrics* **31**: 423.

Hickey J.M., Kinghorn B.P., Teir B. Clark S.A. and van der Werf J.H.J (2013) *J Anim Breed Genet.* **130**: 259.

Kinghorn B.P. (2012). *Proc. Inter. Conf. Quant. Genet.* **32**:80.

Purcell S., Neale B., Todd-Brown K., Thomas L. and Ferreira M.A. (2007). *Am. J. Hum. Genet.* **81**:559.

Thompson E.A. (2008). *Theor. Pop. Biol*. **73**: 369.

VanRaden P.M. (2008). *J. Dairy Sci.* **91**: 4414.

Visscher P.M., Medland S.E., Ferreira M.A., Morley K.I. and Zhu, G. (2006). *PLoS Genet*. **2**:41.

Yang J., Benyamin B., McEvoy B.P., Gordon S.D., and Henders A.K. (2010). *Nat.Genet.* **42**:565.

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:266-269

A STUDY ON EFFECTS OF FAMILY AND HAPLOTYPE BLOCKS ON CONSERVATION OF GENE EXPRESSION TRAITS IN HALF SIB SHEEP FAMILIES

H.A. Al-Mamun[1], P. Kwan[1], R.L. Tellam[2], J.W. Kijas[2] and C. Gondro[3]

[1]School of Science and Technology, University of New England, Armidale NSW 2351, Australia. [2]CSIRO Animal, Food and Health Sciences, Queensland Bioscience Precinct, St. Lucia, Brisbane, QLD 4067, Australia. [3]The Centre for Genetic Analysis and Applications, University of New England, Armidale NSW 2351, Australia

SUMMARY

The objective of this study was to explore the relationship between SNP and haplotype variation on gene expression traits. The data used included expression levels from 24,128 probe sets of *logissimus lumborum* muscle from 38 half-sib Poll Dorset sheep from six families and genotypes from 49,034 SNPs collected from the same animals. The analytical approaches used sought to analyse the effects of family and haplotype blocks on conservation of gene expression traits in this sheep population. Our study indicated that there is a genetic component in gene expression traits and hence gene expression is heritable to non-negligible extent. On average, our estimated heritability for gene expression obtained from skeletal muscle samples of sheep is 0.27 and 0.29 based on two different approaches. These preliminary results are consistent with previous heritability estimates.

INTRODUCTION

A primary goal in molecular biology is to understand how patterns of genetic variation affect the gene expression levels and higher level phenotypes. In recent years, studies of the relationship between genotype and gene expression, or other quantitative traits, have gained considerable attention due to the availability of high throughput technologies in profiling single nucleotide polymorphisms (SNPs) data and global gene expression. Several studies have suggested that the variation in gene expression traits is associated with genetic variation such as SNPs and copy number variants (CNVs) (Spielman *et al.* 2007; Stranger *et al.* 2007), and have demonstrated that a significant proportion of gene expression is heritable both in human (Cheung *et al.* 2003; Price, *et al.* 2011) and in other organisms (Nätt *et al.* 2012; Schadt *et al.* 2003). Most of these association studies comprised a large numbers of SNP from multiple individuals, and made use of the allele frequencies to search for associations with variation in trait data. One potential drawback of this approach is the large number of SNP–wise testing required and the potential for false positive outcomes. Moreover, these methods did not consider the information present in associations between neighbouring SNPs. Neighbouring SNPs tend to be inherited as blocks (Daly *et al.* 2001). These haplotype blocks can be used to find associations with quantitative traits such as gene expression traits. This strategy decreases the impact of multiple testing corrections as fewer hypotheses are tested.

In this study, SNPs and gene expression data obtained from 38 half-sib sheep were used to (i) quantify the heritability of gene expression in a sheep population and (ii) determine the degree of conservation of the gene expression between haplotype blocks within different families of the sheep population.

MATERIAL AND METHODS

Animals. 38 progeny (18 months old ewes) from six Poll Dorset sires (4-8 progeny/sire) were used for genotyping and microarray analysis of skeletal muscle samples. The six sires were

grouped into a high muscling and low muscling sire group (Table 1) based on their yearling trait, Eye Muscle Depth (EMD). Details of these sires have been described in (Kogelman *et al.* 2011).

Table 1 Number of progeny in each family

Family	2	5	7	11	16	17
No. of Progeny	7	8	4	8	8	5
Muscling Group	High	Low	High	High	Low	Low

Data pre-processing and normalization. The Affymetrix GeneChip® Bovine Genome Array (Affymetrix Santa Clara, CA) was used to measure the gene expression of the 40 animals. The Affymetrix GeneChip contains 24,128 probe sets, representing ~19,000 UniGene clusters. The microarray data were initially processed using the statistical software R (http://www.r-project.org) and additional Bioconductor packages (http://www.bioconductor.org). Normalization was performed using the RMA (Robust Multi-chip Average) method. After normalization and removing the control probe sets, 24,016 probe sets remained for further analysis. Linear Models for Microarray Data (limma) package from Bioconductor were used for differential gene expression analysis. Genotyping was undertaken using the Illumina 50K Ovine SNP chip containing 49,034 SNPs and 38 animals genotyped. The SNP data were pre-processed using the software PLINK (http://pngu.mgh.harvard.edu/~purcell/plink/) and 47,680 SNPs remained for further analysis. These 47,680 SNPs were subjected to phasing and haplotype block construction using the method described in (Ferdosi *et al.* 2013).

GRM and IBD estimates. The Genetic Relationship Matrix (GRM) was calculated according to the VanRaden's method (VanRaden, 2007) and the Identity by Decent (IBD) values for each pair of animals were computed using the method described in Price *et al.* 2011. The whole genome was partitioned into 2Mbp blocks and for each block 2x2=4 comparisons were performed between haplotypes from each pair of animals. For each chromosome, 0.5 Mbp from each end were excluded as these data might be noisy and could affect subsequent analysis. We did not consider sex chromosomes in our IBD calculation. Two haplotypes were considered IBD if they matched at > 95% of alleles in the block. Local IBD was defined as the total number of comparisons that produced a match. Genome-wide IBD was computed as the average of the local IBD estimates across all 2Mbp blocks.

Heritability estimates using IBD and GRM. Narrow sense heritability (Visscher *et al.* 2008) was calculated using variance-components analysis (Amos, 1994). We followed the method described in (Price, *et al.* 2011) and used their source code to calculate a heritability estimate for each gene. Let e_{gs} denote normalized gene expression of gene g for each individual animal s and θ_{st} denotes the genome-wide IBD or GRM between the individuals s and t ($0 \leq \theta_{st} \leq 1$). $\Theta = (\theta_{st})$ was assigned the NxN matrix of genome-wide IBD or GRM, where N is the number of animals. Vg was the covariance matrix of normalized gene expression for gene g. We fitted h_g^2, the heritability of gene g, using the model $Vg = h_g^2 \Theta + (1 - h_g^2)I$ to the observed normalization gene expression values e_{gs} by maximizing the likelihood $L(e_g|V_g) \propto \frac{1}{\sqrt{\det(V_g)}} \exp\left(-\frac{1}{2} e_g{}^T V_g^{-1} e_g\right)$, where $e_g = (e_{gs})$.

RESULTS

Variation of gene expression is higher between individuals than within families. After pre-processing and normalization, differential gene expression analysis was performed using the 24,016 probes. Between families, 473 genes were significantly differentially expressed (DE). The lower number of DE genes might be an effect of the small sample size (4-8 animals/sire). For each

DE gene, we calculated the variance among all 38 animals (i.e. total variance) and the variance within each family. As a measure of variability, we then calculated the ratio between the total variance and the variance within each family. For most genes, this ratio had a value greater than one, suggesting higher variation in gene expression among the population than within family. Figure 1 shows a scatter plot of variance in gene expression level among the population and between individuals from Family 11 for the 473 DE genes. As all the progeny were raised in the same places and in the same condition to minimize the environmental variation, the results suggest that a significant portion of the variation in gene expression is genetically determined and thus there exists a heritable component in gene expression.

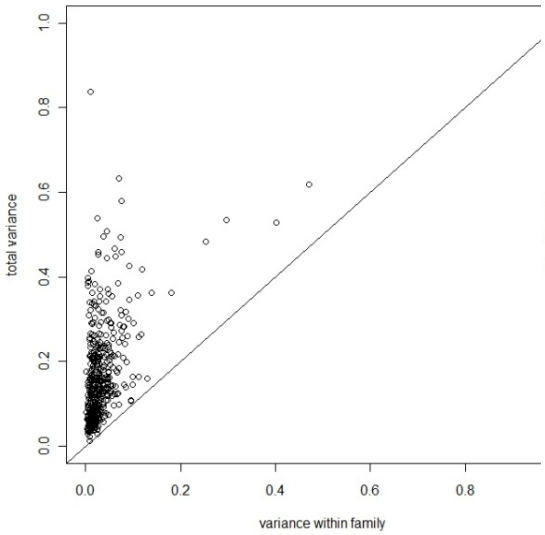


Figure 1. Scatter plot of total variance vs. variance within Family 11 for 473 differentially expressed genes.

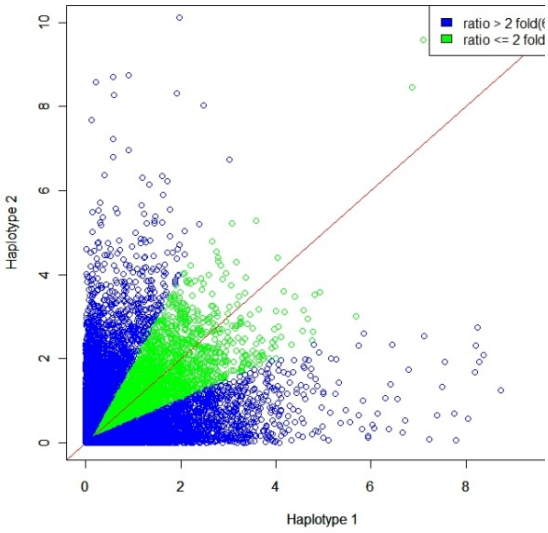


Figure 2. Scatter plot of gene expression level of haplotype 1 group vs. haplotype 2 group in Family 11.

Heritability of gene expression. For the analysis of the gene expression, the normalized intensity values for 22,246 probe sets (probe sets on the X chromosome were removed) were co-analysed along with SNP data from the 38 animals. Two animals were discarded that did not have SNP data. Using the genome-wide IBD and GRM, the overall heritability h_g^2 was estimated for each gene *g* using the variance-component method described in the Methods section. We then computed the overall heritability of gene expression h^2, by averaging all h_g^2 values. The result showed h^2 = 0.25 (standard error ±0.0023) when using the IBD matrix and h^2 = 0.24 (standard error ±0.0027) using the GRM matrix. Some negative values for h_g^2 were observed which do not have any biological interpretation and in most cases these values are very close to zero. These might be attributed to statistical noise. If we ignore negative values and assign each to zero, we obtained h^2 = 0.27 (standard error ±0.0021) and h^2 = 0.29 (standard error ±0.0024) when we used the IBD matrix and the GRM matrix respectively. Both estimates are consistent with previous results which reported that a significant portion of gene expression is heritable at the level of h^2 = 0.3 or higher (McRae *et al.* 2007; Price *et al.* 2011).

Gene expression varies between haplotype groups within families. For each gene within a family, the animals were grouped into two groups (i.e. haplotype 1, if the gene comes

from sire's parental strand and haplotype 2, if the gene comes from sire's maternal strand). Then the variance of gene expression within each haplotype group for each gene was calculated. As a measure of variability of gene expression between two haplotype groups, the variance ratio for each gene was calculated by dividing the variance of the expression levels from the haplotype 1 group by the variance of the expression levels from the haplotype 2 group. This revealed 65% (family 16) to 78% (family 7) of genes showed at least two-fold difference between the variances of the gene expressions in the haplotype 1 group and the haplotype 2 group. These percentages are much greater than expected from random ($P < 10^{-10}$) for every family tested. The results achieved suggested that there are differences in gene expression if the gene is coming from sire's parental or maternal side. Figure 2 shows a plot for the variance of gene expression level of the haplotype 1 group against the variance of gene expression level of haplotype 2 group for Family 11. This demonstrated that a significant number of values deviated from the straight line indicating equal variance for the two groups.

Family effect and haplotype effect on gene expression traits. We wanted to ascertain (1) if family and haplotype affect gene expression levels, and (2) if there is any variation in gene expression between the families. To test the hypothesis that there are family and haplotype effects on gene expression traits, a linear model was fitted in R (expression ~ family + haplotype + family * haplotype). Then, we conducted analysis of variance (ANOVA) test using this linear model. The result was a highly significant effect of family on the gene expression traits ($F = 18.6161$, $P < 2.2e-16$). Further, the effect of the interactions between family and haplotype were also highly significant ($F = 3.6527$, $P < 0.002$), although the haplotypes themselves did not have any significant impact on the gene expression traits.

ACKNOWLEDGEMENTS

This study was funded by SheepGenomics and the Australian Research Council project DP130100542.

REFERENCES

Amos, C. I. (1994) *Am J Hum Genet* **54**: 535.

Cheung, V. G., Conlin, L. K., Weber, T. M., Arcaro, M., Jen, K. Y., Morley, M., and Spielman, R. S. (2003) *Nature genetic.* **33(3)**: 422.

Daly, M. J., Rioux, J. D., Schaffner, S. F., Hudson, T. J., and Lander, E. S. (2001) *Nature Genetics.* **29**: 229.

Ferdosi, M. H., Kinghorn, B. P., Werf, J. H. v. d., and Gondro, C. (2013) *Bioinformatics, Under review*.

Kogelman, L. J., Byrne, K., Vuocolo, T., Watson-Haigh, N. S., Kadarmideen, H. N., Kijas, J. W., Tellam, R. L. (2011) *BMC genomic.* **12:** 607.

McRae, A. F., Matigian, N. A., Vadlamudi, L., Mulley, J. C., Mowry, B., Martin, N. G., Visscher, P. M. (2007) *Human molecular genetics.* **16(4)**: 364.

Nätt, D., Rubin, C.-J., Wright, D., Johnsson, M., Beltéky, J., Andersson, L., and Jensen, P. (2012) *BMC genomics.* **13(1)**: 59.

Price, A. L., Helgason, A., Thorleifsson, G., McCarroll, S. A., Kong, A., and Stefansson, K. (2011) *PLoS genetics.* **7(2)**: e1001317.

Schadt, E. E., Monks, S. A., Drake, T. A., Lusis, A. J., Che, N., Colinayo, V., *et al.* (2003) *Nature.* **422**: 297.

Spielman, R. S., Bastone, L. A., Burdick, J. T., Morley, M., Ewens, W. J., and Cheung, V. G. (2007) *Nature genetics.* **39(2)**: 226.

Stranger, B. E., Forrest, M. S., Dunning, M., Ingle, C. E., Beazley, C., Thorne, N., Dermitzakis, E. T. (2007) *Science.* **315(5813)**: 848.

VanRaden, P. M. (2007) *Interbull Annual Meeting Proceedings*. **37**:33.

Visscher, P. M., Hill, W. G., and Wray, N. R. (2008) *Nature reviews. Genetics.* **9(4)**: 255.

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:270-273

UTILITY OF GRAPHICS PROCESSING UNITS FOR DENSE MATRIX CALCULATIONS IN COMPUTING AND INVERTING GENOMIC RELATIONSHIP MATRICES

Karin Meyer and Bruce Tier

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

SUMMARY

The era of genomic evaluation has brought the need to perform computations involving large, dense matrices. Particular tasks are the computation and inversion of the genomic relationship matrix. This paper investigates the suitability of Graphics Processing Units together with highly optimised software libraries for these computations, using blocked algorithms. It is shown that calculations are readily sped up by parallel processing, using freely available library routines, and that reductions in time by factors of 4 to 5 are achievable even for 'consumer' grade graphics cards.

INTRODUCTION

Computer gaming requires computing of large numbers of pixel values at a fast rate. This computational load has stimulated development of 'co-processors' – so-called Graphics Processing Units (GPU). Modern GPU devices have thousands of cores and capabilities well suited to general purpose computing, providing very high rates of floating point operations. However, GPU cores have limited features and memory, restricting the type of computations that can be accelerated by GPUs. Typically, this requires computations to be executable in subsets and thus to be highly parallelisable.

Cole *et al.* (2012) discuss the potential of GPUs for applications in animal breeding. For a long time, efficient mixed model computations in animal breeding have relied on the sparseness of the pertaining equations. However, the advent of genomic evaluation has resulted in the need for large-scale manipulation of dense matrices. Fortunately, highly optimised software routines are available to perform many of the tasks required, especially in the BLAS (Dongarra *et al.* 1988) and LAPACK (Anderson *et al.* 1999) libraries. Their efficiency in computing the genomic relationship matrix (GRM) and its inverse has been demonstrated by Aguilar *et al.* (2011) and Meyer *et al.* (2013).

Use of a GPU requires a special programming interfaces such as CUDA (Compute Unified Device Architecture), the NVIDIA proprietary platform (NVIDIA Corporation 2013). Matrix computations on GPUs are greatly aided by corresponding software libraries: CUBLAS, part of the CUDA toolkit, provides GPU accelerated BLAS routines, and the equivalents to LAPACK routines are available from the CULA (Humphrey *et al.* 2010) or MAGMA (e.g. Dongarra *et al.* 2012) libraries. This allows for applications using such tools to be readily ported to GPUs, though challenges arise from their limited memory which requires matrices and computations to be broken into blocks accommodated on GPU devices. This paper presents a first investigation into the scope of GPUs to accelerate dense matrix computations, such as required to calculate and invert the GRM.

MATRIX MANIPULATION BY PARTS

Calculation of the GRM, $\mathbf{G}$, involves a matrix product of form $\alpha\mathbf{Z}\mathbf{Z}'$ with dimensions of $\mathbf{Z}$ equal to the number of individuals (n) $\times$ number of alleles (s) and α a scale factor (Van Raden 2008). As $\mathbf{G}$ is symmetric, only one triangle needs to be computed. Calculations represent a rank-k update of a symmetric matrix, a task performed by BLAS routine SYRK. Partition $\mathbf{G}$ and $\mathbf{Z}$ into blocks $\mathbf{G}_{ij}$ $(i, j = 1, r)$ and $\mathbf{Z}_{ik}$ $(i = 1, r$ and $k = 1, t)$, as dictated by memory available on the GPU. Blocks $\mathbf{G}_{ij} = \alpha \sum_{l=1}^{t} \mathbf{Z}_{il}\mathbf{Z}'_{jl}$ can then be computed by repeated calls to SYRK for $i = j$ and BLAS routine GEMM (which evaluates a general matrix by matrix product) for $i \neq j$.

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

Inversion. A standard method to invert a symmetric, positive definite matrix is to carry out a Cholesky decomposition, calculate the inverse of the factor and multiply the latter with its transpose, taking advantage of the triangular nature of these matrices. This can be performed by LAPACK routines POTRF and POTRI. For block-wise inversion, Gauss-Jordan elimination type algorithms have been suggested (Quintana *et al.* 2001; Ezzatti *et al.* 2011; Benner *et al.* 2011). This can be carried out 'in place', overwriting $\mathbf{G}$ with $\mathbf{G}^{-1}$. For each step, partition $\mathbf{G}$ into current (C), previous (P) and trailing (T) blocks with n_1, n_b (chosen block size) and n_2 rows, respectively.

$$\mathbf{G} = \begin{pmatrix} \mathbf{G}_{PP} & \mathbf{G}_{PC} & \mathbf{G}_{PT} \\ \mathbf{G}_{CP} & \mathbf{G}_{CC} & \mathbf{G}_{CT} \\ \mathbf{G}_{TP} & \mathbf{G}_{TC} & \mathbf{G}_{TT} \end{pmatrix}$$

$\mathbf{G}_{CC} \coloneqq \text{chol}(\mathbf{G}_{CC})$	POTRF
$\mathbf{G}_{CC} \coloneqq \mathbf{G}_{CC}^{-1}$	TRTRI
$\mathbf{G}_{PC} \coloneqq \mathbf{G}_{PC}\mathbf{G}_{CC}$	TRMM
$\mathbf{G}_{PP} \coloneqq \mathbf{G}_{PP} + \mathbf{G}_{PC}\mathbf{G}'_{PC}$	SYRK
$\mathbf{G}_{CT} \coloneqq \mathbf{G}'_{CC}\mathbf{G}_{CT}$	TRMM
$\mathbf{G}_{TT} \coloneqq \mathbf{G}_{TT} - \mathbf{G}'_{CT}\mathbf{G}_{CT}$	SYRK
$\mathbf{G}_{PT} \coloneqq \mathbf{G}_{PT} - \mathbf{G}_{PC}\mathbf{G}_{CT}$	GEMM
$\mathbf{G}_{CT} \coloneqq -(\mathbf{G}_{CC}\mathbf{G}_{CT})$	TRMM
$\mathbf{G}_{PC} \coloneqq \mathbf{G}_{PC}\mathbf{G}'_{CC}$	TRMM
$\mathbf{G}_{CC} \coloneqq \mathbf{G}_{CC}\mathbf{G}'_{CC}$	LAUUM

Figure 1. Algorithm for block-wise matrix inversion

At the beginning, current and trailing blocks contain the respective parts of $\mathbf{G}$ given $\mathbf{G}_{PP}$. The algorithm then starts with the Cholesky factorisation of the current, diagonal block, $\mathbf{G}_{CC} = \mathbf{R}'\mathbf{R}$, and inversion of $\mathbf{R}$, an upper triangular matrix. This is followed by steps adjusting previous blocks for the contribution of $\mathbf{G}_{CC}^{-1}$ to their inverses, and trailing blocks by 'absorbing' rows and columns $n_1 + 1$ to $n_1 + n_b$. Finally, $\mathbf{G}_{CC}^{-1}$ is obtained as $\mathbf{R}^{-1}(\mathbf{R}^{-1})'$. Blockwise calculations are repeated, updating n_1 to $n_1 + n_b$ and n_2 to $n2 - n_b$, until $n_1 = n$ and $n_2 = 0$. Pseudo-code adapted from Benner *et al.* (2011) (correcting errors in their description), together with the appropriate BLAS or LAPACK routines for individual calculations are given in Figure 1 (with $\mathbf{A} \coloneqq \mathbf{B}$ denoting replacement of $\mathbf{A}$ by $\mathbf{B}$).

MATERIAL AND METHODS

Time required for both types of matrix operations were compared using simulated matrices. For the matrix product, allele counts were obtained by sampling values 0, 1 or 2 from a uniform distribution for $s = 512000$ and $n = 512$ to 20,480 individuals. For matrix inversion, successive submatrices of a GRM set up as in Meyer *et al.* (2013), were used considering $n = 512$ to 16,384.

Calculations were performed in single precision, using either a single CPU (CPU1), all (4) CPU cores available (CPU4) or the GPU, performing computations in blocks as required by memory limits. For matrix multiplications on the CPU, $\mathbf{Z}$ was processed in up to 5 blocks, splitting $\mathbf{Z}$ adaptively into submatrices $\mathbf{Z}_{il}$ of size $n \times z$ with z chosen that $\mathbf{Z}_{il}$ did not exceed 10 Gb. Corresponding computations on the GPU used 250 blocks of size of $n \times 2048$ for $n \leq 12800$, and $n/2 \times 2048$ otherwise. For matrix inversion, use of LAPACK routines POTRF and POTRI for the complete matrix on both CPU and GPU was contrasted with the block algorithm described above on the GPU (GPUB). This used a block size ($\mathbf{G}_{CC}$) of $n_b = 2048$ and, as suggested by Benner *et al.* (2011), employed a hybrid algorithm with LAPACK routines (POTRF, TRTRI and LAUUM) executed on the CPU, using all 4 cores.

Computing environment. Calculations were carried out on a desktop computer running Linux, with CUDA 5.0. This was equipped with a quad-core Intel I7-960 processor rated at 3.2 Ghz with 8 Mb cache and 12 GB of RAM, and GPU capable NVIDIA GeForce GT240 graphics card with 96 cores, a clock speed of 1.46GHz and 1 Gb of memory. Programs were written in Fortran and compiled using gfortran (gcc 4.4.3), loading BLAS and LAPACK routines from the CUBLAS and CULA libraries and the Intel MKL 11.0 library for computations on the GPU and CPU, respectively.

RESULTS

Computing times required to form the product $\mathbf{ZZ}'$, shown on a logarithmic scale, are constrasted in Figure 2. With $sn(n + 1)$ floating point operations per product, half multiplications and

half additions, these increase quadratically with the number of individuals. As previously shown by Aguilar *et al.* (2011) (though they utilised BLAS routine GEMM which does not exploit the symmetry of **G** and thus requires $2sn^2$ operations), results demonstrate that calculations involved are highly suited to parallel processing. Using all 4 CPU processors available decreased the computing time on average by a factor of 3.69. Employing the GPU reduced times further for all cases, even if n was too large to carry out computations for all n through one call to routine SYRK with the memory available on the GPU device, yielding an average speed-up of 5.16 times. Additional investigations using other values for s (not shown) yielded comparable patterns, suggesting that results are scalable and that similar improvements can be achieved for larger problems.

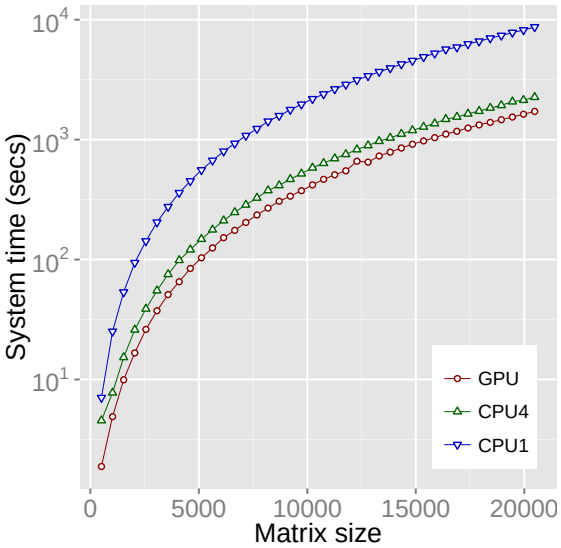


Figure 2. Times for matrix product

Corresponding results for matrix inversion are presented in Figure 3. For this case, parallelisation was slightly less successful with computations using a single CPU for $n > 4000$ requiring on average 3.33 times as long as those utilizing all 4 cores available. For small matrices, processing on the GPU required longer than CPU1. For $n > 4000$, single block computations on the GPU performed best, reducing computing times by factors of 4.14 and 1.23 compared to CPU1 and CPU4, respectively. However, memory available on the GPU restricted these to $n < 15000$. Blockwise inversion on the GPU required similar times than using all CPU cores available in parallel. Other sizes of n_b were tried (not shown), but offered little advantage – indeed for small block sizes, times exceeded those for CPU1. Benner *et al.* (2011) reported greatly increased speeds of computation for their algorithm compared to LAPACK routines, both for parallel computations on the CPU and a hybrid approach, while calculations on the GPU only required matrix sizes of more than 7,000 to be advantageous. Nevertheless, none of their findings could be repeated with our hardware set-up.

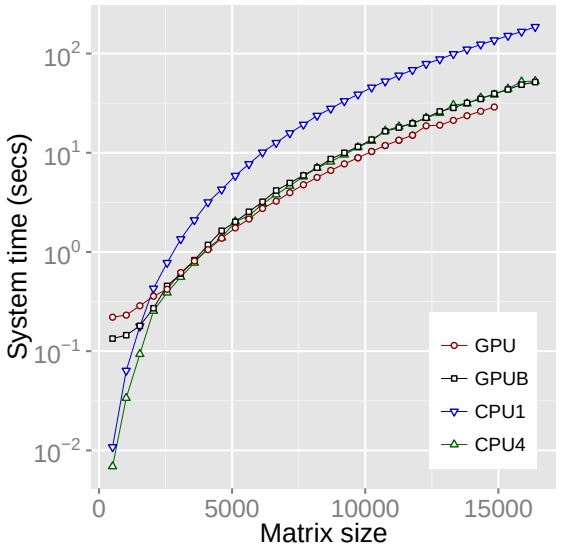


Figure 3. Times for matrix inversion

DISCUSSION

Dense matrix calculations are computationally demanding and the efficiency of computations is greatly influenced by the organisation of loops and memory access. Highly optimised linear

algebra routines are available which perform common types of operations and, together with modern compilers and libraries tuned for specific hardware, can yield very fast computations. These routines are freely available and easy to use and, where possible, should be used when programming such applications.

Moreover, corresponding libraries are available to readily utilise multiple (CPU) or very many (GPU) threads. As shown, performing computations in parallel can markedly speed up calculation of the GRM and its inversion. While the advantages of using the GPU over all CPU cores available shown here might appear modest, it should be born in mind that the graphics card utilised only had very basic GPU capabilities. Hence, results should be regarded more as a 'proof of principle' rather than being indicative. Modern GPUs targeting general purpose computing have up to 6 Gb memory and thousands of cores, and are capable of performing double precision calculations with huge numbers of floating point operations per second, effectively turning a standard desktop computer into a personal supercomputer. Future work will repeat the calculations shown with more powerful hardware, and is likely to achieve substantially higher reductions in computing time for calculations that can be accelerated using the GPU.

CONCLUSIONS

Utilisation of multiple threads can dramatically reduce computing times for dense matrix calculations, such as required in the context of genomic evaluation. Graphic Processing Units provide powerful hardware for parallelisation of computations, and are likely to see increasing use in animal breeding applications in the future.

ACKNOWLEDGEMENTS

This work was supported by Meat and Livestock Australia under grant B.BFG.0050.

REFERENCES

Aguilar I., Misztal I., Legarra A. and Tsuruta S. (2011) *J. Anim. Breed. Genet.* **128**:422.

Anderson E., Bai Z., Bischof C., Blackford S., Demmel J., Dongarra J., Du Croz J., Greenbaum A., Hammarling S., McKenney A. and Sorensen D. (1999) *LAPACK Users' Guide*. Society for Industrial and Applied Mathematics, Philadelphia, PA, Third edition.

Benner P., Ezzatti P., Quintana-Ortí E.S. and Remón A. (2011) In *International Conference on High Performance Computing and Simulation (HPCS)*. IEEE, pp. 640–646.

Cole J.B., Newman S., Foertter F., Aguilar I. and Coffey M. (2012) *J. Anim. Sci.* **90**:723.

Dongarra J., Dong T., Gates M., Haidar A., Tomov S. and Yamazaki I. (2012) In *SC12*. Salt Lake City, Utah, November 14, 2012.

Dongarra J.J., Croz J.D., Hammarling S. and Hanson R.J. (1988) *ACM Trans. Math. Softw.* **14**:1.

Ezzatti P., Quintana-Ortí E.S. and Remón A. (2011) *J. Supercomput.* **58**:429.

Humphrey J.R., Price D.K., Spagnoli K.E., Paolini A.L. and Kelmelis E.J. (2010) In *Modeling and Simulation for Defense Systems and Applications V*, E.J. Kelmelis, ed., Proc. SPIE 770502.

Meyer K., Tier B. and Graser H.U. (2013) *J. Anim. Sci.* **00**:000.

NVIDIA Corporation (2013) CUDA parallel programming platform. `http://www.nvidia.com/object/cuda_home_new.html`. Accessed: February 1, 2013.

Quintana E.S., Quintana G., Sun X. and van de Geijn R. (2001) *SIAM J. Sci. Comput.* **22**:1762.

Van Raden P.M. (2008) *J. Dairy Sci.* **91**:4414.

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:274-277

BREEDS OF NEW ZEALAND SHEEP – AS RECORDED OR BY GENOMIC PREDICTION

K.G. Dodds, S.-A.N. Newman, B. Auvray and J.C. McEwan

AgResearch, Invermay Agricultural Centre, Mosgiel, New Zealand

SUMMARY

The breed composition of the New Zealand sheep industry was examined to help understand the nature of this industry and observe recent trends. The maternal breeds of the New Zealand sheep meat industry are predominantly Romney with Coopworth, Perendale and Texel also common. The last 15 years has seen increased Perendale and Texel and decreased Coopworth numbers and a trend towards composites in ram breeding flocks. A genomic prediction method (gBLUP) was used to predict breed composition. Predicted breed composition was found to be similar to recorded breed for animals with similar breed composition to those in the training set used, and therefore is a useful method of verifying breed recording or predicting breed in unrecorded animals. Genomic prediction tended to over-predict breed components for animals of breed types not included in the training set.

INTRODUCTION

The breed composition of a population reveals the nature of that industry, and can inform research and policy decisions. The composition of the New Zealand (NZ) sheep meat industry is examined, primarily using the Sheep Improvement Limited (SIL; sil.co.nz) database. As more breeding stock are required to resource the national ewe flock than for terminal sires, these results mainly reflect the maternal breed composition.

Genotyping platforms that assay thousands of single nucleotide (SNP) markers have recently been used to predict breed composition (Sölkner *et al*. 2010, Kuehn *et al.* 2011; VanRaden *et al*. 2011; Frkonja *et al*. 2012). One such method was examined in NZ sheep populations, as it offers a method of breed designation without relying on animal recording, and may therefore be useful for validating sample origin and breed recording. It may also allow prediction of breed composition in unrecorded animals.

MATERIALS AND METHODS

Recorded breed. A file of all NZ sheep that have individual records was obtained from SIL's database. The information obtained included year of birth (BYR; 1960-2012) and breed proportions. Up to five different contributing breeds are recorded on SIL for each animal. These are determined by (preferentially): averaging the recorded breeds of the parents, direct recording by owner or by substituting the 'flock breed' for the breed of any unknown parent. The averaging process rounds values up to the nearest 0.5%.

Genomic prediction of breed. The OvineSNP50 SNP genotypes of 13,118 animals that had been genotyped by AgResearch, predominantly as part of an Ovita-funded research programme, up to August 2011 were obtained. Of these 8,705 were recorded on SIL at the time of analysis.

Genomic prediction of breed. Genomic selection (GS) methods were applied to the recorded breed proportions to develop predictions of breed proportions (VanRaden *et al*. 2011). Animals born prior to 2008, and with a recorded breed composition having more than 50% of either Romney (n=2849), Coopworth (n=1007), Perendale (n=290) or Texel (n=168), or more than 50% of Romney, Coopworth or Perendale combined (hereafter denoted "CompRCP"; n=103) were chosen for training. Prediction equations for each breed were calculated using the gBLUP method (Goddard *et al*., 2010) using the model $y_i \sim \mu + u_i + e_i$, where y is the recorded proportion of

Romney, Coopworth, Perendale or Texel), μ is a constant, u_i is the modelled breed proportion and e_i is the residual, $Var(u)=G\sigma^2_a$, $Var(e)=I\sigma^2_e$ and G is the genomic relationship matrix calculated using the first method described by VanRaden (2008). The heritabilities ($\sigma^2_a/(\sigma^2_a+\sigma^2_e)$) of these 'traits' were fixed at 0.95. For animals not in the training set, predicted breed proportions were calculated directly from the animal's SNP data (VanRaden 2008). Principal components were calculated with the prcomp function of R (R Core Team, 2013) using the genomic relationship matrix (as described above) of all 13,118 genotyped animals as a similarity matrix. This was used to graphically display the results.

RESULTS AND DISCUSSION

Recorded breeds. A summary of NZ sheep breeds is shown in Table 1. SIL uses a set of breed definitions as required by its clients, and a number of these represent recent composites, but they are treated as additional 'pure' breeds here. SIL is underrepresented in fine-medium wool breeds (Merino, Corriedale and Halfbred), which use alternative genetic evaluation systems. Trends in recent years (Figure 1) are for stable numbers of Romney, decreasing Coopworth, increased Perendales and Texels (but levelled off), increased Poll Dorset and Suffolks (but now decreasing). The recorded numbers (full animal equivalents) of 'Composite' increased sharply from 2000 to 2005 and then levelled off. In recent years, less than 30% of animals with Coopworth were pure Coopworths. Corresponding figures for Romney, Perendale and Texel were around 70%, 60% and 5%, respectively.

Table 1. Breeds of sheep recorded on SIL from 2005 and estimates in the NZ population.

Breed	% (of purebreeds)	% in SIL[#]	% in genotyped subset	% in NZ*
Romney	53	44	52	58
Coopworth	7	9	23	11
Merino	<1	<1	0	8
Perendale	7	9	7	7
Corriedale	2	2	<1	6
Halfbred	NA	NA	0	4
Drysdale	<0.1	<0.1	0	1
Borderdale	<1	<1	<0.1	1
Texel	2	8	7	1

* http://www.rarebreeds.co.nz/sheepnumbers.html

including part contributions; Other breeds exceeding 2%: Poll Dorset (5%), Suffolk (4%), Composite (4%).

Breed Prediction. Graphical representations of the training set and the remaining animals are shown in Figures 2 and 3 respectively, using the first two principal components (PCs). These explained 66% of the variation (proportion of eigenvalues) in the relationship matrix. The predictions are not always between zero and one, but were seldom more than 0.1 from this range. When the prediction equations were applied to a subset of the SIL recorded animals (whose genotypes were available at that time, and whose SIL breed matched one of the training set breed types), the regression of predicted on recorded breed had correlations ranging from 0.92 (for proportion of Texel) to 0.98 (for proportion of Romney). The slopes of the regressions ranged from 0.96 (Coopworth) to 1.06 (Texel). These results suggest that the genomic selection method is predicting the recorded breed closely, at least within this set of breed types.

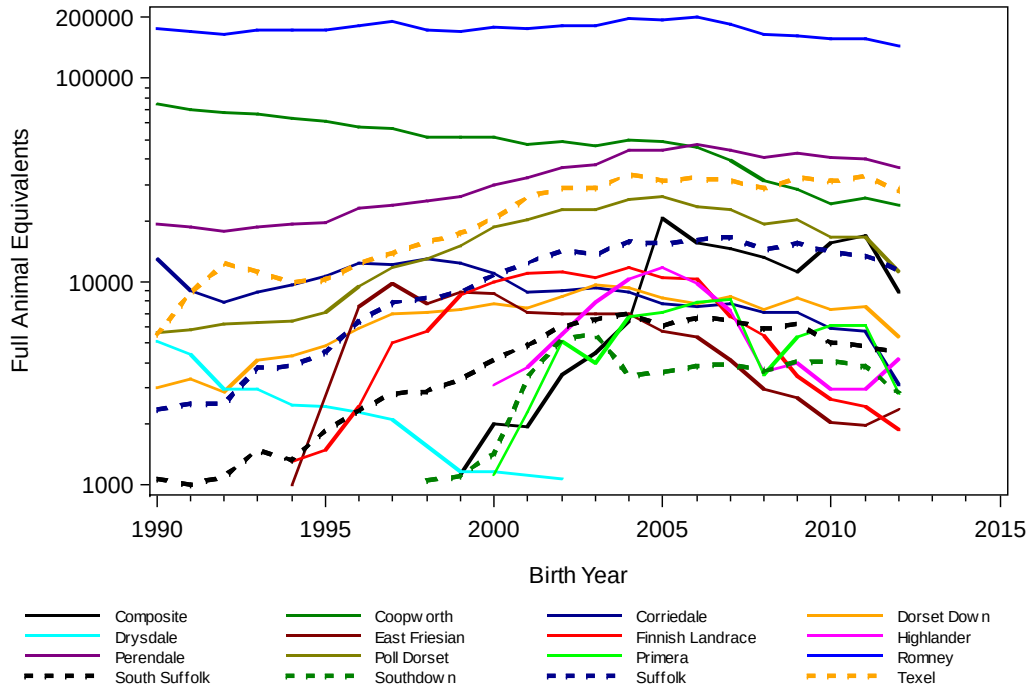


Figure 1. Full animal equivalents of each breed recorded on SIL for each birth year.

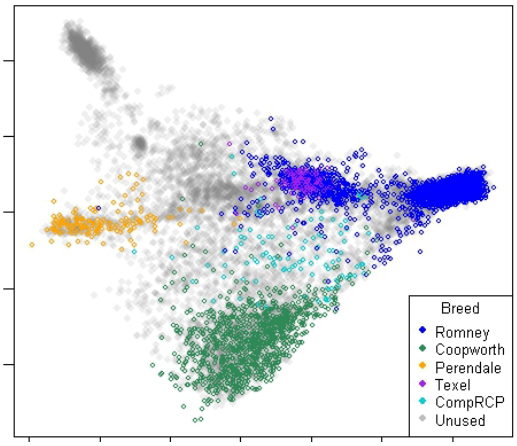


Figure 2. Plot of the first two principal components (PC2 v PC1) of the training set (coloured) for the genomic selection method.

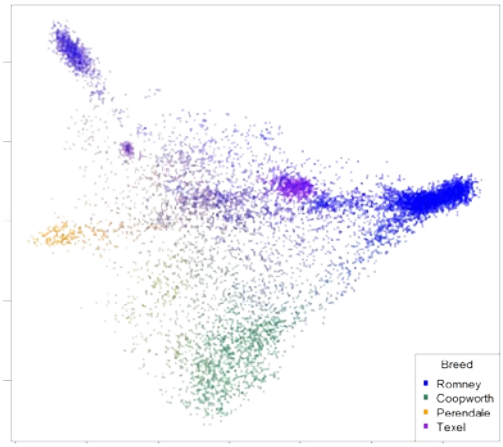


Figure 3. Predicted breed composition of 8776 animals (PC2 v PC1) not in the training set. Points increase in transparency as predicted breed proportions decrease.

The graphs show that breeds cluster together. Coopworths are more spread out than the others breeds, possibly reflecting that a relatively smaller percentage of Coopworth animals are recorded as pure. It may also reflect that the Coopworth breed has allowed some introgression of other breeds, and therefore they are likely to be more diverse than a closed breed. The Romney breed is

mainly located in one region, but some extend across the position of the Perendales (in PC2-PC1 space). This substructure may be a result of differing breeding priorities within this breed.

Animals in the upper left quadrants (Figure 3) are predicted to have a low proportion of each breed being predicted. The animals in this region tend to be meat breeds or composites (mainly Primera, Poll Dorset, Suffolk, and Wiltshire; see Table 2). Therefore the predicted breed proportions are likely to be overestimates. This suggests that the prediction method does not work so well in regions that did not contain any training set animals. It remains to be seen if these predictions would drop if some of these animals were included in training. An intriguing result is the estimated proportions in the four Cheviots genotyped, being about 140% Perendale and -40% Romney, which points to the Romney x Cheviot origins of Perendales.

Table 2. Mean predicted proportions, for Romney (pRom), Coopworth (pCoop), Perendale (pPere) and Texel (pTex) in animals that are recorded as purebred and that were not used for training. Results are shown for Cheviots and breeds with at least 10 animals genotyped.

Breed	n	pRom	pCoop	pPere	pTex
Romney	1496	**0.985**	0.007	0.003	0.002
Coopworth	286	0.022	**0.937**	0.011	0.021
Perendale	262	0.036	0.017	**0.933**	0.009
Texel	57	0.025	0.041	0.037	**0.869**
Corriedale	42	0.084	0.403	0.212	0.145
Poll Dorset	39	0.333	0.090	0.099	0.054
Suffolk	25	0.241	0.127	0.318	0.140
Finnish Landrace	12	0.123	0.167	0.217	0.134
Marshall Romney	10	0.709	0.110	0.116	0.015
Cheviot	4	-0.444	-0.033	1.389	0.046

CONCLUSIONS

NZ meat sheep are predominantly Romney with recent increases in Perendale and Texel and decreases in Coopworth. There is also a trend towards composites. Genomic methods offer a novel method for predicting breed or breed composition, without animal recording. There is a need for additional samples from the base breeds not currently sampled.

ACKNOWLEDGMENTS

We are grateful for support from Ovita Ltd, Beef + Lamb NZ and the NZ sheep breeders.

REFERENCES

Goddard M. E., Hayes B. J. and Meuwissen T. H. E. 2010. *Genet. Res.* **92**: 413.

Frkonja A., Gredler B., Schnyder U., Curik I. and Sölkner J. (2012) *Animal Genetics* **43**:696.

Kuehn L.A., Keele J.W., Bennett G.L., McDaneld T.G., Smith T.P.L., Snelling W.M., Sonstegard T.S. and Thallman R.M. (2011) *Journal of Animal Science* **89**:1742.

R Core Team (2013) http://www.R-project.org.

Sölkner J., Frkonja A., Raadsma H.W.R., Jonas E., Thaller G., Egger-Danner E. and Gredler B. (2010). Interbull Bulletin 42: 62.

VanRaden P.M. (2008) *J. Dairy Sci.* **91**:4414.

VanRaden P.M., Olson K.M., Wiggans G.R., Cole J.B. and Tooker M.E. (2011) *Journal of Dairy Science* **94**:5673.

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:278-281

SUCCESS RATES OF COMMERCIAL SNP BASED PARENTAGE ASSIGNMENT IN SHEEP

A. M. Bell[1,2], J.M. Henshall[1,2], S. Gill[2,3], K. Gore[2,4] and J.W. Kijas[5]

[1]CSIRO Animal, Health and Food Sciences, FD McMaster Laboratory, Armidale NSW 2350
[2]Sheep Cooperative Research Centre, Armidale NSW 2351
[3]Meat Livestock Australia, Sydney NSW 2059
[4]Animal Genetics and Breeding Unit, University of New England, Armidale NSW 2351
[5]CSIRO Animal, Health and Food Sciences, Queensland Bioscience Precinct, Brisbane QLD 4067

SUMMARY

Single nucleotide polymorphism (SNP) based parentage assignment is attractive as SNP are abundant in the sheep genome and amenable to high throughput and therefore lower cost genotyping. To examine the minimum number of SNP required to obtain high accuracy parentage assignment, blood samples were collected from 4 industry flocks and genotyping was undertaken. A maximum likelihood approach was applied to the genotypes to predict sire, dam and progeny within 3 of the sampled flocks, and dams within 1 sampled flock. A SNP based, flock specific methodology utilizing differing numbers and types of SNPs for estimating assignment rates was developed. Rates of assignment ranged from 99.5% to 77.7% across 3 flocks, with 0% incorrect assignments, with the exception of one panel in one flock for sire assignment, where the incorrect assignment rate was 0.1%. Rates of assignment varied from 62.2% to 28.3% with 0% incorrect assignments in the fourth flock, with the exception of one panel for dam assignment where the incorrect assignment rate was 0.1%, but only 60% of dams and 50% of sires within this flock were genotyped. Using 2 out of a potential 6 multiplexed panels of SNP markers gave high rates of correct paternity, but using 3 panels provided higher confidence and is recommended. This maximum likelihood approach using SNPs provides the basis for delivering highly accurate parentage determination for under AUD20, increasing the affordability of this as a powerful tool for industry.

INTRODUCTION

When parentage in a breeding enterprise is known, the rate of genetic progress can be improved by information from relatives when estimating breeding values by best linear unbiased prediction (BLUP EBVs) and accounting for maternal effects. Generating parentage records can be laborious and expensive due to the large amount of infrastructure required for artificial insemination programs, single sire mating strategies and mothering up or pedigree recording at lambing events. Inadvertent misallocation of lambs to dams can also occur particularly if dams are not scanned in lamb to obtain knowledge of the number of lambs expected and cross-mothering or mismothering occurs. DNA based methods of predicting parentage have been the focus of research in recent years, and the utilization of SNPs has reduced the cost of genotyping. Successful parentage testing requires a robust and technically accurate SNP genotyping platform coupled with a marker set containing SNP with high minor allele frequencies (MAF). The objective of this study was to develop an industry applicable low cost DNA based tool utilising SNP for determining sheep parentage.

MATERIALS AND METHODS

The SNP markers used in this project were identified by the International Sheep Genomics Consortium (ISGC). SNP were prioritised for use in parentage testing following analysis within a

spectrum of breeds using multiple genotyping platforms (Kijas *et al*. 2009; Kijas *et al*. 2012). In this experiment, 3 types of SNP were used to design sets or "multiplexes" of SNPs. A multiplex is a combination of SNP that can be assayed in a single reaction. The 3 SNP types were 1) ISGC parentage SNPs (those identified by the ISGC as suitable for parentage testing); 2) performance SNP (SNP identified as directly causing variation in phenotype or linked to mutations that cause phenotypic variation, for example, SNP linked to the Poll locus); and 3) filler SNP (SNP used to fill in around the first two types of SNP with a high MAF across a broad spectrum of breeds). A total of 383 SNPs were assigned to 6 multiplexed panels (named W1-W6). The number of markers in each multiplex ranged from 63 (W2, W3 and W4) to a maximum of 66 (W5). Details of each multiplex are shown in Table 1.

Table 1. SNP types within multiplexes

Panel	Total SNP	SNP Type		
		ISGC	Performance	Filler (MAF)
W1	64	38	6	20
W2	63	28	3	32
W3	63	18	1	44
W4	63	2	0	61
W5	66	0	0	66
W6	64	1	0	63
Total	383	87	10	286

Sheep Genetics (www.sheepgenetics.org.au) identified 4 industry representative flocks with varying levels of genetic relatedness between candidate sires, dams and progeny, and where knowledge of parentage was essential for breeding program purposes. Flocks 1, 3 and 4 are Merino enterprises, and Flock 2 comprises Dohne sheep. Blood cards were distributed by Sheep Genetics to the targeted producers. The blood cards were returned to Sheep Genetics, and sent to GeneSeek (USA) for DNA extraction and SNP genotyping using the SEQUENOM platform. Details of the flocks and number of sheep within each group in each flock are given in Table 2.

Table 2. Number of genotypes per flock and sheep group – details supplied by each flock owner

Flock	Sires		Dams		Lambs		Unknown	
	Total	Genotyped	Total	Genotyped	Total	Genotyped	Total	Genotyped
1	11	11	302	302	415	415	0	0
2	0	0	111	111	122	118	0	0
3	32	32	111	111	103	103	44	44
4	7	7	21	21	180	180	3	3

SNP data was only used if the sheep was known to be a sire, dam or lamb. Analysis of the genotype data was conducted without knowledge of the relationships between sires or the long term level of inbreeding. A maximum likelihood method was used for pedigree assignment (Marshall *et al.* 1998; Kalinowski *et al.* 2007; Kalinowski *et al.* 2010). Given the SNP data for a sire and a lamb, the likelihood that the sire is the parent is evaluated, along with the likelihood that the sire is not the parent of the lamb. The estimations use an assumed allele frequency for each SNP in the population, and an assumed genotyping error rate. As in Marshall *et al.* (1998) the logarithm (log) of the ratio (likelihood that the sire is the parent / likelihood that the sire is not the

parent) is referred to as the LOD score (from log odds). LOD scores were also estimated for lamb-dam pairs, and for lamb-sire-dam trios.

Simulation was used to derive an appropriate LOD threshold for each test type (lamb-sire pair, lamb-dam pair, or lamb-sire-dam trio) for each flock based on the observed allele frequencies. A total of 1000 progeny were simulated, each with randomly chosen sire and dam from the flock. Missing parents were simulated using allele frequencies estimated for the flock. For each simulated lamb, LOD scores were estimated for each sire and dam. For the most likely 5 sires and most likely 5 dams the LOD score was estimated for each of the 25 possible parent pairs. For sire parentage, the most likely sire was identified and the LOD score stored (mLOD1), along with the difference between mLOD1 and the LOD score for the next most likely sire. This difference was referred to as Δ1. The LOD score for the second most likely sire (mLOD2) and associated Δ2 were stored. The same method was used for dam parentage and for sire-dam parentage.

For Δ, a threshold (TΔ) was declared at TΔ = 3, and was used in all flocks. Parentage was only assigned if the most likely parent was at least 3 times more likely than the second most likely parent. Given the threshold TΔ = 3, a threshold for mLOD, (TmLOD) was found that balanced the number of false positives (i.e. mLOD2 > TmLOD) and false negatives (i.e. mLOD1 < TmLOD), subject to the constraint that the percentage of false positives was less than 10%.

For the real lambs, mLOD and Δ were compared to the thresholds TmLOD and TΔ, and parentage assigned if mLOD ≥ TmLOD and Δ ≥ TΔ, or not assigned if mLOD < TmLOD and Δ < TΔ. In all simulations and analyses we assumed a genotyping error rate of 1%.

RESULTS

Table 3. Assignment rates (AR %) for real data, False negative (-ve %), False positive (+ve %) and TmLOD (simulated data) using varying number of SNP and panels - 127 (W12), 190 (W123) or 191 (W126)

Panel	W12				W123				W126			
	AR	Tm LOD	-ve	+ve	AR	Tm LOD	-ve	+ve	AR	Tm LOD	-ve	+ve
Flock 1 Sire	97.3	5.2	1.7	1.7	98.3	9.4	0.7	0.8	99.5	7.1	0.0	0.4
Flock 1 Dam	88.2	7.7	3.7	1.6	94.7	10.7	0.7	0.7	95.4	10.2	0.8	0.9
Flock 1 Trio	96.4	22.8	0.8	1.2	97.8	33.7	0.1	0.2	98.6	33.0	0.2	0.5
Flock 2 Dam	81.4	5.9	5.4	5.0	81.4	9.7	0.9	2.1	90.7	9.1	2.5	2.2
Flock 3 Sire	80.6	4.6	0.9	0.8	86.4	4.7	0.0	0.0	79.6	9.8	0.0	0.2
Flock 3 Dam	81.6	6.1	1.2	0.8	91.3	6.7	0.1	0.1	88.3	8.0	0.1	0.2
Flock 3 Trio	77.7	19.6	0.1	0.5	78.6	33.3	0.0	0.0	78.6	30.1	0.0	0.1
Flock 4 Sire	49.4	2.1	1.1	1.2	48.3	4.3	0.8	0.5	48.3	30.3	0.5	0.2
Flock 4 Dam	60.0	4.3	1.7	1.3	62.2	4.6	0.2	0.2	60.0	7.8	0.4	0.9
Flock 4 Trio	28.3	18.4	0.5	0.4	28.9	32.8	0.2	0.2	28.9	30.3	0.1	0.1

False negative and false positive rates decreased (false positives decreased from 1.7% to 0.8% in Flock 1 sires simulated data) when changing from W12 to W123. This trend was evident across most flocks and groups (sire, dam or trio). In the real data, generally the assignment rate increased as the number of panels changed from 2 to 3. The exception is Flock 4, for which assignment rates were lower in some of the groups. There was not a significant difference in assignment rate across the panels examined (p=0.95). Upon investigation, it was discovered that 40% of dams and 50% of sires of Flock 4 had not been genotyped for this particular study. The number of genotypes for dams in this flock was the smallest across all the flocks at 21, but the use of embryo transfer in this flock has allowed the dams to have a large number of progeny (range of 0-15, average of 5).

DISCUSSION

This study demonstrated that a small number of SNP panels (2-3) generate adequate parentage assignment rates in Australian sheep flocks. The results also indicate that the design of the SNP panels is technically robust. Their performance across 4 flocks and 2 breeds in this study showed high rates of assignment where genotypes of potential sires, dams and lambs were available. Assignment rates were lower when fewer sires and dams within that flock were genotyped. The parentage assignment methodology developed allows for the assumption of a genotyping error rate, and it can be set to account for the observed error rate in any given SNP genotyping platform. This prevents the true sire from being eliminated on the basis of a single genotyping error. Importantly, the approach also uses the allele frequency at each SNP within the flock to generate population specific thresholds. This ensures that the thresholds for assignment are specific for each breed and flock.

The availability of parentage SNP panels with inbuilt performance SNPs is attractive to industry. The benefits for producers can include obtaining additional information such as the Poll status of the animal at no additional cost. The candidate SNPs that comprise the performance SNPs in parentage panels will be an area of focus for future research.

Utilising 2 panels of multiplexed SNP (or 127 loci) gave high rates of correct parentage and may be sufficient for many flocks, but 3 panels (or 190 SNP) provided higher confidence and is the recommendation for initial commercial application of a DNA based parentage testing product for less than AUD20.

ACKNOWLEDGEMENTS

The authors wish to thanks Felice Driver and MLA for financial support. Darryl Irwin (Sequenom, Queensland) performed multiplex assay design and testing. Julius van der Werf contributed to the methods development and assisted in the transfer of the working SNP panels to Daniel Pomp (GeneSeek) for production genotyping.

REFERENCES

Kalinowski S.T., Taper M.L. and Marshall T.C. (2007) *Molecular Ecology* **16**(5):1099.
Kalinowski S.T., Taper M.L. and Marshall T.C. (2010) *Molecular Ecology* **19**(7):1512.
Kijas J.W., Townley D., Dalrymple B.P., Heaton M.P., Maddox J.F., McGrath A., Wilson P., Ingersoll R.G., McCulloch R., McWilliam S., Tang D., McEwan J., Cockett N, Oddy V. H., Nicholas F.W. and Raadsma H. for the International Sheep Genomics Consortium. (2009) *PLoS ONE* **4:** e4668.
Kijas J.W., Lenstra J., Hayes B., Boitard S., Porto Neto L., San Cristobal M., Servin B., McCulloch R., Whan V., Gietzen K., Paiva S., Barendse W., Ciani E., Raadsma H., McEwan J., Dalrymple B., and the International Sheep Genomics Consortium (2012) *PLoS Biology* **10**(2):e1001258.
Marshall T.C., Slate J., Kruuk L.E. and Pemberton J.M. (1998) *Molecular Ecology* **7**(5):639.

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:282-285

NOVEL PHENOTYPING TECHNIQUES FOR ENHANCING GENETIC AND GENOMIC PREDICTIONS OF TRAITS THAT ARE DIFFICULT TO MEASURE IN GRAZING LIVESTOCK

I.W. Purvis[1], P. Valencia[2], L. Overs[2] and P.L. Greenwood[1,3]

[1]CSIRO Animal, Food and Health Sciences, Armidale NSW 2350
[2]CSIRO Information and Communication Technologies, Pullenvale QLD 4069
[3]NSW Department of Primary Industries, University of New England, Armidale NSW 2351

SUMMARY

The development of miniaturised wireless sensors and data capture systems now offers the capability to study livestock in their commercial production environment, and to do this in a way that does not constrain the animal from expressing its full range of genetic drivers for the traits under study. In this way, variation in the traits of economic importance which form the breeding objective, can be directly assessed. This will allow appropriate genetic parameters to be estimated for novel and hard to measure traits.

This paper discusses the issues underlying the need for new and novel phenotyping methods and presents early results from studies utilising sensors and sensor networks for predicting feed intake and feed efficiency of individual cattle on pasture.

INTRODUCTION

With recent advances in high-throughput genotyping technologies for livestock, the rate-limiting step in the conduct of large-scale genetic investigations has become the collection of complex phenotype information from relevant populations (Pollak *et al*. 2012). This is particularly so for grazing ruminants. To date, variation in some of the difficult-to-measure attributes of grazing livestock that have significant economic impacts, such as feed intake on pasture, aspects of reproductive performance and quantification of disease status, have only been quantifiable by constraining animals in artificial (non-grazing) environments. For example, the measurement of feed intake in beef cattle is now usually conducted in a feedlot environment where animals are maintained in group pens and fed diets of very different composition and availability than the pasture swards that often constitute the normal production system in Australia. Similarly, in assessing disease traits, the phenotype is not always measured in the environment where the proximal causes of the disease state are found (Houle *et al*. 2010).

The large international research effort that has delivered a high quality map of the bovine genome has been accompanied by a similar effort in phenotyping. However, some of the traits of major economic importance to beef cattle breeders have not been able to be measured, in part, due to the lack of practical measurement technologies and a focus on early age selection criteria. Whilst indirect or proxy traits have been utilised to acquire some knowledge of the associations between genotype and breeding objective traits, there remains a significant phenotype gap that needs to be filled to improve the return on the investment in genotyping. The capacity to directly measure traits of importance in breeding objectives relevant to pasture-based beef enterprises is critical. This will give breeders the capacity to identify the selection strategies that lead to the most cost effective means of achieving optimal and sustainable progress in the aggregate outcome.

NOVEL METHODS OF PHENOTYPING GRAZING LIVESTOCK FOR ECONOMICALLY IMPORTANT TRAITS

The development of electronic sensing capability has the potential to allow the measurement of traits of economic importance that previously had not been measurable in the commercial environment. Historically, on-animal logging devices for sensors used for phenotyping were bulky, and often heavy enough to raise concerns that the animal may not have exhibited its normal behaviours. However, over recent decades there have been considerable advances in miniaturisation and reduction of power use in electronic devices, such as microcontrollers, Global Positioning System (GPS) chips and in radio technologies. This has allowed ecologists and environmental scientists to collect high quality traces of the movements of free-ranging animal over, often lengthy, time-frames.

Recent technical advances in digital radio communications and microcontrollers has led to the evolution of Wireless Sensor Networks (WSN) which offers the potential for lightweight, small sensing devices for measuring a wider variety of traits relevant to grazing livestock (Hancock *et al*. 2009). However, the constraints imposed by a device that can be practically deployed on livestock, introduce limitations on local storage and communications throughput, which in turn, makes transmission of high-temporal, low-level sensory information difficult, particularly as the system is scaled up to a larger number of devices. This limitation has motivated the development of classifiers on the WSN nodes, which change the high temporal-resolution, but low-level sensed-data, into temporally-sparse high-level behavioural activities. Such an approach can produce a significant reduction of information whilst retaining enough information to still accurately classify phenotype behaviours. This reduction of information saves bandwidth and energy, which positions this approach to enable measurement of large numbers of livestock over long-term periods.

The intersection of the capability to have accurate knowledge of phenotype behaviour in their natural environments, over long periods on large numbers of animals, provides the novel methodology for phenotyping livestock in a practical and economically viable way.

AN EXAMPLE: DIRECT PHENOTYPING OF FEED INTAKE AND EFFICIENCY OF GRAZING RUMINANTS USING WIRELESS SENSORS AND SENSOR NETWORKS

In Australia the major cost of beef cattle production is associated with the cow-calf unit. It has been estimated that the feed costs of the breeding female and her calf can be 60-70% of the total herd feed costs, and as much as 90%, when account is taken of rearing replacement females. As such, this is a critical component of the input costs of the beef enterprise, and genetic variation in this trait and its association with production efficiency should be key elements of the breeding objectives of breeders of beef bulls.

To date, the focus for measuring variation in feed intake has been evaluation of young animals in a feedlot environment where test animals are maintained in group pens and fed *ad lib*. diets of grain-based high energy concentrates. Alternatively, under pasture systems, chemical markers such as N-alkanes (Dove and Mayes 2006), have been used to predict intake, selectivity, and digestibility of the pasture. However, marker methods have limitations, and are difficult to apply for the lengthy periods needed to get robust estimates of an animal's underlying intake of pasture.

The development of a practical measure of feed intake for all classes of animals maintained in a pasture-based environment, would provide a means of estimating the heritability and genetic correlations necessary to evaluate the utility of direct and indirect selection criteria for a range of breeding objectives.

Application of new technologies to measurement of feed intake and efficiency in livestock. The particular challenge associated with developing a robust and precise method of measuring feed intake in grazing animals is the absence of an existing methodology to use as a high quality benchmark against which to train the predictive algorithms developed from the sensor data.

The first step in developing means of economically estimating feed intake *en masse* is to identify a suite of sensors that are likely to exhibit a response that correlates with pasture intake and determining where these sensors are best located in and on the animal. A study jointly initiated by CSIRO and NSW Department of Primary Industries at Armidale in NSW is being employed to test a range of sensors, from the perspective of size, cost, weight, energy usage, sensor longevity and impact of the sensor on the pasture intake of the animal. Two locations have initially been trialled to assess likely survivability (longevity of the device) and non-obtrusiveness. One approach was to mount a suite of sensors on an eartag and the other as a device attached to a halter, adjacent to the mouth of the animal or on the back of the head.

With respect to sensors, two sensing modalities are being initially employed. An Inertial Measurement Unit (IMU), comprised of 3-axis accelerometers, 3-axis magnetometers and a pressure sensor for gross-height change detection were selected due to the need of the animal to move its head (and mouth) in order for feed intake to occur. Similarly, the ripping and chewing of feed matter (and drinking) will necessarily produce sounds which could potentially be used to estimate feed intake (Galli *et al*. 2011).

The second step in developing a practical and economic means of estimating pasture intake is development of algorithms that classify the low-level, high-sample-rate *input* sensor data into *output* behaviours such as foraging, biting, chewing, ripping, ruminating, drinking, sleeping etc. Mapping the inputs directly to pasture intake was deemed impractical as this would have required measuring the pasture intake at a frequency similar to input sample rate (faster than 1Hz). The input features (every accelerometer and magnetometer axis and the pressure value as well as a number of audio statistics over various window sizes) inherently provide different levels of predictive power and so need to be scaled and weighted. However, in order to determine the appropriate predictive power of any feature using a supervised learning approach requires a training dataset to be compiled which is comprised of the potential *input* features and the *outputs* (behaviours / traits). Therefore, a multi-day, multi-animal trial, recording the raw data (IMU and audio) was performed with simultaneous recording of benchmark methods of measuring feed intake and animal behaviour. These benchmark methods of biomass disappearance, chemical markers and highly annotated video by experts, required significant human and technological resourcing per reading and is a key motivation for our use of sensors and predictive algorithms.

Figure 1 shows the uncalibrated ("raw") accelerometer and magnetometer traces over a 45 minute period for an Angus steer in a field-grazing environment. The animal exhibits a variety of behaviours (as evidenced from video footage), ranging from foraging, standing still, visual searching, and continuous episodic grazing. The low accelerometer variation when the animal is still, or visually searching, highlights the advantage of utilising multiple sensors as the magnetometer trace can be used to differentiate these two different behaviours. Similarly magnetometer readings alone do not clearly differentiate the continuous episodic grazing activity correctly.

Based on evidence from other studies the inclusion of acoustic data provides significant additional power to discriminate between biting and chewing actions which allowed accurate estimations to be made of dry matter intake in grazing sheep (Galli *et al.* 2011).

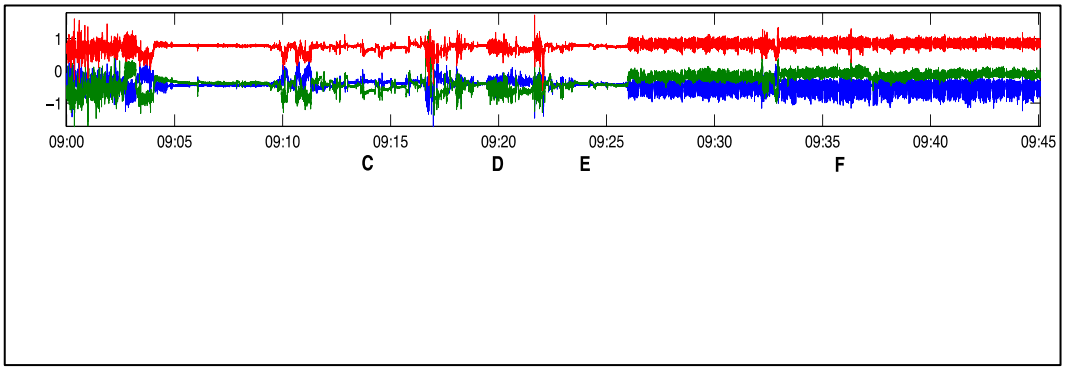


Figure 1. Accelerometer (100Hz sampling) and Magntometer (10Hz sampling) trace for 45 minute period for a grazing steer. Annotation of time refers to A) foraging, B) stationary, C) visual searching, D) foraging again, E) visual searching again, and F) continuous episodic grazing.

CONCLUSIONS

The use of technologies built around electronic sensors and sensor networks offers great promise for the phenotyping of large numbers of animals in their normal commercial environment. Initial experiments on small numbers of animals using video data to benchmark behaviours associated with grazing have provided a platform from which to develop robust predictive algorithms. The particular challenges in the further development of this phenotyping method lie in the management of the very large volumes of data that are an integral part of this methodology and the design and management of the power source.

REFERENCES

Dove H. and Mayes R. (2006) *Nat. Protoc.***1**:1680.

Galli J.R., Cangiano C.A., Milone D.H., and Laca E.A. (2011) *Livest. Sci.* **140**:32.

Hancock R.N., Swain D.L., Bishop-Hurley G.J., Patison K.P., Wark T., Valencia P., Corke P., and O'Neill, C.J. (2009) *Sensors* **9**: 3586.

Houle D., Govindaraju D.R., and Omholt, S. (2010) *Nat. Rev. Genet.* **11**: 855.

Pollak E.J., Bennett G.L., Snelling W. M., Thallman R.M., and Kuehn L. A. (2012) *Anim. Prod. Sci*. **52**: 92.

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:286-289

PHENOTYPIC ASSOCIATIONS BETWEEN METHANE PRODUCTION TRAITS, VOLATILE FATTY ACIDS AND ANIMAL BREEDING TRAITS

R.M. Herd[1], S.H. Bird[1], K.A.Donoghue[2], P.F. Arthur[3] and R.F. Hegarty[4]

[1] NSW Department of Primary Industries, Beef Industry Centre, Armidale, NSW, 2351
[2] NSW Department of Primary Industries, Agricultural Research Centre, Trangie, NSW, 2823
[3] NSW Department of Primary Industries, Agricultural Institute, Menangle, NSW, 2568
[4] Environmental and Rural Science, University of New England, Armidale, NSW, 2351

SUMMARY

This paper reports results for 532 young Angus bulls and heifers measured for methane production in respiration chambers. The animals were tested on a roughage ration offered at 1.2-times maintenance, based on their pretest weight (WT). Rumen fluid was collected for analysis of volatile fatty acids (VFA). Mean WT was 410 ± 93 kg (sd), daily methane production (MP) 204 ± 31 L/day, MP per unit weight (MI) was 0.52 ± 0.10 L/kg WT, and methane produced per unit dry-matter intake (DMI; MY) was 29.9 ± 4.4 (L/kg DMI). Pearson correlations showed that WT and feed intake were moderately positively correlated with MP, and negatively correlated with MI and MY. Concentrations of the three most abundant VFA (acetate, propionate and butyrate), total VFA, and the molar proportions of propionate and butyrate were associated with variation in MI and MY, but less so with variation in MP. There were statistically-significant associations for MP and MI, but not MY, with standard BREEDPLAN weight and carcase weight EBV, but either no or weak associations with other carcase EBV. These preliminary results show that BREEDPLAN EBV could be used to reduce the intensity of cattle methane emissions, but not MY which appears to be independent of genetic variation in growth and carcase traits.

INTRODUCTION

Cattle and sheep emit methane, a potent greenhouse gas, as part of the fermentation process in their rumen. There is a strong positive relationship between feed intake and methane production. Direct selection for lower daily methane production (MP) may not be desirable because it could favour lower feed intake and/or lighter and slower growing animals. Methane intensity (MI) and methane yield (MY), being methane produced per unit of weight and per unit of feed intake respectively, measure methane production that is independent of size and feed intake.

MATERIALS AND METHODS

The Angus cattle measured were born in 2009 and 2011 in research herds at Agricultural Research Centres at Glen Innes and Trangie, NSW. The cattle were measured for methane production in 2011 and 2012 in the 10 open circuit respiration chambers at the University of New England campus, Armidale, NSW. The cattle were trucked to Armidale, with about 40 animals constituting a safe weight load for transport. Each year, within each herd and sex, cohorts of 40 head in 4 groups of 10 were formed and prepared for measurement. Progeny of individual sires were stratified across groups and cohorts. Before transport, the cohort of 40 animals was weighed and then fed in the groups of 10 an amount calculated to provide about 1.2 to 1.5-times their estimated energy requirement for maintenance, based on equation 1.21 of SCA (1990). The test ration was a commercial lucerne and oaten hay chaff (Manuka “Blue Ribbon” Chaff®), chosen to mimic good, dry pasture. Regular samples were taken for feed analysis over the 2 years and had an average content of 88% dry matter (DM), a crude protein content of 14% DM and metabolizable energy content of 9MJ/kg DM. After 10 days the animals were weighed again, with this weight used as their pretest weight (WT), and then transported to Armidale. There cattle were kept in their

groups of 10 and fed the same amount of the same chaff ration for a minimum of 4 days. Then the first of the 4 groups was moved into the animal house and each animal fed in an individual pen at 1.2-times maintenance based on its WT. Feeding a restricted daily allowance proportional to WT was done to avoid feed refusals, minimise day to day variation in daily MP and to avoid 'level of feeding' effects on MY.

Methane production was measured over 2 x 24h consecutive periods. Animals were placed in their chambers by 10.00, with their daily feed allowance in a feed bin and water available from a drinker inside the chamber. After 24 hours the chambers were briefly opened and the feed bin replaced with a clean bin and fresh feed. After 48 hours the animals were let out of the chambers, briefly restrained, and a sample of rumen fluid aspirated through a flexible stomach tube. The rumen fluid was preserved by acidification and then stored at -18^0C for subsequent analysis by gas liquid chromatography of VFAs, being products of the fermentation in the rumen. Most animals consumed their daily feed allowance within 8 to 12 hours, so for most animals these VFA concentrations represent levels following a short period of 12-or-so hours without fresh feed. The open circuit respiration chamber consists of an enclosed pen (1.8m x 3m) within a polycarbonate shell (3.6m x 2.4m x 2.4m), each with an individual mass flow meter and airflow subsample line connected by a multiplexer to a Servomex gas analyser.

Data on 218 animals of approximately 2-years of age tested in 2011 and 314 yearling-age animals in 2012 was used. The magnitude of phenotypic associations between methane production over the second-24hr period and VFA were determined by calculating Pearson correlation coefficients. Variation in methane traits associated with genetic variation in cattle breeding traits, as measured by standard BREEDPLAN® EBV, was assessed by the calculating regression coefficients for the traits against EBV in a general linear model. These EBV are described in BREEDPLAN (2010). Only EBV reported for at least 525 of the 532 animals are analysed. Fixed effects fitted were year, herd and sex, and their 2-way interactions. The interactions of EBV with the fixed effects (that is, differences in the slopes of the relationships) were mostly non-significant and are not reported. Pearson correlations and regression coefficients that differed from zero ($P<0.05$) were taken as evidence for a statistically-significant association. BREEDPLAN EBV for weight and carcass traits, extracted in May 2013, were used.

RESULTS

Summary statistics are presented in Table 1. There was substantial variation in the methane production traits measured. Pretest weight and DMI were strongly positively correlated with MP, and negatively correlated with MI and MY (Table 2). Significant correlations for the three most abundant VFA (acetate, propionate and butyrate), total VFA, and the molar proportions of propionate and butyrate, showed that variation in VFA production was associated with variation in MI and MY, but less so for variation in MP.

Fitting the fixed effects of year of test, then herd and sex each explained part of the variation in MP (3%, 17%, 8% respectively), MI (71%, 0.1%, 3%) and MY (55%, 1%, 4%), and together with their interactions explained 32%, 76% and 63% of the variation in MP, MI and MY, respectively. Variation in MP had statistically-significant associations with weight EBV and the EBV for carcase_wt and fat (Table 2). Variation in MI had statistically-significant associations with weight EBV, EBV_carcase_wt and EBV_ema, whereas MY was not associated with any EBV except EBV_ema.

DISCUSSION

These preliminary results show large natural variation between animals in MP, MI and MY. Heavier animals were offered more feed and had a greater MP, which might be considered undesirable, but had lower MI and MY, which is desirable in terms of greenhouse gas emissions.

Should these correlations hold at a genetic level, then selection for lower MP could reduce DMI and animal performance, and increase both MI and MY.

Table 1. Summary statistics for n=532 young Angus bulls and heifers tested for methane production in 2011 and 2012

Trait	Mean	SD	Maximum	Minimum
Pretest weight (WT kg)	410	93	670	229
Dry-matter intake DMI; kg/d)	6.9	1.1	9.5	4.6
Methane production (L/d)	204	31	350	115
Methane intensity (L/kg WT)	0.52	0.10	0.68	0.24
Methane yield (L/kg DMI)	29.9	4.4	41.2	15.0
Acetate (mmoles/L)	38.6	10.9	82.5	6.4
Propionate (mmoles/L)	8.1	2.9	36.2	1.1
Isobutyrate (mmoles/L)	0.6	0.2	1.4	0.1
Butyrate (mmoles/L)	4.2	1.5	10.5	0.7
Isovalerate (mmoles/L)	1.1	0.3	2.4	0.2
Valerate (mmoles/L)	0.4	0.2	1.9	0.0
Total VFA (mmoles/L)	53.0	15.2	131	8.6
Acetate%	73.0	1.6	77.3	63.1
Propionate%	15.1	1.7	27.7	12.4
Butyrate%	7.9	1.0	13.5	3.2
EBV_BWT (kg)	3.9	1.9	10.2	-0.7
EBV_200d_wt (kg)	23.8	7.0	42.0	3.0
EBV_400d_wt (kg)	45.8	12.2	78.0	4.0
EBV_600d_wt (kg)	56.6	16.1	104	9.0
EBV_carcase_wt (kg)	28.3	9.6	57.0	-4.0
EBV_ema (cm^2)	0.9	1.4	5.4	-2.8
EBV_rib_fat (mm)	0.8	1.7	6.6	-4.0
EBV_rump_fat (mm)	0.6	1.8	6.8	-3.4
EBV_rby (%)	-0.1	0.9	2.0	-3.1

Higher concentrations of VFA were associated with higher MI and MY, but not with variation in MP. The molar proportions of the major VFA have been previously shown to be related to MY in cattle, with propionate being negatively, and butyrate positively, related to MY (Whitelaw *et al.* 1984), as found in this experiment. A high propionate pattern is a desirable as hydrogen from rumen fermentation is used in its formation, rather than being used for CH_4 formation. These relationships open the possibility of using VFA as indicator traits for methane emissions. However, the strength of their phenotypic association, with correlation coefficients of 0.4 or lower, are too low to be either an accurate screening test for high or low emitting cattle, or for genetic improvement based on indirect selection, conclusions also reached for sheep by Robinson *et al.* (2010). In this experiment, the VFA concentrations for most animals represent levels following a short period of 12-or-so hours without fresh feed, and may have been different had rumen fluid been sampled during peak fermentation that occurs immediately following feed consumption. More careful scrutiny of VFA as markers for methane emissions is still warranted.

The statistically-significant associations for MP and MI, but not MY, with standard BREEDPLAN weight EBV provides evidence that MP and MI could be altered by appropriate emphasis on these EBV. Increase in weight EBV was associated with greater MP (usually undesirable), lower MI (desirable) and no change in MY. The carcase EBV, except carcase_wt,

had little or no association with variation in methane emission traits. Preliminary estimates of genetic correlations are reported in Donoghue *et al*. (2013).

These results show that substantial natural variation in methane emissions is present in these Angus cattle, and that in BREEDPLAN the beef industry has a system which could be used to reduce the intensity of methane greenhouse gas emissions, but perhaps not MY which appears to be independent of genetic variation in the recorded growth and carcase composition traits.

Table 2. Pearson correlations for methane production (MP), methane intensity (MI) and methane yield (MY) with weight, feed intake and volatile fatty acids (VFA), and regression coefficients (±se) with BREEDPLAN EBV

	MP (L/day)	MI (L/kg WT)	MY (L/kg DMI)
Pretest weight (WT; kg)	0.53***	-0.79***	-0.59***
Dry-matter intake (DMI; g/d)	0.55***	-0.76***	-0.59***
Acetate (mmoles/L)	-0.07	0.33***	0.29***
Propionate (mmoles/L)	-0.05	0.16***	0.13**
Butyrate (mmoles/L)	-0.09*	0.40***	0.35***
Total VFA (mmoles/L)	-0.07	0.32***	0.28***
Acetate %	0.00	0.06	0.07
Propionate %	0.05	-0.41***	-0.39***
Butyrate %	-0.14**	0.44***	0.38***
EBV_BWT (kg)	3.5 ± 0.6***	-2.6 ± 1.2*$^{+}$	0.07 ± 0.06
EBV_200d_wt (kg)	1.6 ± 0.2***	-0.9 ± 0.3**$^{+}$	0.03 ± 0.02
EBV_400d_wt (kg)	.89 ± 0.08***	-0.5 ± 0.2**$^{+}$	0.02 ± 0.01
EBV_600d_wt (kg)	.69 ± 0.06***	-0.5 ± 0.1***$^{+}$	0.01 ± 0.01
EBV_carcase_wt (kg)	1.1 ± 0.1***	-0.6 ± 0.2**$^{+}$	0.02 ± 0.01
EBV_ema (cm^2)	0.3 ± 0.8	-5.0 ± 1.5***$^{+}$	-0.19 ± 0.09*
EBV_rib_fat (mm)	-1.6 ± 0.7*	-1.1 ± 1.2^{+}	-0.11 ± 0.07
EBV_rump_fat (mm)	-1.5 ± 0.6*	-1.3 ± 1.2^{+}	-0.11 ± 0.07
EBV_rby (%)	1.2 ± 1.3	-0.6 ± 2.3^{+}	0.01 ± 0.13

*P<0.05; **P<0.01; ***P<0.001. $^{+}$These coefficients and se have been multiplied by 1,000.

ACKNOWLEDGEMENTS

This project was supported by funding from the Australian Government Department of Agriculture, Fisheries and Forestry as part of its Carbon Farming Futures Filling the Research Gap Program, and by Meat & Livestock Australia. We thank David Mula, Karen Dibley, Peter Newman, Reg Woodgate and Kim Quinn for skilled technical assistance.

REFERENCES

BREEDPLAN (2010) BREEDPLAN EBVs "The traits explained". ABRI, Armidale.

Donoghue K.A., Herd R.M., Bird S.H., Arthur P.F. and Hegarty R.F. (2013) *Proc. Assoc. Advmt. Anim. Breed. Genet*. **20**: (This proceedings).

Herd R.M., Donoghue K.A., Bird S.H., Hegarty R.F. and Arthur P.F. (2011) *Proc. Assoc. Advmt. Anim. Breed. Genet*. **19**: 187.

Robinson D.L., Goopy J.P., Hegarty R.S. and Vercoe P.E. (2010) *9th Wld Congr. Genet. Appl. Livest. Prod*. Leipzig, Germany. http://www.kongressband.de/wcgalp2010/assets/pdf/0712.pdf.

SCA. (2000) 'Feeding Standards for Australian Livestock. Ruminants' CSIRO Publications, East Melbourne.

Whitelaw F.G., Eadie J.M., Bruce L.A. and Shand W.J. (1984) *Brit. J. Nutr.* **52**: 261.

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:290-293

PRELIMINARY GENETIC PARAMETERS FOR METHANE PRODUCTION IN AUSTRALIAN BEEF CATTLE

K.A. Donoghue[1], R.M. Herd[2], S.H. Bird[2], P.F. Arthur[3] and R.F. Hegarty[4]

[1] NSW Department of Primary Industries, Agricultural Research Centre, Trangie, NSW 2823
[2] NSW Department of Primary Industries, Beef Industry Centre, Armidale, NSW 2351
[3] NSW Department of Primary Industries, Elizabeth Macarthur Agricultural Institute, Menangle, NSW 2568
[4] Environmental and Rural Science, University of New England, Armidale, NSW 2351

SUMMARY

This paper reports the first heritability estimates for methane traits in beef cattle, using records from 530 young Angus bulls and heifers measured for methane production in respiration chambers. Weight (WT) and ultrasound scan traits (eye muscle area: EMA; rump fat depth: P8; rib fat depth: RIB; intramuscular fat percentage: IMF) were also recorded on these animals in order to investigate the relationships between methane and production traits. Heritabilities for daily methane production (MP), methane production per unit feed intake (methane yield: MY) and methane production per unit weight (methane intensity: MI) were low to moderate (0.21, 0.19 and 0.23, respectively). Methane traits (MP, MY and MI) were not correlated, either phenotypically or genetically, with the body composition traits. These preliminary results show that there may be some potential to use genetic improvement to reduce methane emissions in beef cattle.

INTRODUCTION

Cattle and sheep emit methane, a potent greenhouse gas, but currently there are few technologies available to mitigate methane emissions in extensive beef production systems. Genetic improvement is capable of producing small but permanent and cumulative changes in performance, and is particularly useful in extensive production systems as found in majority of Australian beef herds. Thus, genetic improvement is an attractive approach for the mitigation of methane emissions in Australian beef cattle. In order to assess the viability of this mitigation approach, genetic variation in methane traits along with relationships with important production traits must be quantified. This paper provides preliminary heritability estimates for methane traits, along with estimates of phenotypic and genetic relationships with production traits.

MATERIALS AND METHODS

Progeny born in 2009 (n=218) and 2011 (n=312) from Angus cows in 2 research herds at the Agricultural Research Centre, Trangie NSW, were measured for methane production in 2011 and 2012 in 10 respiration chambers on the University of New England campus, Armidale NSW. For progeny born in 2009, males from both herds and females from one of the herds were measured for methane, while for progeny born in 2011, animals from both sexes in both herds were measured. Each year, animals were allocated into cohorts within herd and sex (n=40), fed a restricted diet (1.2-times the estimated energy requirement for maintenance) and groups of 10 animals were individually measured in the respiration chambers. The 530 animals were progeny of 38 sires (average 14 progeny per sire, range 1-33). Progeny of individual sires were stratified across groups and cohorts. Herd *et al.* (2013) provides details on the diet and measurement procedure.

Data. Methane production was measured over 2 x 24h consecutive periods. For animals born in 2011 these measurements were taken at approximately yearling age (mean=369 days). However, for animals born in 2009, these measurements were taken at approximately two years of

age (mean=748 days) due to delays in construction of the facility. Traits measured included pre-test weight (WT), dry matter intake (DMI), daily methane production (MP; litres of methane per day), methane production per unit feed intake (methane yield: MY) and methane production per unit weight (methane intensity: MI). Editing of records included removal of animals with incomplete pedigrees, missing birth date and trait measurements greater than 4 standard deviations from the contemporary group mean.

Live animal ultrasound scans were collected on all animals in the research herds at approximately 600 days of age by a certified ultrasound technician, including animals which had not been measured for methane. There were ultrasound scan records available on 750 animals, who were the progeny of 38 sires (average 20 progeny per sire, range 1-38). Traits recorded included eye muscle area (EMA), fat depth at the rump (P8) and rib (RIB) sites and intramuscular fat percentage (IMF).

Model of analysis. Variance components were estimated using ASReml (Gilmour *et al*. 2009). For methane and scan traits, the fixed effect of contemporary group (CG) was included in the model and single record contemporary groups were excluded from the analysis. For methane traits, CG definition included cohort and methane group and management group. Age of the animal on the date of measurement was included as a covariate for all methane traits, and WT was also included for the MP trait. For scan traits, CG definition was the same as for methane traits for those animals with a methane record. For animals without a methane record, CG definition included birth year, sex and management group. Age of the animal on the date of measurement was included as a covariate for all scan traits. 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. Bivariate analyses of all trait combinations were also conducted.

RESULTS AND DISCUSSION

Table 1 contains summary statistics for the methane test and body composition scan data. A large amount of variation was observed for both WT and age of methane measurement due to the older age of measurement of the animals born in 2009. The methane traits (MP, MY and MI) exhibited substantial phenotypic variation even after adjustments for DMI (MY) and WT (MI).

Table 1. Descriptive statistics for methane and body composition scan traits

	No. records	Average (SD)	Minimum	Maximum
Methane traits				
Age (days)	530	525 (192)	264	822
WT (kg)	530	410 (93)	229	670
DMI (kg/d)	530	6.9 (1.2)	4.6	9.5
MP (L/d)	530	205 (30)	122	350
MY (L/kg DMI)	530	29.9 (4.2)	15.9	41.2
MI (L/kg WT)	530	51.6 (9.4)	25.8	67.8
Scan traits				
Age (days)	750	613 (83)	483	791
EMA (cm^2)	750	59.9 (7.6)	35.0	96.0
P8 (mm)	750	5.5 (3.9)	1.0	32.0
RIB (mm)	750	3.8 (2.6)	1.0	18.0
IMF (%)	730	3.5 (1.2)	1.5	8.1

Genetic parameters for methane and scan traits are reported in Table 2. This study provides the first heritability estimates for methane traits in beef cattle. Heritabilities for methane traits (MP, MY and MI) were low to moderate (0.19-0.23), with relatively large associated standard errors.

Robinson *et al.* (2010) reported a low heritability (0.13) for MI in sheep, while Pinares-Patino *et al.* (2011) reported a moderate heritability (0.30) for MY, also in sheep. Several dairy studies have predicted MP using DMI, and heritabilities reported range from 0.12 (Cassandro *et al*., 2010) to 0.35 (de Haas *et al*., 2011) for MP and 0.58 (de Haas *et al*., 2011) for MP adjusted for milk production (similar to MI in this study). Results from this study, along with other published estimates, indicate that there may be some potential to use genetic improvement to reduce methane emissions in livestock. However, further investigations involving larger numbers of animals are needed.

Heritabilities reported in this study for WT, EMA, P8 and IMF were very similar to published estimates in Australian Angus animals (Jeyaruban *et al*., 2009; Meyer, 2005). While the heritability for RIB (0.63) was higher than reported estimates in young Angus animals (0.28-0.45; Meyer, 2005), it was similar to published estimates in Angus cows (Donoghue *et al*., 2009).

Table 2. Genetic parameters (SE) for weight, methane and scan traits

	σ^2_a (SE)	σ^2_e (SE)	σ^2_p (SE)	h^2 (SE)
WT	612 (210)	803 (172)	1,415 (101)	0.43 (0.13)
MP	72 (38)	271 (37)	343 (23)	0.21 (0.11)
MY	1.1 (0.6)	4.6 (0.6)	5.7 (0.4)	0.19 (0.10)
MI	4.1 (1.9)	13.7 (1.9)	17.8 (1.2)	0.23 (0.10)
EMA	13.0 (4.1)	18.5 (3.3)	31.5 (1.9)	0.41 (0.12)
P8	3.7 (0.9)	3.2 (0.7)	6.9 (0.4)	0.53 (0.12)
RIB	1.8 (0.4)	1.1 (0.3)	2.9 (0.2)	0.63 (0.12)
IMF	0.24 (0.07)	0.40 (0.06)	0.63 (0.04)	0.37 (0.10)

Phenotypic (r_p) and genetic (r_g) correlations and their associated standard errors between all traits are reported in Table 3. Large positive r_p (0.89-0.96) were observed among methane traits, indicating that, phenotypically, animals with higher MP also had higher MY and MI. Heavier animals had higher MP (0.58) and lower MI (-0.28), but no difference in MY (0.05) than lighter animals. In a dairy study, de Haas *et al*. (2011) reported a phenotypic correlation between predicted MP and milk production of 0.26. As expected, large positive phenotypic relationships were observed between the two measures of external fat (0.91; P8-RIB) and between external and internal fat measures (0.73-0.75; P8/RIB-IMF). Results from this study indicate that, phenotypically, there was no relationship between methane traits (MP, MY and MI) and scan traits (-0.16 to 0.07).

Large positive r_g (0.87-0.96) were observed among methane traits, indicating that, genetically, animals with higher MP also had higher MY and MI. While WT was highly positively correlated with MP (0.79), it was lowly positively correlated with MY (0.18) and lowly negatively correlated with MI (-0.23). Previous literature estimates for genetic correlations between predicted MP and milk production range from 0.31 (de Haas *et al*., 2011) to 0.92 (Cassandro *et al*., 2010), while de Haas *et al*. (2011) reported a large negative correlation (-0.87) between MI and milk production. As expected, large positive genetic relationships were observed between the two measures of subcutaneous fat (0.99; P8-RIB) and between subcutaneous and intramuscular fat measures (0.97-0.98; P8/RIB-IMF). In these preliminary results, no evidence of strong genetic relationships between methane traits and scan traits was observed (-0.23 to 0.29), indicating that selection for methane traits would have little impact on body composition. It should be noted that most of the genetic correlation estimates in this study have large standard errors and further investigations are warranted once more data are available. High correlations between the different methane traits (methane production, methane yield and methane intensity) indicate that, phenotypically and

genetically, methane is independent of feed intake.

Table 3. Genetic (above diagonal) and phenotypic (below diagonal) correlations (SE) for methane and scan traits

Trait	WT	MP	MY	MI	EMA	P8	RIB	IMF
WT	-	0.79	0.18	-0.23	0.55	-0.07	-0.01	0.08
		(0.12)	*(0.30)*	*(0.28)*	*(0.16)*	*(0.21)*	*(0.20)*	*(0.22)*
MP	0.58	-	0.96	0.95	0.17	0.18	0.16	0.29
	(0.03)		*(0.04)*	*(0.04)*	*(0.29)*	*(0.25)*	*(0.25)*	*(0.27)*
MY	0.05	0.93	-	0.87	-0.02	0.12	0.08	0.21
	(0.05)	*(0.02)*		*(0.09)*	*(0.30)*	*(0.26)*	*(0.26)*	*(0.28)*
MI	-0.28	0.96	0.89	-	-0.23	0.15	0.14	0.21
	(0.05)	*(0.03)*	*(0.01)*		*(0.27)*	*(0.25)*	*(0.24)*	*(0.26)*
EMA	0.44	-0.01	-0.03	-0.16	-	0.19	0.21	0.41
	(0.04)	*(0.05)*	*(0.05)*	*(0.05)*		*(0.19)*	*(0.18)*	*(0.18)*
P8	0.18	0.04	0.06	-0.01	0.17	-	0.99	0.97
	(0.05)	*(0.05)*	*(0.05)*	*(0.05)*	*(0.04)*		*(0.01)*	*(0.04)*
RIB	0.17	-0.004	0.01	-0.05	0.19	0.91	-	0.98
	(0.05)	*(0.05)*	*(0.05)*	*(0.05)*	*(0.04)*	*(0.01)*		*(0.03)*
IMF	0.19	0.07	0.07	-0.01	0.19	0.75	0.73	-
	(0.05)	(0.05)	(0.05)	(0.05)	(0.04)	(0.02)	(0.02)	

CONCLUSIONS

These preliminary results show that genetic variation in methane emissions is present in these Angus cattle. No antagonistic phenotypic or genetic relationships between methane and body composition traits were identified. Thus, genetic improvement to reduce methane emissions may be possible, but further investigations involving larger numbers of animals are needed.

ACKNOWLEDGMENTS

Project funds came from the Australian Government Department of Agriculture, Fisheries and Forestry as part of its Carbon Farming Futures Filling the Research Gap Program, and Meat & Livestock Australia. We thank David Mula, Karen Dibley, Peter Newman, Reg Woodgate and Kim Quinn for skilled technical assistance.

REFERENCES

Cassandro, M., Cecchinato, A., Battagin, M. and Penasa, M. (2010) Proc. 9th WCGALP, www.kongressband.de/wcgalp2010/assets/pdf/0837.pdf.

de Haas, Y., Windig, J.J., Calus, M.P.L, Dijkstra, J., de Haan, M., Bannink, A. and Veerkamp, R.F. (2011) *J. Dairy Sci.* **94**:6122.

Donoghue, K.A. and Parnell, P.F. (2009) *Proc. Assoc. Advmt. Anim. Breed. Genet*. **18**: 117.

Herd, R.M., Bird, S.H., Donoghue, K.A., Arthur, P.F. and Hegarty, R.F. (2013) *Proc. Assoc. Advmt. Anim. Breed. Genet*. **20**: (This proceedings)

Gilmour A.R., Cullis B.R., Welham S.J. and Thompson R. (2009) 'ASREML. Program user manual' NSW Agriculture, Orange Agricultural Institute, Forest Road, Orange, NSW, 2800, Australia.

Jeyaruban, M.G., Johnston, D.J. and Graser, H.-U. (2009) *Anim. Prod. Sci.* **49**:1.

Meyer, K.(2005) *Animal Science*. **81**:337.

Pinares-Patino, C.S., McEwan, J.C., Dodds, K.G., Cardenas, E.A., Hegarty, R.S., Koolaard, J.P. and Clark, H. (2011) *Animal Feed Science and Technology.* **166-67**:210.

Robinson, D.L., Goopy, J.P., Hegarty, R.S. and Vercoe, P.E. (2010) Proc. 9th WCGALP, www.kongressband.de/wcgalp2010/assets/pdf/0712.pdf.

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:294-297

A POTENTIAL PRACTICAL SYSTEM TO ESTIMATE PASTURE INTAKE OF INDIVIDUAL ANIMALS

D.J. Cottle

School of Rural and Environmental Science, University of New England, Armidale, NSW, 2351

SUMMARY

Pasture feed intake of individual animals is very difficult to estimate. A practical measurement system would better enable the selection of livestock for pasture feed use efficiency and lower maintenance requirements, which are very important biological and economic traits. A prototype (Proway-CottleDove) feed bin system was trialed by comparing chaff intakes by cattle in a feedlot measured by a race fed autofeeder with intakes estimated by marker analysis of faecal and feed samples following the controlled daily consumption of wax labeled supplement. Intake of a native pasture was also estimated by use of the bin. Autofeeder-recorded daily chaff intakes were very variable and unreliable and so the accuracy of the bin system in estimating intake could not be determined. The repeatabilities of chaff intakes estimated from marker concentrations from sequential faecal samples were 0.2-0.3. Chaff intake predictions were in a feasible range, based on cattle liveweight. When the pasture grasses were combined in analyses, following a principal component analysis of markers, the diet composition, digestibility and daily intake of pasture plus supplement estimates, on average, were consistent with the prediction of intake from liveweight and liveweight gain. The average total intakes estimated from days 5, 7 and 10 faeces marker concentrations were 7.8kg, 6.9kg and 9.7kg/head respectively. The bin system used in this trial would estimate pasture intake at an approx. cost of $122/head. Multi-bin systems using Sapien Technology components and databases are being developed for further testing.

INTRODUCTION

The cost of feed is second only to capital costs in importance to the profitability of commercial beef operations. About 70-75% of the total dietary energy cost in a beef cow herd is used for maintenance and these requirements of beef cattle have remained largely unchanged over the last 100 years (Basarab *et al.* 2005). Pasture intake (and the efficiency of its use for liveweight gain) has always been difficult to measure under field conditions. Methods for estimating pasture intake have been extensively reviewed (Langlands 1987; Dove and Mayes 1996; Mayes and Dove 2000; Dove and Mayes 2005; Dove 2010; Crews and Carstens 2012; Cottle 2013). Measurements can be based on plant biomass or be made on animals. Estimates of plant biomass before and after grazing by a mob or herd do not provide estimates of individual animal intake. Some measurement methods can disturb normal grazing behavior and interfere with intakes. Livestock selectively graze (Hanley 1982), so their diet cannot be easily quantified using plant-based measurements.

Residual feed intake (RFI) can be used to directly select for feed use efficiency (Cottle 2011) however the high cost of RFI measurement in a feedlot (~$A500/head) and RFI's interaction with feed type and level (Herd *et al.* 2011) has limited the use of RFI by industry. Hugh Dove and co-workers developed the approach of feeding weighed amounts of wax-labeled supplement to dose animals with natural markers. This has been turned into a more practical approach by enabling the animal to self-dose in the paddock with labelled supplement via a purpose-built feed bin with an electronic identity device tag reader (patent pending). The bin has mechanisms to control and record the daily labeled supplement intake of each animal and in-house algorithms are used to calculate individual pasture intake.

MATERIALS AND METHODS

Animals. In trial 1, Angus-Charolais cross heifers grazed at pasture and in trial 2 these heifers were fed lucerne chaff followed by oaten:lucerne chaff in the feedlot. All cattle that ate labeled supplement regularly were kept in the trial paddock or feedlot pens in each feeding trial.

Feed and bins. The grazing paddock in trial 1 mainly contained wallaby, parramatta and red grass, paspalum, setaria and white clover. Cattle were fed *ad lib* 100% lucerne chaff or *ad lib* 50% lucerne chaff:50% oaten chaff (Manuka Chaff, Tamworth) through race auto-feeders in feedlot pens at 'Tullimba'. The repeatability of daily chaff intake measured by the auto-feeders was less than 0.1, probably due to competition for access to the race and feed bin. The labeled supplement was cottonseed meal (CSM) pellets (Supreme Stockfeed, Guyra) containing 0.75% (w/w) beeswax and 30% oat hulls. The trials used a bin/race system modified from the initial prototype to control daily supplement intake.

Sampling and measurements. After many prototype bin/race technical issues were resolved and individual, daily CSM intakes were consistent, faecal and feed samples were taken at days 5 or 6, 8 or 9 and 10 or 11. Samples were oven dried at 65°C for 7 days, ground through a sieve and mailed to CSIRO, Canberra for analysis of alkane and alcohol content (ppm/OM) (Dove and Mayes 1996).

Statistical analyses. Pasture composition and organic matter intake (OMI) were estimated from the marker concentrations in feed, supplement and faeces using a modification of the methods described by Dove and Moore (1995), Dove (2010) and Cottle (2013). The repeatability of chaff intake prediction from the different faecal samples was calculated from the variances between and within animals for marker predicted chaff intake. The different grasses in the pasture trial were combined in analysis following a principal component analysis (PCA) of the markers (Dove 2010).

RESULTS AND DISCUSSION

Data from six alkanes (C25, 27, 29, 30, 31, 33) and four alcohols (C24, 26, 28, 30) were chosen for analyses. Autofeeder chaff intake measurements were unreliable, however the correlations between them and marker predicted lucerne chaff intakes were 0.96 from day 5 faeces samples; 0.64 from day 8 faeces samples; and 0.89 from day 11 faeces samples. The marker predicted lucerne chaff intakes were higher than autofeeder chaff intakes. These correlations for mixed chaff intake were 0.69 from day 6 faeces samples; 0.45 from day 9 faeces samples and 0.62 from day 11 faeces samples. Predicted chaff intakes of animals with at least 2 faecal samples are shown in Figure 1.

PCA analyses of plant alkane and alcohol concentrations were carried out to establish *a priori* if the 10 markers could distinguish between plant species. PCA scores 1 and 2 accounted for 98% of the variance in marker profile (Figure 2) and their biplot showed that the labelled CSM, wallaby grass and white clover were easier to distinguish than red grass, setaria, paspalum and parramatta grass.

A marker profile for a single component called 'grass' was calculated and the diet composition, whole-diet digestibility and total daily intake of pasture plus pellet estimates, on average of 11kg/day, were consistent with the prediction of total intake from liveweight and gain using the equation of Minson and McDonald (1987).

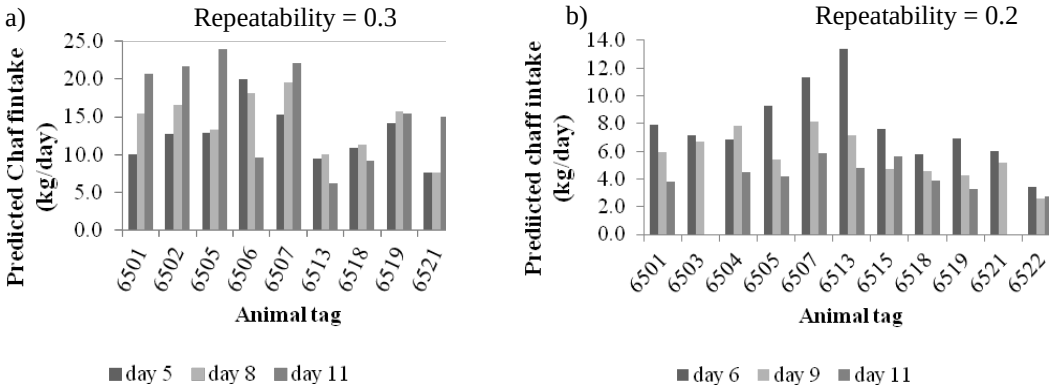


Figure 1. Predicted a) lucerne chaff and b) mixed chaff intake (kg/day) using C27, C29, C31, C33 alkanes and C24, C26, C28, C30 alcohols from 3 different faecal sampling days.

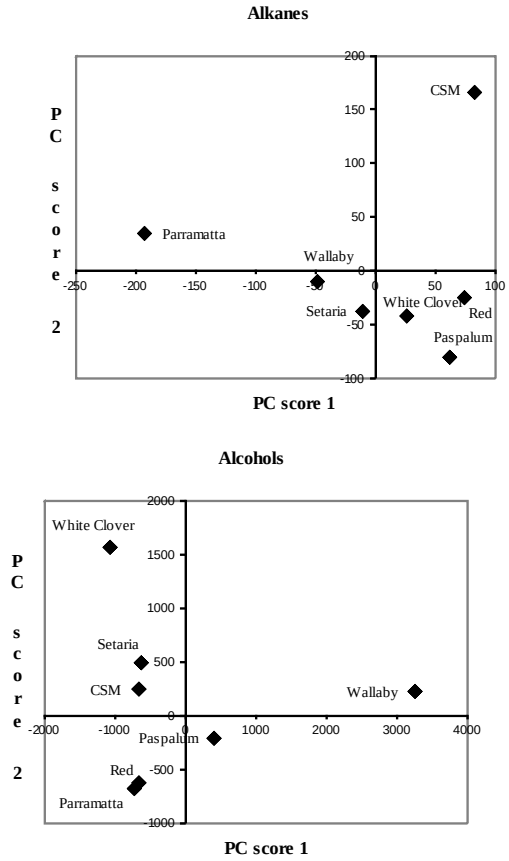


Figure 2. Clustering of plant species based on the first two principal components of alkane and alcohol marker concentrations.

The average total intakes predicted from days 5, 7 and 10 faecal marker concentrations were 7.8 kg, 6.9 kg and 9.7 kg/head respectively. There was a high between animal variation in predicted pasture intakes. These total intakes appeared lower than expected from liveweights and suggest that animals may have been grazing plant species that were not collected (e.g. demeter fescue) as the paddock contained diverse plant species that changed with season. The daily allowance of labelled supplement was consumed rapidly and this may have affected the steady state kinetics of herbage markers.

CONCLUSIONS

The intakes measured by the autofeeders were unreliable, so the accuracy of the intake estimates from the Proway system was unknown. However, the intakes predicted from feeding labelled supplement were realistic given the average liveweight of the heifers (455kg). The prototype bin system tightly controlled the maximum daily supplement intakes (daily supplement intake repeatability >0.9). Assuming a bin life of 20 years and a conservative 20 head tested per bin and a marker test cost of $71.50/sample, the cost per animal tested was $122, which is much less than a RFI test. The Proway-Sapien system shows promise as a practical means of measuring pasture intake and feed use efficiency.

ACKNOWLEDGEMENTS

I am grateful to Dr Hugh Dove for his advice, Andrew Hunter for cattle supply, Col Crampton for cattle husbandry, Adam Fitzpatrick and Jim Cook for autofeeder data processing, Bruce Hunt and Brad Dawson for bin engineering, Ken Bray for bin programming, Guy Ballard for supply of movement detection cameras, Scott McDonald for laboratory marker analyses and MLA for R&D trial funding.

REFERENCES

Basarab J.A., Ramsey P., French N., Crews D., Lyle K.L., Moore S.S. and Okine E.K. (2005) AARI Project Number: 2002A090R. Alberta Agriculture, Food and Rural Development, Lacombe, Canada.

Cottle D.J. (2011) *Proc. Aust. Assoc. Anim. Breed*. **19**: 423.

Cottle D.J. (2013) *Recent Advances in Animal Nutrition - Australia*. **22**: in press.

Crews D.H. and Carstens G.E. (2012) In 'Feed Efficiency in the Beef Industry', pp. 21-28, editor R.A. Hill, Wiley and Blackwell, Oxford, UK.

Dove H. and Mayes R.W. (1996) *Journal of Nutrition* **126:** 13.

Dove H. and Mayes R.W. (2005) *Small Ruminant Research* **59**: 123.

Dove H. and Moore A.D. (1995) *Australian Journal of Agricultural Research* **46**: 1535.

Dove H. (2010). In 'Proceedings 4th Grazing Livestock Nutrition Conference', pp. 31-56., editors B. W. Hess, T. DelCurto, J.G.P. Bowman and R.C. Waterman, Western Section American Society of Animal Science, Champaign, USA.

Hanley T.A. (1982) *Journal of Range Management* **35:** 146.

Herd R.M., Arthur P.F. and Archer J.A. (2011) *Proc. Assoc. Advmt. Anim. Breed. Genet.* **19**: 47.

Langlands J.P. (1987) In 'The Nutrition of Herbivores', pp. 363-390, editors J.B. Hacker and J.H. Ternouth, Academic Press, Sydney.

Mayes R.W. and Dove H. (2000) *Nutrition Research Review* **13:** 107.

Minson D.J. and McDonald C.K. (1987) *Tropical Grasslands* **21**: 116.

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:298-302

ACCURACY OF GENOMIC PREDICTION FOR RESIDUAL FEED INTAKE IN A MULTI-BREED CATTLE POPULATION

M. Khansefid[1,2,3], J.E. Pryce[1,2], S.P. Miller[4] and M.E. Goddard[1,2,3]

[1]Department of Environment and Primary Industries, AgriBio, 5 Ring Road, Bundoora, VIC, 3083, Australia
[2]Dairy Futures Cooperative Research Centre (CRC), AgriBio, 5 Ring Road, Bundoora, VIC, 3083, Australia
[3]Melbourne School of Land and Environment, The University of Melbourne, Parkville, VIC, 3010, Australia
[4]Department of Animal and Poultry Science, University of Guelph, Guelph, Ontario, N1G 2W1, Canada
email: m.khansefid2@student.unimelb.edu.au and majid.khansefid@depi.vic.gov.au

SUMMARY

Combining information from different cattle breeds is a potential way to improve the accuracy of genomic estimated breeding values (GEBVs) by increasing the size of the reference population. However, the phase of linkage disequilibrium between SNPs and quantitative trait loci for traits such as residual feed intake (RFI) may vary from one breed to another, which would erode the value of combining breeds. RFI is a selection criterion for feed efficiency and is the difference between actual intake and expected intake for maintenance and production. The aim of this research was to evaluate the accuracy of GEBVs when RFI records were combined from 5,614 animals of different breeds including 842 Holstein heifer and 2,009 Australian beef cattle (1,134 Angus, 217 Herford, 79 Murray Grey and 579 Shorthorn) and 2,763 Canadian beef cattle (534 Angus, 384 Charolais and 1,845 mixed synthetic breed) and their genotypes (606,096 SNPs) were used. We estimated the variance explained by the SNPs and the variance explained by SNP x breed interactions. The model with the highest likelihood was when SNP effects within two groups of breeds in addition to pedigree was fitted. The first group comprised Holsteins and the Angus cattle from the Trangie Research Station in NSW, Australia and the second group included all the other cattle. The difference between these two groups is that the cattle in group 1 were measured for RFI on a pelleted diet shortly after weaning while those in group 2 were measured on a feedlot diet at >1 year of age. According to the best model, the SNP effects were not significantly different between the two breeds fed a similar diet and measured at a similar age. However, the SNP effects differed between groups that were fed different diets and measured at different ages. The GEBVs of the validation animals were calculated using their SNP genotypes and the estimated SNP effects and correlated with their actual RFI phenotypes to estimate the accuracy of the GEBV. The average accuracy was 0.31 which was near to expected from the BLUP equations (0.34). Thus an across breed reference population appears to be promising for genomic prediction of RFI provided the animals are at about the same age and on a similar diet. However, there is only a small increase in accuracy by adding animals of another breed because the relationships between animals in different breeds are low. The BLUP equations correctly predict this limited increase in accuracy.

INTRODUCTION

Residual feed intake is an important trait relevant to feed efficiency in beef and dairy cattle but it is difficult to improve genetically because it is expensive to measure (Arthur *et al.* 2004). It is hoped that genomic selection using DNA markers might be used to achieve genetic improvement in RFI. Since the introduction of genomic selection (Meuwissen *et al.* 2001) there has been much research into the accuracy with which genomic estimated breeding values (GEBVs) predict true breeding values. The most common method to estimate the accuracy of GEBV has been to put aside a proportion of the population (a validation group) and not use them in the estimation of SNP effects. Then the estimated SNP effects are used to calculate GEBVs for the excluded animals which are then correlated with their phenotypic records. This correlation is the accuracy with

which the GEBVs predict new phenotypes. This method has several disadvantages. For instance, accuracies (r) or reliabilities (r^2) are not available for individual animals. When conventional BLUP is used to predict breeding values, the reliabilities of individual EBVs are calculated from the BLUP equations and it would be useful if this could also be done for GEBV but to date this approach is not well accepted. Theory and experimental results show that the reliability of GEBVs depends mostly on the precision of phenotypic data and number of genotyped animals in the reference population (VanRaden 2009). One way of increasing the number of individuals with phenotypes and genotypes is using a multi-breed reference population. However, the gain in accuracy from multi-breed reference populations has been found to be low, although a convincing explanation for this finding has not been offered. Three possible explanations are: 1. the effect of a quantitative trait locus (QTL) varies from breed to breed (*i.e.* breed x QTL interaction). This could be due to a true interaction between breed and the QTL or to an interaction between QTL and the way the trait was measured in different breeds (*e.g.* at different ages). 2. the linkage disequilibrium (LD) between the QTL and the SNPs that are assayed varies between breeds. 3. the across breed LD is low and limited to SNPs very close to the QTL so that there is limited information which can be transferred across breeds. The first two reasons result in a breed x SNP interaction. The LD between SNPs and QTL is only likely to be consistent across breeds for SNPs very close to the QTL and therefore we need very dense markers. In this research we have used around 700,000 SNPs which should be dense enough because LD phase is conserved across breeds at distances of 5 kb (deRoos *et al.* 2009). The aim of this research is to explain the accuracy of GEBV for RFI using a multi-breed reference population and to assess if using prediction error variances (PEVs) of GEBVs from the BLUP equations can correctly predict the accuracy.

MATERIAL AND METHODS

Cattle and RFI measurement. RFI records of 5,614 animals including 842 Holstein heifer, 2,009 beef cattle of Australia and 2,763 Canadian beef cattle were available for analysis. The Australian beef cattle included different breeds, 1,134 Angus, 217 Herford, 79 Murray Grey and 579 Shorthorn) and RFI data of Canadian beef consisted of 534 Angus, 384 Charolaise and 1,845 mixed synthetic breed (average breed compositions were formed by Angus (45.9%), Simmental (20.7%), Piedmontese (5%), Gelbvieh (4.2%), Charolais (2%) and Limousin (1.4%). The Holstein heifers were fed with cubed alfalfa *ad libitum* (Pryce *et al.* 2012) and the Angus cattle from Trangie Research Station were fed a pelleted diet *ad libitum* shortly after weaning. The other beef cattle used in this study were fed a feedlot diet at > 1 year of age. Residual feed intake phenotypes for the animals were obtained from 3 different studies (Australian dairy cattle: Pryce *et al.* 2012; Australian beef cattle: Bolormaa *et al.* 2013; Canadian beef cattle: Montanholi *et al.* 2009).

SNP data. The SNP marker data was from Illumina HD Bovine SNP chip, with 777,963 SNPs for Holstein heifers or imputed from lower density SNP chips (7K, 10K and 50K) to HD (800K) with BEAGLE (Browning and Browning 2009) for beef cattle. The genotypes passed quality control procedures including Illumina Genetrain (GC) score greater than 0.6 and rare minor allele frequencies higher than 0.5 % (Pryce *et al.* 2012). In order to construct genomic relationship matrix (GRM) for genomic evaluation (Yang *et al.*, 2010), common SNPs (606,096 SNPs) in the 3 datasets (Holstein heifers, Australian beef cattle and Canadian beef cattle) were used.

Statistical analysis. There were two types of GRM in the analyses: 1.using all estimated genomic relationships between all animals in the data and 2. where genomic relationships between animals of different breeds were set to zero to indicate the lack of relationship between animals of different breeds. A pedigree relationship matrix was also added to some of the models to see whether adding a polygenic term improved the log likelihood. The statistical model when the fixed effects and all three random terms were used in the analysis was:

(1) $\mathbf{y} = \mathbf{Xb} + \mathbf{Z_1u_1} + \mathbf{Z_2u_2} + \mathbf{a} + \mathbf{e}$

where, **y** is the vector of RFI records, **X** and $\mathbf{Z_{1,2}}$ are design matrixes relating phenotypes to their corresponding fixed effects and random effects, **b** is the vector of fixed effects including dataset (source of data), herd, feed management group prior to and on trial, contemporary group, cohort, month of birth, sex and age, $\mathbf{u_1}$ are SNP effects ~ $N(0,I\ \sigma^2_{SNP})$, $\mathbf{u_2}$ are SNP effects within breed ~ $N(0,I\ \sigma^2_{SNP*breed})$ and **a** are polygenic effects ~ $N(0,A\ \sigma^2_{polygenic})$. In order to fit this model, an equivalent model was used, that is:

(2) $\mathbf{y = Xb + g_1 + g_2 + a + e}$

where, $\mathbf{g_1=Z_1u_1}$ ~ $N(0, \mathbf{Z_1Z_1'\sigma^2_{SNP}})$, $\mathbf{g_2=Z_2u_2}$ ~ $N(0, \mathbf{Z_2Z_2'\sigma^2_{SNP*breed}})$ $\mathbf{Z_1Z_1'}$ is the GRM and $\mathbf{Z_2Z_2'}$ is the GRM within breed, that is all relationships between animals in different breeds have been set to zero. To test the significance of the $\mathbf{g_1}$ and $\mathbf{g_2}$ terms, the log of likelihood of the model was calculated using the full model and after dropping either $\mathbf{g_1}$ or $\mathbf{g_2}$ from the model. To test the significance of a change in log of likelihoods, two times the difference in log of likelihoods was compared to Chi squared with 1 degree of freedom. To find the best GRM within breed in the model, some breeds were treated as part of the one "super breed" in the analysis. Murray Grey and Australian Angus cattle were always grouped together and treated as one breed due to the small number of Murray Grey animals. Conversely, the Trangie Angus animals were treated as a separate breed to other Angus because RFI was measured at a younger age and using different feed at Trangie. In order to calculate the accuracies of GEBVs in a genotyped population without phenotypes, 5 subsets of the main population were generated. The animals of subsets were selected randomly but for each validation no animals with common sires were allowed to be present in both validation and reference groups. The phenotypes of each validation group were removed and after estimating GEBVs by BLUP, the correlation between GEBVs and phenotypes adjusted for fixed effects in the validation population was calculated which was divided by the square root of estimated heritability to form the empirical accuracy of estimated breeding values in each validation population.

(3) $\text{Empirical Accuracy} = r_{\text{GEBVs, Corrected\_Phenotypes}} / \sqrt{h^2_{\text{Pedigree}}}$

The empirical accuracies were compared to theoretical accuracies calculated without a validation population directly from the mixed model equations. The empirical accuracies were correlations within breed and to be consistent the theoretical accuracies were also calculated within breed. To do this, the prediction error variances for the animal effects were calculated from the mixed model equations in the standard way and used to predict the theoretical accuracy of GEBVs in the validation population.

RESULTS AND DISCUSSION

After fitting a model with an overall effect of the SNPs ($\mathbf{g_1}$) instead of the polygenic term (**a**), the log of likelihood improved significantly ($P < 0.01$) and adding SNP x breed ($\mathbf{g_2}$) further improved log of likelihood ($P < 0.01$). The results indicated that keeping the relationship between Holstein and Trangie Angus while setting the relationship between them and non-Trangie Angus and other breeds to zero (model 6) improved the log of likelihood ($P < 0.01$). However, model 6 was not significantly better than model 7 in which only Trangie Angus and Holstein relationships were kept and the relationships between different breeds were set to zero (Table 1). One of the main differences of Trangie cattle compared with the other beef animals in the experiment was their age at RFI measurement time, it seems that the effect of age is more important than the effect of breed in RFI evaluation because by treating Trangie cattle and Holstein heifers as a super breed a better log of likelihood was achieved. Therefore, the best model was reached by applying 3 relationship matrixes; an overall GRM, super breed GRM when keeping relationships between Trangie beef cattle and Holstein heifers and setting all other breed by another breed relationships to zero and pedigree relationship matrix. In this model (model 9) the genetic variance was almost

entirely explained by the overall GRM (SNP effect) and within breed GRM (SNP x breed effect). The accuracies of GEBVs were also estimated with this model. The average accuracy for RFI in 5 validations was 0.31 which was near to expected from the BLUP equations (0.34). It seems that an across breed reference population can be used provided the animals are measured for RFI at about the same age and on a similar diet. However, there is only a small increase in accuracy by adding animals of another breed because the relationships between animals in different breeds are all low. The BLUP equations correctly predict this limited increase in accuracy (about 2%).

Table1. Application of different models to find the best fitted one (highest log of likelihood)

Model	Log of Likelihood	σ^2_{SNP}	$\sigma^2_{SNP*Breed}$	$\sigma^2_{polygenic}$	σ^2_e	h^2
1. $\mathbf{Xb + g_1}$	-2853.50	0.3010	-	-	0.7024	0.3000
2. $\mathbf{Xb + g_{2\_superbreed1}}$	-2853.95	-	0.3280	-	0.6730	0.3277
3. $\mathbf{Xb + g_{2\_superbreed2}}$	-2850.83	-	0.3197	-	0.6832	0.3188
4. $\mathbf{Xb + g_{2\_superbreed3}}$	-2852.61	-	0.3312	-	0.6702	0.3308
5. $\mathbf{Xb + a}$	-2901.49	-	-	0.3023	0.7022	0.3010
6. $\mathbf{Xb + g_1 + g_{2\_superbreed2}}$	-2849.18	0.1237	0.1959	-	0.6832	0.3187
7. $\mathbf{Xb + g_1 + g_{2\_superbreed3}}$	-2847.93	0.1537	0.1790	-	0.6693	0.3320
8. $\mathbf{Xb + g_1 + g_{2\_superbreed2} + a}$	-2848.21	0.1246	0.1697	0.0522	0.6569	0.3453
9. $\mathbf{Xb + g_1 + g_{2\_superbreed3} + a}$	-2847.33	0.1523	0.1588	0.0421	0.6495	0.3522

a=pedigree relationship matrix

$\mathbf{g_1}$=(DD+TT+NT+MG+HH+SS+AA+CC+XX); $\mathbf{g_{2\_superbreed1}}$=DD,TT,(NT+MG),HH,SS,AA,CC,XX

$\mathbf{g_{2\_superbreed2}}$=(DD+TT),(NT+MG+HH+SS+AA+CC+XX);

$\mathbf{g_{2\_superbreed3}}$=(DD+TT),(NT+MG),HH,SS,AA,CC,XX

* In each model the relationships between the breeds in the same brackets were kept while relationships of the breed with another breed were assigned to zero. (DD=Holstein heifers; Australian beef cattle: NT=Non-Trangie Angus, TT=Trangie Angus, MG=Murray Grey HH=Herford, SS=Shorthorn; Canadian beef cattle: AA=Angus, CC=Charolaise, XX= Mixed synthetic breed)

CONCLUSIONS

According to the best fitting model, it seems the SNP effects were not significantly different between Holstein and Trangie cattle, fed a similar diet and measured at a similar age. However, the SNP effects probably differed between groups fed different diets and measured at different ages. So, it is important to consider feed and age at measurement time in RFI evaluations.

REFERENCES

Arthur P.F., Archer J.A. and Herd R.M. (2004) *Aust. J. Exp. Agric.* **44**:361.

Bolormaa S, Pryce J.E., Kemper K., *et al* (2013) *J. Anim. Sci.* Published online before print May 8.

Browning, B.L. and Browning S.R. (2009) Am. J. Hum. *Genet.* **84**:210.

deRoos A.P.W., Hayes B.J. and Goddard M.E. (2009) *Genet.* **183**:1545.

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

Montanholi Y.R., Swanson K.C., Palme R., Schenkel F.S., McBride B.W., Lu D., and Miller, S.P. (2009) *Anim.* **4**:692.

Pryce J.E., Arias J., Bowman P.J., Davis S.R., Macdonald K.A., Waghorn G.C., Wales W.J., Williams Y.J., Spelman R.J. and Hayes B.J. (2012) *J. Dairy Sci.* **95**:2108.

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.

Yang J., Benyamin B., McEvoy B.P., Gordon S., Henders A., Nyholt D.R., Madden P.A., Heath A.C., Martin N.G., Montgomery G.W., Goddard M.E. and Visscher P.M. (2010) *Nat. Genet.* **42**:565.

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:303-306

GENETIC SOLUTIONS TO IMPROVE RESOURCE EFFICIENCY IN DAIRY CATTLE

Y. de Haas[1], J.E. Pryce[2], J. Dijkstra[3], E. Wall[4] and R.F. Veerkamp[1]

[1]Animal Breeding and Genomics Centre, Wageningen UR Livestock Research, P.O. Box 65, NL-8200 AB Lelystad;
[2]Biosciences Research Division, Department of Primary Industries Victoria, 5 Ring Road, Bundoora, VIC 3083, Australia;
[3]Animal Nutrition Group, Wageningen University, P.O. Box 338, NL-6700 AH Wageningen;
[4]Scottish Rural College, Animal & Veterinary Sciences Group, Roslin Institute Building, Easter Bush, Midlothian EH25 9RG, UK.

SUMMARY

Examples of traits related to feed resource efficiency are residual feed intake (RFI) and methane (CH_4) emission. In an experimental dataset of 588 heifers, we showed that it is possible to decrease CH_4 emission (predicted on dry matter intake (DMI) and ration) by selecting more efficient cows (genetic correlation of 0.3). Feed efficiency phenotypes are difficult and expensive to measure on a large scale, but genomic selection is a promising tool to make progress in breeding resource efficient cows, since it relaxes the need for information on performance of all animals. Using genomic selection, a reduction in predicted CH_4 in the order of 15% in 10 years is theoretically possible. To double this genetic gain, a large reference population is needed. Therefore, an international collaboration between 9 countries in Europe, US and Australiasia is set up to assemble data on >6,000 cows with high quality phenotypes and genotypes. The next step is to predict the genomic breeding values with this extended dataset, and report the accuracies. This way, a combined approach, including feeding, management and genetic selection, can be set up, which is likely to be the best approach to successfully improve feed resource efficiency.

OVERALL INTRODUCTION

Climate change is a growing international concern and it is well established that the release of greenhouse gases (GHG) is a contributing factor. The general aim of the Kyoto protocol is to reduce GHG emissions by 20% by the year 2020 relative to 1990 levels. The global livestock sector, particularly ruminants, contributes approximately 18% of total anthropogenic GHG emissions (Steinfeld *et al.* 2006). One way to reduce the environmental impact of dairy cattle is to improve their resource efficiency. Examples of traits related to feed resource efficiency are dry matter intake (DMI), residual feed intake (RFI), and methane (CH_4) emission. This paper provides genetic parameter estimates for feed resource efficiency traits, and examines the value of creating an international data set for these traits.

GENETIC PARAMETERS FOR FEED RESOURCE EFFICIENCY TRAITS

Introduction. Nutritional and microbial opportunities to reduce CH_4 emissions have been extensively researched, but there is little knowledge regarding the use of natural variation to breed for animals with lower CH_4 yield (Wall *et al.* 2010). Measuring CH_4 emission rates directly from animals is difficult and hinders direct selection on reduced CH_4 emission. However, improvements can be made through selection on associated traits (e.g. RFI (Verbyla *et al.* 2010)), or through selection on CH_4 predicted from feed intake and diet composition (de Haas *et al.* 2011).

Aim. The objective of this study was to quantify phenotypic and genetic variation in RFI and predicted CH_4 emission (PME), and to examine the potential use of genomic selection to facilitate the inclusion of resource efficiency phenotypes in selection programmes (de Haas *et al.* 2011).

Material and methods. Data from previous experiments were used, and records on daily DMI, weekly live weights and weekly milk productions were available from 588 heifers (Veerkamp *et al.* 2000). RFI (MJ/d) is the difference between net energy intake and calculated energy requirements for maintenance as a function of live weight and for fat and protein corrected milk production. PME (g/d) is 6% of gross energy intake (method of International Panel on Climate Change (IPCC)) corrected for energy content of methane (55.65 kJ/g). All heifers were genotyped using the Illumina 50K SNP panel (54,001 SNP in total; Illumina, San Diego, CA). Genetic parameters were determined using a random regression sire-maternal grandsire model in ASREML (Gilmour *et al.* 2009). Effects of SNPs were estimated using Bayesian stochastic search variable selection (SSVS; (George and McCulloch 1993)). Genomic breeding values were predicted for these heifers using a model that included the genotypic information. A polygenic model was used to estimate breeding values using only pedigree information. A 10 fold cross-validation approach was employed to assess the accuracies of the two sets of predicted breeding values by correlating them with the phenotypes.

Results and discussion. The estimated heritabilities for PME and RFI were 0.35, and 0.40, respectively (Table 1). Both heritability estimates fit well in the range recently reviewed by Berry and Crowley (2013). PME has not been analysed before, but it can well be compared with DMI.

Table 1. The estimated heritability (on diagonal), phenotypic (above diagonal) and genetic correlation (below diagonal) for residual feed intake (RFI) and predicted methane emission (PME). The corresponding standard errors are shown in parentheses

	RFI	PME
Residual feed intake (RFI)	0.40 (0.11)	0.72 (0.08)
Predicted methane emission (PME)	0.32 (0.06)	0.35 (0.12)

The positive genetic correlation between RFI and PME indicated that cows with lower RFI have lower PME as well. Hence, it seems possible to decrease methane production of a cow by selecting more efficient cows, and the genetic variation suggests that reductions of the order of 11 to 26% in 10 years are theoretically possible, and in a genomic selection program even higher (de Haas *et al.* 2011). For both feed resource efficiency phenotypes (RFI and PME) the genomic model produced breeding values with reliability double, or even triple, that of the breeding values produced by the polygenic model (Table 2). No other studies have published accuracies of genomic predictions of these new traits, but achieved accuracies were lower than theoretically expected accuracies (Daetwyler *et al.* 2010).

Table 2. Reliabilities of estimated breeding values (EBV) based on pedigree information only, and direct genomic values (DGV) based on both pedigree and marker (SNP) information for residual feed intake (RFI) and predicted enteric methane emission (PME)

	RFI	PME
Pedigree	0.14	0.04
Pedigree + SNP	0.27	0.14

ADDED VALUE OF INTERNATIONAL COLLABORATOIN – A FIRST ATTEMPT

Introduction. A number of countries have started to record DMI data, but not enough records are available to get accurate breeding values for this trait to be used in their national breeding programme. One way to obtain estimated breeding values (EBVs) in a population is to use

genomic selection, where phenotypes, e.g., DMI, are measured in a subset of the population and genomic predictions are calculated for other animals that have genotypes, but no phenotypes. While this approach is appealing, allowing selection for improved efficiency, the size of the reference populations from which the genomic prediction equations are derived are currently too small within each country to achieve satisfactory levels of accuracy of genomic breeding values (Verbyla *et al.* 2010). One way to increase the accuracy of the genomic prediction is to combine datasets from multiple populations. Challenges when combining phenotypes from several countries include genotype by environment (GxE) interactions and differences in trait definitions. A multi-trait model can handle traits that are measured in different environments as separate traits, and therefore treat both the GxE interaction and differences in trait definitions properly.

Aim. The aim of this study was to estimate the accuracy of genomic prediction for DMI, when analysed together in a single-trait run, or in a multi-trait run, using both Australian data on growing heifers and European data on lactating heifers (de Haas *et al.* 2012).

Material and methods. In total, DMI records were available on 1801 animals; 843 Australian (AU) growing heifers with records on DMI measured over ±70 days at 200 days of age (Williams *et al.* 2011, Pryce *et al.* 2012), 359 Scottish (UK) and 588 Dutch (NL) lactating heifers with records on DMI during the first 100 days in milk (Banos *et al.* 2012, Veerkamp *et al.* 2012). The genotypes used in this study were obtained from the Illumina Bovine 50k chip. The AU, UK and NL genomic data were matched using the SNP name. Quality controls were applied by carefully comparing the genotypes of 40 bulls that were available in each dataset. This resulted in a total of 30,949 SNPs being used in the analyses. Genomic predictions were estimated with genomic REML (G-REML), using ASReml (Gilmour *et al.* 2009). The accuracy of genomic prediction was evaluated in 11 validation sets. The reference set (where animals had both DMI phenotypes and genotypes) were either within AU or Europe (UK and NL), or with a multi-country reference set consisting of all data except the validation set.

Results and discussion. When DMI for each country was treated as the same trait (i.e., univariate analysis), using a multi-country reference set (uni-multi) increased the accuracy of genomic prediction for DMI for UK, compared to the accuracy achieved with a univariate analysis with the national reference set. The accuracy did, however, not increase for AU and NL (Table 3).

Table 3. The average of the approximated accuracy (and corresponding standard error) of genomic prediction of dry matter intake (DMI), calculated as the correlation between genomic breeding value (GEBV) and the true breeding value (TBV), estimated in a univariate, bivariate or trivariate run between Australia (AU), Europe (EU), United Kingdom(UK) and the Netherlands (NL), where "uni within" refers to the current situation with a national reference set. In all other analyses, a multi-country reference set was taken consisting of all data except the validation set.

Country	uni within	uni multi	bi: AU-EU	tri: AU-UK-NL
AU	0.378 (0.027)	0.336 (0.046)	0.388 (0.041)	0.389 (0.042)
EU	0.313 (0.050)	0.323 (0.051)	0.322 (0.048)	0.330 (0.049)
UK	0.301 (0.042)	0.333 (0.059)	0.315 (0.048)	0.332 (0.032)
NL	0.326 (0.098)	0.312 (0.093)	0.329 (0.092)	0.328 (0.094)

Extending the model to a bivariate (AU-EU) or trivariate (AU-UK-NL) model increased the accuracy of genomic prediction for DMI in all countries (de Haas *et al.* 2012). Highest accuracies were estimated for all countries when data was analysed with a trivariate model, with increases of up to 5.5%.

This first attempt has shown that it is worthwhile setting up an international collaboration and sharing data, but the increase in accuracy was not enough to get accurate breeding values for this trait to be used in their national breeding programme. Therefore, an initiative has started to combine DMI data from 9 countries in Europe, US and Australiasia. Pooling DMI data across countries can establish if this is a viable way to estimate genomic prediction equations that give breeding values with sufficient accuracy, so that these can be used for demonstration by the collaborators in the project. First results of this collaboration are expected late 2013.

OVERALL CONCLUSIONS

Examples of traits related to feed resource efficiency are residual feed intake (RFI) and methane (CH_4) emission. Our studies on national data have shown that genetic solutions to improve these feed resource efficiency traits is possible. However, international collaboration to assemble data on more cows will improve the accuracy and genetic gain. A combined approach, including feeding, management and genetic selection, can then also be set up, which is likely to be the best approach to successfully improve resource efficiency.

REFERENCES

Banos G., Coffey M.P., Veerkamp R.F., Berry D.P. and Wall E. (2012) *Animal* **6:** 1040.

Berry D.P. and Crowley J.J. (2013) *J. Anim. Sci.* **91:** 1594.

Daetwyler H.D., Pong-Wong R., Villanueva B. and Woolliams, J.A. (2010) *Genetics* **185:** 1021.
de Haas Y., Calus M.P.L., Veerkamp R.F., Wall E., Coffey M.P., Daetwyler H.D., Hayes B.J. and Pryce J.E. (2012) *J. Dairy Sci.* **95:** 6103.

de Haas Y., Windig J.J., Calus M.P.L., Dijkstra J., de Haan M., Bannink A. and Veerkamp R.F. (2011) *J. Dairy Sci.* **94:** 6122.

George E.I. and McCulloch R.E. (1993) *J. Am. Stat. Ass.* **88:** 881.

Gilmour A.R., Cullis B.R., Welham S.J. and Thompson R. (2009) 'ASREML. Program user manual' NSW Agriculture, Orange Agric. Inst., Forest Road, Orange, NSW, 2800, Australia.

Pryce J.E., Arias J., Bowman P.J., Davis S.R., MacDonald K.A., Waghorn G.C., Wales W.J., Williams Y.J., Spelman R.J. and Hayes B.J. (2012) *J. Dairy Sci.* **95:** 2108.

Steinfeld H., Gerber P., Wassenaar T., Castel V., Rosales M. and de Haan C. (2006) 'Livestock's Long Shadow: Environmental Issues and Options' Food and Agric. Org. of the United Nations.

Veerkamp R.F., Oldenbroek J.K., van der Gaast H.J. and van der Werf J.H.J. (2000) *J. Dairy Sci.* **83:** 577

Veerkamp R.F., Coffey M.P., Berry D.P., De Haas Y., Strandberg E., Bovenhuis H., Calus M.P.L. and Wall E. (2012) *Animal* **6:** 1738.

Verbyla K.L., Calus M.P.L., Mulder H.A., de Haas Y. and Veerkamp R.F. (2010) *J. Dairy Sci.* **93:** 2757.

Wall E., Simm G. and Moran D. (2010) *Animal* **4:** 366.

Williams Y.J., Pryce J.E., Grainger C., Wales W.J., Linden N., Porker M. and Hayes B.J. (2011) *J. Dairy Sci.* **94:** 4715.

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:307-310

FEMALE, MALE AND GENOMIC MEASURES FOR USE IN GENETIC SELECTION TO IMPROVE LIFETIME WEANING RATE OF BRAHMAN CATTLE

S.A. Barwick[1], D.J. Johnston[1], R.G. Holroyd[2], J.R.W. Walkley[3] and H.M. Burrow[4]

Cooperative Research Centre for Beef Genetic Technologies
[1]Animal Genetics & Breeding Unit (AGBU*), Univ. of New England, Armidale, NSW, 2351
[2]Queensland Department of Agriculture, Fisheries & Forestry, Rockhampton, Qld, 4701
[3]Livestock Systems Alliance, University of Adelaide, Roseworthy, SA, 5371
[4]CSIRO Livestock Industries, Rockhampton, Qld, 4702

SUMMARY

Female and male measures with potential to be practical, early-in-life genetic indicators of female reproduction in Brahmans were chosen from earlier reports and compared, along with genomic measures, for multi-trait use to improve Brahman lifetime annual weaning rate (*LAWR*). Results suggested substantial genetic gains in *LAWR* may be possible in 10 years using these measures, but need confirming in other data. Female hip height and coat score and male preputial eversion and liveweight were measures that could warrant wider recording for *LAWR* improvement. Recording of pregnancy test outcomes from matings 1 and 2 should also be encouraged. A genomic EBV in combination with other measures added to the gain in *LAWR*, but needed an accuracy approaching 60% to be the most important contributor to gains in the combinations of measures studied.

INTRODUCTION

Low reproduction limits productivity in Brahmans, a major beef breed of tropical environments including in Australia (Johnston *et al.* 2013a). In a larger study, Barwick *et al.* (2013) examined the potential of numerous early-in-life measures for multi-trait use as selection criteria to improve female reproductive performance of Brahmans. This report focuses on lifetime annual weaning rate (*LAWR*) and on the multi-trait use of only the potentially most practical measures for recording in harsh tropical environments. The basis for comparing measures was the estimated genetic gain in *LAWR* from selection. *LAWR* aligns with the weaning rate trait of beef cattle breeding objectives (Barwick and Henzell 2005). Estimates of the changes expected in the individual criteria contributing to the gains in *LAWR* are also presented.

METHODS

Definitions. *LAWR* and the female and male measures studied were from an experiment with Brahmans in northern Australia. Environments and management were described by Barwick *et al.* (2009), Corbet *et al.* (2013) and Johnston *et al.* (2013a). Females were by 54 sires and male progeny of the females by a further 60 sires. Females calved first at 3 years and were culled if they failed to wean a calf in any two consecutive years. *LAWR* was the average weaning rate of cows based on the number of annual mating opportunities they experienced over 6 possible matings.

The female and male measures studied were chosen for their potential to be both indirect genetic criteria for *LAWR*, based on earlier bi-variate analyses, and practical to be recorded by industry. Female adaptive measures were coat colour and navel score at 9 m of age, coat score (*COAT*) at 12 m and rectal temperature at 13 m, from Prayaga *et al.* (2009) and Wolcott *et al.* (2013a); female 18- and 24-m measures were liveweight and hip height (*HH18*, *HH24*), scanned fat depth (*SFAT18*) and eye muscle area at 18 m, from Barwick *et al.* (2009) and Wolcott *et al.*

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

(2013b); mating 1 and 2 measures were pregnancy-test outcomes from mating 1 (*PREG1*) and 2 (*PREG2*) at 28 and 40 m, days to 1st calving (*DC1*) at 37 m, and success at both weaning a calf from mating 1 and being pregnant from mating 2 (*W1P2*) at 40 m, from Johnston *et al.* (2013a, 2013b); male non-semen measures were flight time at 6 m, rectal temperature at 12 m, scrotal size at 12 and 18 m, liveweight (LWT15), hip height, scanned fat depth, scanned eye muscle area and body condition score at 15 m, sheath score and preputial eversion (EV) at 18 m, from Corbet *et al.* (2013) and Johnston *et al.* (2013b); and semen measures were sperm mass activity score, sperm progressive motility (MOT) and percent normal sperm at 18 m, from Corbet *et al.* (2013) and Johnston *et al.* (2013b). Genomic measures examined were genomic EBVs for *LAWR* of 30, 40 or 60% accuracy (GEN30, GEN40, GEN60) based on genotyping of males, and genomic EBVs for selected other measures of 40% accuracy. Female measures are shown in italics throughout this report; male measures are not italicised.

Evaluation of measures. Genetic gains were estimated assuming index selection across the four pathways of Rendel and Robertson (1950). Step-down analyses systematically eliminated measures that contributed least to the estimated gain in *LAWR* while retaining those contributing 5% or more of the gain. Analyses were conducted in blocks, with retained measures carried forward to be considered with other measures. Inbreeding and other contributors to long-term response were not considered. Estimates were adjusted to gains per 10 yrs for presentation and should be viewed as approximations of the selection gains in Brahmans that may be possible.

Parameters. These were based on bi-variate estimates from Barwick *et al.* (2009), Johnston *et al.* (2009), Prayaga *et al.* (2009), Corbet *et al.* (2013), Johnston *et al.* (2013a, 2013b) and Wolcott *et al.* (2013a, 2013b). For positive definite matrices it was necessary to reduce genetic correlations with *LAWR* to approximately 70% of their bi-variate values. This reduction was applied to all measures except genomic EBVs, which were retained at their assumed accuracies. Derivations of genomic EBV variances and genetic and phenotypic correlations involving genomic EBVs are described by Barwick *et al.* (2013). Zero environmental correlations were assumed, including between female and male measures. The genetic standard deviation for *LAWR* was 0.0877 (Johnston *et al.* 2013a).

Selection accuracies and correlated responses. Calculation of accuracies and correlated responses used MTIndex of J. van der Werf and assumed animals had a parent record, 20 half-sib records and their own record (depending whether selection was of sires or dams) for all measures.

Selection intensities. The selection described was for a Brahman population of 13000 cows and assumed a concerted selection effort could be made across the breed (Barwick *et al.* 2013). Sires to breed sires were the top 20 of 2000 (1%; i = 2.665), sires to breed dams the top 135 of 2000 (6.75%; i = 1.9345), dams to breed sires the top 800 of 8000 (10%; i = 1.755), and dams to breed dams the top 3800 of 4000 (95%; i = 0.1086). These levels were used to estimate genetic gains for comparing all combinations of measures. For the identified best combinations, estimated gains were also calculated assuming no selection of dams.

Generation intervals. The age structure assumed was derived from Brahman industry data and included cow calving age groups of from 3 to 13 years. Generation interval (*L)* was the average age of selected parents at the birth of progeny. For direct selection on *LAWR* (requiring outcomes from 6 matings) *L's* for breeding sires and dams were 10.88 and 11.32 years. These were respectively 6.49 and 6.93 years for selection on *PREG2* or *W1P2*, 6.07 and 6.07 years for selection on *PREG1* or *DC1*, and 5.70 and 5.26 years for selection on any of the other measures considered. For selection on a combination of measures, *L's* were decided by the last available measure.

RESULTS AND DISCUSSION

Results suggested substantial gains in *LAWR* may be possible in Brahmans from multi-trait selection on a best combination of practical, early-in-life measures. The estimated gains in *LAWR* in 10 years from this were 8 to 12% from sire selection and 12 to 15% (i.e. 0.12 to 0.15 calves weaned/cow) from selection of sires and dams (Table 1). Gains using combinations of measures were greater than using individual measures and greater than for direct selection. Gains from selection on individual genomic EBVs for *LAWR* were 0.066, 0.089 and 0.133 calves weaned/cow for GEN30, GEN40 and GEN60. Gains using an individual genomic EBV for a correlated trait were considerably less; and the gain using an individual fatness measure (*SFAT18*) was small.

Female *HH24* (and *HH18*) and *COAT* and male EV and LWT15 were among the most important measures of combinations (Table 1) and may warrant wider industry recording. Greater *HH24* and lower LWT15, less preputial eversion and a sleeker coat were preferred for *LAWR*. *COAT* was also most important for reducing female age at puberty (Barwick *et al.* 2013).The estimated changes in individual criteria and in *LAWR* (Table 1) for measurement combinations 1) and 5), respectively, represented rates of gain of 0.08, 0.10, 0.03, 0.09 and 0.14 genetic standard deviations per year for *HH24*, EV, LWT15, *COAT* and *LAWR*, and 0.24, 0.12, 0.09 and 0.18

Table 1. Combinations of female, male and genomic measures giving the greatest estimated genetic gain in lifetime annual weaning rate (*LAWR*, calves weaned/cow) in Brahmans from selection of sires and dams, for differing categories of measures available[1,2]. Also shown are the estimated 10 year gains in *LAWR* and the associated changes in each of the measures

Measures	Estimated change	Measures	Estimated change	Measures	Estimated change
1) Female + male non-semen		*2) Female + male*		*3) Female + male* + GEN30	
HH24 (cm)	2.1	*HH24* (cm)	2.3	*HH24* (cm)	2.7
EV (mm)	-11.2	EV (mm)	-10.2	EV (mm)	-9.1
LWT15 (kg)	-4.0	MOT (%)	12.6	MOT (%)	11.3
COAT (score)	-1.1	LWT15 (kg)	-4.9	GEN30[3]	0.035
		COAT (score)	-1.2	LWT15 (kg)	-4.9
LAWR:	**0.123**	*LAWR:*	**0.132**	*LAWR:*	**0.136**
4) Female + male + GEN40		*5) Female + male* + GEN60			
HH24 (cm)	2.8	GEN60[3]	0.124		
GEN40[3]	0.058	*HH24* (cm)	3.3		
EV (mm)	-9.3	*COAT* (score)	-1.1		
MOT (%)	11.6				
LWT15 (kg)	-4.9				
LAWR:	**0.143**	*LAWR:*	**0.154**		

[1]See earlier text for details of measures in each category. 'Female' & 'male' include relevant sub-categories.
[2]Combinations are the end result of step-down analyses of the estimated genetic gain. All lists showmeasures in their order of greatest importance to genetic gain. Female measures are italicised, male measures are not.
[3]GEN30, GEN40 and GEN60 are genomic EBVs for *LAWR* of 30, 40 and 60% accuracy, respectively.

genetic standard deviations per year for GEN60, *HH24*, *COAT* and *LAWR*. Semen MOT added to gains when it was available. Adding *PREG1* and *PREG2* increased accuracies but not gains, as *L's* were increased. Industry recording of *PREG1* and *PREG2* should be encouraged; pregnancy testing is commonly practiced and the marginal cost of the recording would be small. A genomic

EBV for *LAWR* added to gains especially at an accuracy of 40% but needed an accuracy approaching 60% to be the most important measure of combinations (Table 1).

The results need confirming since they depend on many estimates and step-wise procedures are susceptible to bias. The results may apply only to Brahmans or perhaps to *Bos indicus*. The genetic relationships utilised for hip height and liveweight with *LAWR* in the Brahman (Wolcott *et al.* 2013b), in particular, need confirming. Measures were also considered separately between the sexes, which meant there was little opportunity for a measure to be important in both sexes. Where this separation can be relaxed, the gains may be greater.

CONCLUSIONS

Results need confirming but suggested substantial genetic gains in *LAWR* may be possible in Brahmans from selection on combinations of practical, early-in-life measures. Female *HH24* (or *HH18*) and *COAT* in females and male EV and LWT15 are measures that could warrant wider recording for this purpose. Recording of *PREG1* and *PREG2* should also be encouraged. A genomic EBV in combination with other measures added to gains, but would need to have an accuracy approaching 60% to be the most important individual contributor to gains in the combinations of measures studied.

ACKNOWLEDGEMENTS

We thank the many participants in the original experiment, and the Commonwealth CRC program, CRC core-partner organisations, co-operating breeders and Meat & Livestock Australia for financial and in-kind support.

REFERENCES

Barwick S.A. and Henzell A.L. (2005) *Aust. J. Exp. Agric.* **45**: 923.

Barwick S.A., Johnston D.J., Burrow H.M., Holroyd R.G., Fordyce G., Wolcott M.L., Sim W.D. and Sullivan M.T. (2009) *Anim. Prod. Sci.* **49**: 367.

Barwick S.A., Johnston D.J., Holroyd R.G., Walkley J.R.W. and Burrow H.M. (2013) *Anim. Prod. Sci.* (submitted)

Corbet N.J., Burns B.M., Johnston D.J., Wolcott M.L., Corbet D.H., Venus B.K., Li Y., McGowan M.R. and Holroyd R.G. (2013) *Anim. Prod. Sci.* **53**: 101.

Johnston D.J., Barwick S.A., Corbet N.J., Fordyce G., Holroyd R.G., Williams P.J. and Burrow H.M. (2009) *Anim. Prod. Sci.* **49**: 399.

Johnston D.J., Barwick S.A., Fordyce G., Holroyd R.G., Williams P.J., Corbet N.J. and Grant T. (2013a) *Anim. Prod. Sci.* (in press)

Johnston D.J., Corbet N.J., Barwick S.A., Wolcott M.L. and Holroyd R.G. (2013b) *Anim. Prod. Sci.* (in press)

Prayaga K.C., Corbet N.J., Johnston D.J., Wolcott M.L., Fordyce G. and Burrow H.M. (2009) *Anim. Prod. Sci.* **49**: 413.

Rendel J.M. and Robertson A. (1950) *J. Genet.* **50**: 1.

Wolcott M.L., Johnston D.J. and Barwick S.A. (2013a) *Anim. Prod. Sci.* (in press)

Wolcott M.L., Johnston D.J., Barwick S.A., Corbet N.J. and Burrow H.M. (2013b) *Anim. Prod. Sci.* (submitted)

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:311-314

LIVEWEIGHT LOSS IN ADULT EWES IS AFFECTED BY THEIR SIRES BREEDING VALUES FOR FAT AND MUSCLE

S.E. Blumer[1,2], G.E. Gardner[1,2], M.B. Ferguson[3] and A.N. Thompson[1,2,4]

[1]CRC for Sheep Industry Innovation, Homestead Building, UNE, Armidale 2351
[2]School of Veterinary and Biochemical Sciences, Murdoch University, Murdoch 6150.
[3]New Zealand Merino, Level 2, 114 Wrights Road, Christchurch 8024.
[4]Department of Food and Agriculture Western Australia, South Perth 6151.

SUMMARY

Ewes that lose more weight at times of nutritional pressure may decrease farm profitability through reduced production but also through reduced stocking rates, increased supplementary feeding costs and labour. Liveweight profiles were derived from the splined liveweight data of adult ewes from the Sheep CRC Information Nucleus Flock and liveweight loss was analysed. This paper reports on the response of liveweight loss to sire breeding values for fat and muscle in two contrasting environments. Overall liveweight loss was affected by significant differences between site, sire breed, ewe age, and previous and current reproductive performance. Sire breeding values for fat were significant, and interacted with site. There was a reduction in liveweight loss as sire breeding values for fat increased at Katanning in WA but an increase in liveweight loss at Kirby in NSW. Sire breeding values for muscle were also significant and different for each site, with the effects being opposite to fat at each site. These results suggest that selection against fat or selection for increased muscling may compromise the ability of ewes to maintain weight during summer and autumn in dry Mediterranean climates, however this may not be applicable for all environments.

INTRODUCTION

The storage and mobilisation of fat is an important mechanism for all animals to cope with fluctuating environments. Fat is stored during favourable times and then mobilised to provide energy for fundamental functions when requirements exceed supply, such as during periods of limited nutrition or during late pregnancy and lactation. Adams *et al.* (2002) found that when grazing dry feed, a strain of sheep with a greater proportion of fat tissue lost less weight than those with smaller fat reserves and that liveweight change was correlated with the change in weight of fat tissue. Conversely the opposite was true on short green feed, with the leaner sheep losing less weight and this was associated with greater intake of green feed. Hopkins *et al.* (2007) showed that in crossbred lambs the amount of fat stored in the carcass was correlated with the Australian Sheep Breeding Values (ASBVs) for fat. We therefore hypothesise that when liveweight loss occurs on dry feed, adult ewes from sires with higher breeding values for fat will lose less weight, whereas when liveweight loss occurs on green feed the response to fat will be the opposite.

MATERIALS AND METHODS

The Information Nucleus Flock was comprised of eight flocks located at different sites across Australia, and a description is provided by van der Werf *et al.* 2010. We analysed data from 2060 Merino and 712 Border Leicester x M ewes born in 2007, 2008 and 2009. Ewes of both genotypes were run under the same conditions at all times throughout the year, although separated during joining and lambing. Ewes at each site were managed by adjusting grazing pressure and altering supplementary feeding according to the recommendations developed for Merinos (Young *et al.* 2011), although actual liveweight losses were greater than recommended

especially at the Katanning site. Liveweight data from 2009, 2010 and 2011 when ewes were two, three and four years old were used and repeated measurements on ewes meant that 5216 annual records were analysed. Of the 83 sires used, 27 Border Leicester and 43 Merino sires were common between the two sites. In this paper, we report predictions from sites at Katanning in WA and Kirby in NSW.

Ewes were weighed on average 5.8 times per year. Liveweights were corrected for a) wool weight, calculated from greasy fleece weights and assuming constant wool growth rates during the year; and b) conceptus weight (Freer *et al.* 1997). The liveweight profile for each ewe over 12 months following each weaning period was produced in Genstat using a random coefficient regression model including a cubic polynomial for time. The model used was:
Liveweight = μ + day + ewe + ewe.day + spline(day) + ewe.spline(day).
Day was included as a fixed effect and all other terms were included as random effects. A covariance between the ewe intercept (ewe) and slope (ewe.day) was also included. Data was analysed in blocks (sire breed x site x year of birth x year).

The splined liveweight profile was used to derive the average liveweight, minimum, maximum and range in liveweight during each year. Liveweight loss (maximum to subsequent minimum) was analysed using SAS. Fixed effects in the base model were site (Kirby, Katanning), sire breed (BL, M), year (2009, 2010, 2011), age (2, 3, 4), birth type and rearing type of the ewe, and birth and rearing type of lambs raised by the ewe in the previous year and the current year. Year delineated the year in which each annual block of splined weight measurements commenced. Age described maturity and differentiated between parities. Ewe identification was included as a random term to account for repeated measures of ewes across years, and the sire random term allowed measures for each ewe to be utilised as replicates for sires. In a separate analysis ASBVs of the ewes sire for muscling (PEMD) and fat (PFAT) were included as covariates. Due to the significant correlations between these ASBVs (0.53 for BL sires and 0.71 for M sires) the breeding values were also tested individually to confirm their effects. First and second order interactions were tested and removed in a stepwise manner if non-significant ($P>0.05$).

RESULTS

Liveweight loss at Katanning occurred primarily between weaning and mid pregnancy. This period coincides with summer and early autumn, and limited availability of dry pasture and crop stubbles. At Kirby, liveweight loss occurred predominantly between joining and mid pregnancy which coincides with winter, and limited availability of green pasture.

The average liveweights of ewes from the different sites were 53.9kg ±0.6 at Katanning and 47.1kg ±0.6 at Kirby ($P<0.001$; Fig. 1). On average ewes from Border Leicester sires were heavier than ewes from Merino sires (54.1kg vs. 46.9kg, $P<0.001$). For both sire breeds, liveweight also increased with age, with three year olds ewes being heavier than two year old ewes (52.8kg and 47.3kg respectively, $P<0.001$). Average liveweight was also significantly affected by birth type and rearing type of the ewe (singles heavier than multiples); previous birth type and rearing type (non-productive ewes heavier than ewes that produced and reared multiples); current birth type and rearing type (ewes producing and rearing multiples heavier than non-productive ewes); and year (heavier in 2009 than in 2010).

Liveweight loss was significantly different between the sites ($P<0.001$, Fig. 1). Ewes at Katanning lost 11.5kg ±0.25 and ewes at Kirby lost 4.9kg ±0.33 between their maximum and minimum liveweights. Year had a significant effect with liveweight loss greater in 2009 than in 2010 at Kirby (10.7kg ±0.27 and 2.0kg ±0.37) and greater in 2010 than in 2009 at Katanning (8.1kg ±0.24 and 13.2kg ±0.26). Interactions between site, year and age of ewe ($P<0.001$) were as large as 9.5kg between two year old ewes in 2009 and 2010 at Kirby, and

as small as 1.4kg between two year old ewes in 2009 and three year old ewes in 2010 also at Kirby. The interaction between site by sire breed was also significant (P<0.001). Ewes at Kirby lost more weight if they had Merino sires than if they had Border Leicester sires (7.0kg *versus* 2.8kg), and ewes at Katanning lost more weight if they had Border Leicester sires than if they had Merino sires (12.3kg *versus* 10.7kg). Ewes that had previously raised a lamb lost less weight in the current year than those that were dry or were pregnant but did not raise a lamb to weaning. Ewes that were producing and rearing multiples in the current year lost more weight than those that were non-productive.

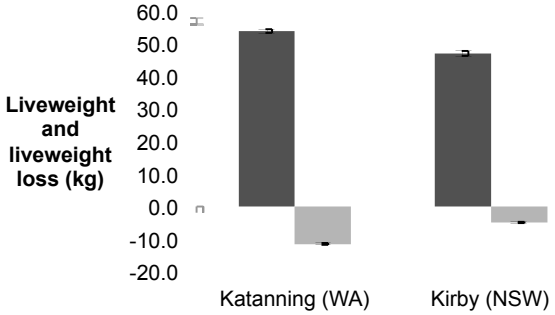


Figure 1. Predicted average liveweight (dark grey) and liveweight loss (mid grey) in ewes grazed at Katanning (WA) and Kirby (NSW) over three years (±SE).

On average, sire breeding values for fat had a significant impact on total liveweight loss (P<0.05), however there was a significant breeding value by site interaction (P<0.01). Ewes from sires with higher breeding values for fat lost less liveweight at Katanning (WA), with a reduction in liveweight loss of 3.2kg across the 2.5 mm range of sire PFAT values (Fig. 2). By contrast at Kirby (NSW) there was a negative relationship with an increase in liveweight loss of 1.5kg across the same range of sire PFAT values.

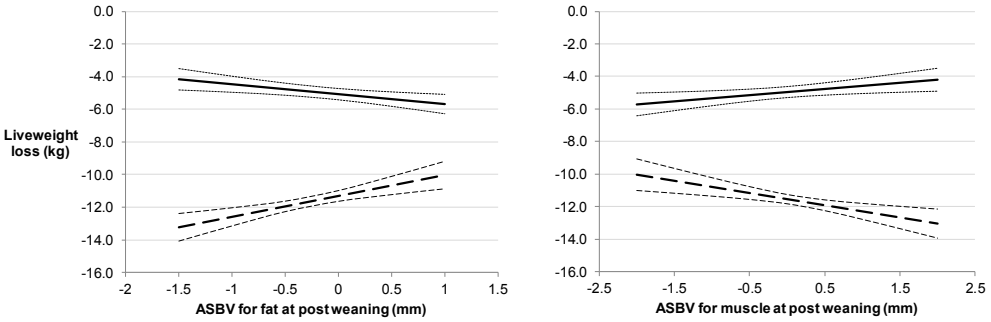


Figure 2. Predicted relationship between liveweight loss and sire breeding values for fat and for muscle for ewe progeny grazed at Katanning, WA (solid line), Kirby, NSW (broken line) over three years (±SE).

Sire breeding values for PEMD were also significant (P<0.01) and interacted significantly with site (P<0.01), with an increase in liveweight loss as sire PEMD increased at Katanning and a decrease in liveweight loss as sire PEMD increased at Kirby. Across the 4mm range of sire PEMD, liveweight loss increased by 3.0kg at Katanning but decreased by 1.5kg at Kirby.

There was no significant interaction between these sire breeding values and sire breed, so these liveweight loss responses to sire PFAT and PEMD at each site were equally evident for ewes from both Merino and Border Leicester sires.

DISCUSSION

Sire breeding values for fat and muscle influenced liveweight loss in their ewe progeny and the response differed between sites and these results support our hypothesis. These differences in the liveweight response of the progeny could be due to a number of additional factors including dam genetics and environmental conditions. At Katanning, ewes from sires with higher breeding values for fat lost less liveweight during summer and autumn, which is prior to the break of season and germination of annual pastures. By contrast, ewes from sires with higher breeding values for fat lost more weight during winter at Kirby, where they grazed on limited amounts of green pasture due to cold temperatures and slow pasture growth. This interaction between sire fat and environment is consistent with Ferguson *et al.* (2010) who reported that ewes with higher ASBVs for fat had a higher reproductive performance in some years but not others. The sire ASBVs by site interaction in the current study could be explained by differences in the quality of the grazed pastures and or the size of the nutritional stress and weight loss. Adams *et al.* (2002) concluded that fat stores would be more important in Mediterranean climates where sheep lose weight on protein poor pastures in contrast to higher rainfall regions where the feed supply is more consistent throughout the year. While feed quantity and quality were not measured at either site, ewes at Kirby consistently lost less liveweight. Adams *et al.* (2002) also concluded that those animals with a greater proportion of lean had a greater drive to eat on green feed and so high muscling may be more important in regions where green feed is more consistently available.

The positive effects of sire fat on liveweight loss quantified at the Katanning site could have broader application across southern Australia, especially as the losses in liveweight during summer and autumn are typically greater for autumn/early winter lambing flocks which still dominate. Young *et al.* (2011) quantified the potential economic value of genetic differences in liveweight loss during summer and autumn and concluded that reduced weight loss could be worth up to $2.30/kg per ewe. In the current study, a 1mm change in sire PFAT reduced liveweight loss by 1.3kg at the Katanning site which could equate to $2.90 per ewe. Further work is needed to establish the genetic correlations between fat and other production and carcass traits before advocating what selection pressure should be placed on fat for different production systems and environments, but the results of the current study do suggest it could be more important than previously considered in environments where ewes lose significant weight during summer/autumn.

ACKNOWLEDGMENTS

Funding for this work is being provided by the Australian Sheep Industry CRC, Murdoch University and the Department of Agriculture and Food Western Australia.

REFERENCES

Adams N.R., Blache D. and Briegel J.R. (2002) *Aust J of Exp Ag* **42**: 399.

Ferguson M.B., Young J.M., Kearney G.A., Gardner G.E., Robertson I.R.D. and Thompson A.N. (2010).. *Anim Prod Sci* **50**: 1011.

Freer M., Moore A.D. and Donnelly J.R. (1997) *Ag Sys* **54**: 77.

Hopkins D.L., Stanley D.F., Martin L.C., Ponnampalam E.N., and Van De Ven R. (2007) *Aust J of Exp Ag* **47:** 1208.

van der Werf J.H.J., Kinghorn B.P. and Banks R.G. (2010). *Anim Prod Sci* **50**: 998.

Young J.M., Thompson A.N., Curnow M. and Oldham C.M. (2011). *Anim Prod Sci* **51:** 821.

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:315-318

GENETICS OF BODY CONDITION SCORE AND ITS RELATIONSHIP WITH FERTILITY, MILK AND SURVIVAL IN AUSTRALIAN HOLSTEIN CATTLE

M. Haile-Mariam[1,2], R. Butler[3] and J.E. Pryce[1,2]

[1]DEPI Victoria, 5 Ring Road, La Trobe University, Bundoora, VIC, 3083, Australia
[2]Dairy Futures CRC, 5 Ring Road, La Trobe University, VIC, 3083, Australia
[3]Holstein Australia, 24-36 Camberwell Road, Hawthorn East VIC 3122, Australia

SUMMARY

First parity body condition score (BCS) data from the Holstein Association of Australia recorded between 1999 and 2012 were used to determine the heritability and genetic and environmental correlations of BCS with other economically important traits. Heritability estimates of BCS were 0.16 and 0.22 when estimated using single and multiple-trait (with other type traits) sire model analyses, respectively. Genetic correlations between BCS and milk yield traits were negative (~ -0.2). The genetic correlation of BCS with fertility and lactation length shows that BCS could be used as a predictor of fertility. Residual correlations of BCS with almost all the traits were in the same direction as the equivalent genetic correlations. The genetic trend in BCS as well as chest width (a highly correlated type trait) show a small decline in recent years, perhaps due to inclusion of liveweight breeding values in the Australian Profit Ranking. Although this trend is based on a small dataset and a short time span, there is a need to evaluate the consequences of selection for reduced live weight on fertility and health traits.

INTRODUCTION

Body condition of cows scored during the lactation is associated with milk yield, fertility and health of cows (Roche *et al*. 2009). However, the strength of these associations may vary depending on the production system, such as pasture based grazing or indoor feeding systems. For example, in the pasture-based dairy production system in New Zealand the genetic correlation of BCS with milk yield traits is near zero or positive (Pryce and Harris, 2006). But is negative in the US (Dechow *et al*. 2004) and Europe (Veerkamp *et al*. 2001). In Australia, pasture-based production systems dominate the States of Victoria and Tasmania, whereas more concentrate relative to forage is fed in the other States. Also condition scoring methods vary among countries (Roche *et al*. 2009), possibly as a result correlations with other traits and also variation in BCS may vary countries.

In Australia, monitoring the genetic trend of BCS is of interest because the Australian Profit Ranking (APR), the economic index, which was introduced about a decade ago, includes liveweight (LWT) predicted from type traits (stature, body depth and chest width). This was done to take account of variation in maintenance feed requirements related to body size. However, selecting for reduced LWT may have the unintended consequences, such as favouring cows with relatively poor condition score. Literature estimates (e.g. Veerkamp and Brotherstone 1997) show that in particular chest width is highly genetically correlated with BCS.

MATERIALS AND METHODS

In Australia, BCS is measured on a scale of 1 (thin) to 8 (fat) (Earle 1976) by professional type classifiers of the HAA once in the first parity. Data were available for cows that were also scored for type between 1999 and 2012. For comparison purposes and to examine the relationship between BCS and other traits, data of cows that calved from January 1994 were extracted from the ADHIS database. Other type data of cows with missing BCS or with BCS of below 1 and above 8

were included with missing BCS to avoid exclusion of data due to unintended selection. Finally approximately 430,000 cows of which 45% had valid BCS data were selected. About 90% of the type classifications over the years were carried out by 27 classifiers. Thus data of 34 classifiers, who scored less than 300 cows each, were excluded. Days in milk at classification varied from 1 to 500 days but for analyses reported in this study classification after 365 days were set to missing. Cows that were classified after 49 months of age and that calved for the first time after 38 months of age were also excluded.

To assess the relationship between BCS and other dairy traits of economic importance, the type data of cows were merged with milk yield and fitness data. Because the type data is managed by the breed society while the data on other traits by data processing centres (DPC) only about 55% of the cows with type data could be merged with their data for other traits. In the merged data, survival from first to second lactation was higher (i.e. 87%) than the average in Australian Holstein cows (i.e. 83%) which could be because cows culled early in the lactation are not classified. To examine if genetic correlations between BCS and other traits are influenced by the exclusion of data of cows not type classified, data of other cows were added if they were progeny of sires with type data and if they were contemporaries (the same herd-year-season-age, HYSA) to type scored cows with their type data coded missing.

A sire model was used to estimate the h^2 of BCS. The fixed effects fitted when analysing BCS (or type traits) included Herd-Classifier-Round, month of calving, age and days in milk at classification. Bi-variate models were used to analyse BCS with chest width in order to minimise the effect of selection on chest width on the h^2 of BCS. To examine the relationship between BCS and other type traits, BCS data of cows were analysed in multiple-trait models with type traits that were reported to be highly correlated. To examine the relationship of BCS with fertility (calving interval, calving to first service interval, pregnancy, first service non-return rate), production (daily milk yield close to 90-days, 305-day milk, protein, and fat yield) and survival, BCS data of cows were analysed using a set of tri-variate sire models (i.e. each analyses included BCS and lactation length (LL) because there was less selection on LL as almost all cows had LL data. The fixed effects fitted when analysing fertility, production, LL and survival included HYSA, month of calving and age at calving. For calculating EBVs for BCS an animal model was used. To illustrate the genetic trends, EBV of sires and cows with BCS data were plotted by birth year for BCS, chest width and bone quality.

RESULTS AND DISCUSSION

Month of calving, age and days in milk at classification had significant effect on BCS. Cows calving between June and Sept. were in poorer condition compared to those calving in Oct. to Dec. and Feb. to May. Older cows had higher BCS than younger cows. Cows classified early in lactation and late in lactation were in better condition than those scored in mid-lactation.

The estimated h^2 of BCS was 0.16 when analysed using a single trait sire model. The h^2 of BCS from a multiple-trait sire model was higher (Table 1). The h^2 for other type traits in Table 1 were within the range of estimates elsewhere (Veerkamp and Brotherstone 1997) and those used by ADHIS for genetic evaluation. The h^2 estimate of BCS was lower than estimates from some European countries (e.g. Veerkamp *et al*. 2001) but were similar to estimates from New Zealand (Pryce and Harris 2006), the US (Dechow *et al*, 2004) and Canada (Bastin *et al*. 2010).

Genetic correlations between BCS and selected type traits are shown in Table 1. Chest width had the highest genetic correlation with BCS, followed by bone quality and then angularity. Of the type traits used to predict LWT, BCS was least correlated with stature.

Table 1. Heritability (on diagonal), genetic (above diagonal) and residual correlation (below diagonal) and among body condition score and selected type traits

Traits	BCS	Bone quality	Stature	Angularity	Chest width	Body depth
BCS	0.22±0.01	-0.76±0.02	0.01±0.03	-0.70±0.02	0.81±0.01	0.48±0.02
Bone quality	-0.38[†]	0.28±0.01	0.07±0.02	0.71±0.01	-0.60±0.01	-0.19±0.02
Stature	0.04	0.00	0.38±0.01	0.12±0.02	0.14±0.02	0.17±0.02
Angularity	-0.30	0.48	0.06	0.23±0.01	-0.37±0.02	0.11±0.02
Chest width	0.39	-0.32	0.15	-0.13	0.24±0.01	0.71±0.01
Body depth	0.24	-0.13	0.15	0.13	0.47	0.34±0.01

[†]All standard error of residual correlations are approximately zero.

Table 2. Genetic (r_g) and residual (r_e) correlation between body condition score and fertility, yield and survival traits

Traits	r_g	r_e
Close to 90-day daily milk	-0.22±0.04	-0.06±0.0
305-day milk	-0.25±0.04	-0.06±0.0
305-day protein	-0.19±0.05	-0.04±0.0
305-day fat	-0.21±0.05	-0.04±0.0
Lactation length	-0.30±0.06	-0.05±0.0
Calving interval	-0.28±0.06	-0.05±0.0
Calving to 1st service interval	-0.45±0.09	-0.04±0.01
Pregnancy rate	0.10±0.13	0.04±0.01
1st service non-return rate	0.02±0.14	0.01±0.01
Survival to 2nd lactation	-0.02±0.06	0.01±0.0

Correlations of BCS with production, fertility and survival are presented in Table 2. All residual correlations regardless of the trait were close to zero, but most of them were significant as residual correlations were estimated with small standard errors. All genetic correlations with fertility traits were favourable meaning better condition cows had better fertility. Both milk yield early in lactation and 305-day milk yield have unfavourable correlations with BCS. These correlations were weaker than European (Veerkamp *et al*. 2001) and US studies (Dechow *et al*. 2004) but stronger than a study of New Zealand dairy cattle (Pryce and Harris 2006) and with the range of those reported by Bastin *et al*. (2010) from Canada. The genetic relationship between fertility traits such as CI and CFS with BCS were of similar magnitude to those observed in the US (Dechow *et al*., 2004) and the UK (e.g. Wall *et al*. 2007). Others have reported genetic correlations that are more favourable than the current study suggesting that the value of BCS as predictor of fertility could be higher (Pryce and Harris 2006). Of all correlations, those involving survival and LL were different from those observed in New Zealand where a genetic correlation of 0.35 with LL and 0.26 with survival were reported (Pryce and Harris 2004). Both our results and those in New Zealand are different from those in the US (Vallimont *et al*. 2013) where the correlation between productive life and BCS were negative (-0.48). The near zero genetic correlation with survival may mean that both cows with poor condition (for poor fertility) and good condition (for low milk yield) are possibly culled in Australian. Wall *et al*. (2007) found that life span in the UK had a positive genetic correlation with BCS.

Figure 1 shows the genetic trend for BCS and chest width which appears to have started to decline with the animals born in 2004 which coincides with the inclusion of predicted LWT into the APR. It is also worth noting that at about the same time, the US also included LWT using similar predictors in their index, Net Merit (VanRaden 2004) which may also have contributed to the decline. Based on animals born before 2004 there was no clear trend in BCS and chest width showing that selection on milk yield traits is not the main reason for the decline. However, it is worth noting that these results are preliminary given the time period and the amount of data on BCS but suggest that there is a need to evaluate the inclusion of predicted LWT in the APR and its possible impact on health and fertility traits.

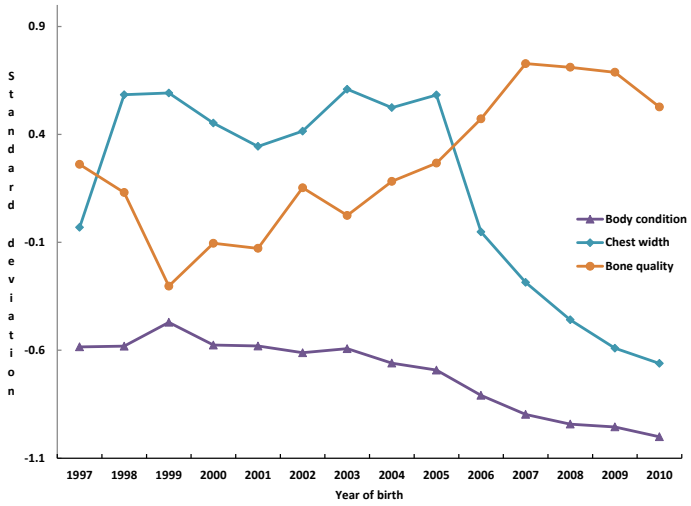


Figure 1. Genetic trend for body condition score, chest width and bone quality per genetic standard deviation

ACKNOWLEDGEMENTS

Funding for this work is provided by Dairy Future CRC. Type and body condition data were extracted by Dr Kevin Beard from the ADHIS database.

REFERENCES

Bastin C., Loker S., Gengler N., Sewalem A. and Miglior F. (2010) *J. Dairy Sci.* **93**:2215.

Dechow C.D., Rogers G.W., Klei L., Lawlor T.J. and VanRaden P.M. (2004) *J. Dairy Sci*. **87:** 3534.

Earle D. (1976) *J. Agric. Victoria* **74:** 228.

Pryce J.E. and Harris B.L. (2004) *Interbull Bull.* **32**:82.

Pryce J.E. and Harris B.L. (2006) *J. Dairy Sci.* **89**:4424.

Roche, J.R., Friggens, N.C., Kay, J.K., Fisher, M.W., Stafford, K.J. and Berry D.P. (2009) *J. Dairy Sci*. ***92:***5769.

Vallimont J.E., Dechow C.D., Daubert J.M., Dekleva M.W., Blum J.W, Liu W., Varga G.A., Heinrichs A.J. and Baumrucker C.R. (2013) *J. Dairy Sci.* **96**: 1251.

VanRaden, P.M. (2004) *J. Dairy Sci.* **87**: 3125.

Veerkamp, R.F. and Brotherstone S. (1997) *Anim. Sci*. **64**:385.

Veerkamp R. F, Koenen E.P.C. and de Jong G. (2001) *J. Dairy Sci.* **84:**2327.

Wall E., Coffey M.P. and Brotherstone S. (2007) *J. Dairy Sci*. **90**:1527.

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:319-322

PRODUCTIVE AND GENETIC DIFFERENCES BETWEEN COWS MANAGED ORGANICALLY OR CONVENTIONALLY

J.G. García-Muñiz[2], N. Lopez-Villalobos[1], R.G. Sherlock[3], N. Martin[1], C.W. Holmes[1] and N.M. Shadbolt[4]

[1]IVABS and [4]Agricomerce, Massey University, Private Bag 11222,
Palmerston North 4442, New Zealand
[2]On sabbatical leave from Departamento de Zootecnia, Universidad Autónoma Chapingo,
Chapingo, Estado de México
[3]Livestock Improvement Corporation, 605 Ruakura Road, Newstead,
Hamilton 3284, New Zealand

SUMMARY

International demand for organic dairy products has been on the rise. The productivity of organic production systems relative to traditional systems is therefore of increasing interest. In August 2001 the Dairy Cattle Research Unit at Massey University allocated 44 cows managed as a conventional herd and 44 Holstein-Friesian cows managed as an organic herd to monitor the differences between organic and conventional dairy farming systems. Replacements for each herd were produced using semen from high breeding worth (BW) bulls but in the case of the organic herd the number of bulls available was restricted by the organic standards. To compensate for this restriction, a mate-selection program allowing crossbreeding was used to maximise the future production worth (PW) of progeny in the organic herd. Management of stocking rate and use of purchased supplements attempted to equalise total feed offered per cow in both herds. The objective of this study was to compare the productive and genetic differences between the two herds, for the production season 2008-09, the 7th season of organic management. Compared to the conventional herd, the organic herd had a higher proportion of Jersey genes (0.38 vs 0.20, P=0.001), similar milk production (4,019 vs 3,899 L, P=0.28), similar fat production (207 vs 208 kg, P=0.93), similar protein production (145 vs 143 kg, P=0.66), similar lactation length (268 vs 272 days, P=0.38), similar liveweight (465 vs 464 kg, P=0.97), lower BCS (4.09 vs 4.26 units, <0.05), slightly lower BW ($113 vs $125, P=0.13), similar PW ($128 vs $126, P=0.89). Results from this study show that, in organic herds, strategic selection of sires for crossbreeding to maximise PW of future replacements can compensate for the lower BW of the sires, caused by the limited number of high BW sires available for use on organic farms.

INTRODUCTION

Demand for organic dairy products has been on the rise in United States of America (McBride and Greene 2007) and Europe (Nauta *et al.* 2005). Organic milk production differs from conventional milk production systems in several ways. In Norway, organic dairy herds have an accentuated spring calving period, lower production intensity, a generally older herd, and a more complex breed composition than is found in conventional dairy herds. Also, Norwegian certified organic milk producers are allowed a maximum of 30% total energy intake per year from concentrates (Harden and Edge 2001). The objective of organic milk production is to employ local natural resources without use of chemical fertilizers and pesticides. A key focus of the organic farming system is the maintenance of health and well-being of cows in the herd, without the routine use of conventional treatments methods (e.g., antibiotics).

Milk production per cow in organic herds has been reported as lower than or similar to production per cow in conventional herds. A Norwegian study (Harden and Edge 2001)

reported that milk production per cow in organic herds was lower than milk production per cow in conventional herds with only small differences in milk somatic cell counts. Similarly, a Dutch study (Nauta *et al*. 2005) showed that milk production was lower and somatic cell counts were higher in long-standing-organic dairy farms compared with conventional and recently converted organic farms.

A Danish study (Kristensen and Kristensen 1998) of 13 organic and 18 conventional herds over a three-year period found that the peak milk yield was lower in organic cows but lactation persisted at a higher level for longer in the organic herds, leading to only marginal differences in annual herd production levels.

In New Zealand, little information is available about milk production and composition of milk from cows that experience a farm's conversion from a conventional to an organic grazing system. In addition, the availability of semen meeting organic standards was from bulls that on average had lower Breeding Worth (BW) than the bulls from the premier sire dairy team that farmers normally used in conventional herds. Now (2013), certified organic dairy farms can use semen from the premier sire team. Breeding worth (BW) is an economic selection index that estimates a cow's or sire's ability to breed profitable replacements. Production worth (PW) is another economic selection index that estimates the efficiency with which a cow converts feed into farm profit.

The objective of this study was to compare the productive and genetic differences between an organic and a conventional herd for the production season 2008-2009. These research herds were part of the long term experiment set up at the Dairy Cattle Research Unit, Massey University, to establish and monitor the performance of grazing dairy cows managed on a conventional or an organic system of milk production.

MATERIALS AND METHODS

Description of data. Historical data for the 2008-2009 season, on calving and drying off dates, calving number (primiparous or multiparous), lactation lengths and herd-test records of individual cows from a whole-farm systems experiment comparing organic with conventional milk production at Massey University were analysed. Cows from the experimental herds contributed data on yields (kg/day) of milk (MY), fat (FY), and protein (PY); additionally, monthly records of liveweight (LWT) and body condition score (BCS) per cow were also available during the corresponding production season. Throughout the experiment, semen from bulls that had been progeny tested in New Zealand was used to produce replacement heifers. Cows in the conventional herd were inseminated with high BW premier sires and cows from the organic herd with semen of New Zealand dairy bulls that met the organic standards of BioGro New Zealand; organic bulls were fewer and had lower BW than premier sires. To compensate for the lower BW of the organic bulls, a mate selection strategy (Lopez-Villalobos *et al.* 2004) was implemented for the organic herd. The mate selection used a multiple objective optimization (Tozer and Stokes 2001) to maximise PW and fertility and minimise somatic cell score (SCS) of future replacements.

Statistical analyses. The MIXED procedure (SAS 2008) was used to fit Legendre polynomials of 4^{th} (for MY) or 3^{rd} (for FY and PY) order to the herd-test data using random regression analysis to estimate the lactation curve parameters of individual cows (Brotherstone *et al.* 2000). The resulting random coefficients of the Legendre polynomials and the cows' actual duration of lactation were used to calculate the yields per lactation of milk, fat, and protein. The yields per lactation of milk and milk components, as well as the cows' LWT, BCS, lactation length, the percentage of Jersey genes, BW and PW were subjected to analysis of variance to test for differences due to production system and calving number, using the GLM procedure (SAS 2008).

RESULTS

Based on the size of the Akaike information criterion of sequentially fitted Legendre polynomials (SAS, 2008), 3rd (for FY and PY) and 4th (for MY) degree Legendre polynomials provided the best fit for the description of the individual lactation curves of cows from the experimental herds. The random individual cow lactation curves, the fixed regression line for the overall lactation curve, and the respective scatter plots for each variable and herd are displayed in Figure 1. Least squares means from the analysis of variance comparing the experimental herds, after accounting for differences in cow calving number, are presented in Table 1. Cows in both herds had similar values for lactation length, liveweight, BW, PW and yields of milk, fat, and protein. Cows in the organic herd were thinner and had higher percentage of Jersey genes.

DISCUSSION

Milk production per cow in organic herds has been reported lower than in conventional herds (Hardeng and Edge 2001; Nauta *et al.* 2005). This is often attributed to lower intake of concentrates (Hardeng and Edge 2001) and lower total dry matter intake for cows on these systems (Sehested *et al.* 2003). In the present experiment, however, lactation curves and the corresponding lactation yields of milk, fat, and protein did not differ for the conventional and organic herds, even though the mean BW of the conventional herd was higher than the mean BW of the organic herd. However the most representative measure of phenotypic potential for farm profit is PW rather than BW, because PW accounts for permanent and heterosis effects during the productive life of the cow. Despite cows from the organic herd being of lower BW they were able to match their counterparts' production figures in the conventional herd by being thinner and, on average, of similar PW.

Multiple objective optimisation has been useful to select sires for multiple objectives. For example Tozer and Stokes (2001) illustrated the use of this technique to select sires that maximise Net Merit and minimize inbreeding in the future progeny. A similar technique was implemented in the selection of sires for the Massey University organic herd. The results from the present study show that, in organic herds, strategic selection of sires for crossbreeding to maximise PW of future replacements can compensate for the lower BW of the sires that satisfied standards for organic production.

Table 1. Least squares means by production system and calving number for the yields of milk, fat, and protein, the percentage of Jersey genes, Breeding Worth, Production Worth, liveweight, body condition score (BCS) and lactation length of grazing dairy cows managed on a conventional or an organic system of milk production during 2008-09, the 7th season of organic management

	Production system			
Trait	Conventional	Organic	SED*	P
Milk yield (kg/cow/lactation)	3,899.0	4,019.0	110.4	0.28
Fat yield (kg/cow/lactation)	208.0	207.0	5.9	0.93
Protein yield (kg/cow/lactation)	143.0	145.0	3.2	0.67
Jersey genes (%)	20.4	37.8	0.05	0.001
Breeding Worth ($)	124.5	112.5	7.9	0.13
Production Worth ($)	125.7	127.9	15.5	0.89
Liveweight (kg)	464.3	464.8	11.2	0.97
BCS (units, scale 1 to 10)	4.26	4.09	0.07	0.05
Lactation length (days)	272.0	268.2	4.5	0.38

**SED = standard error of the difference.*

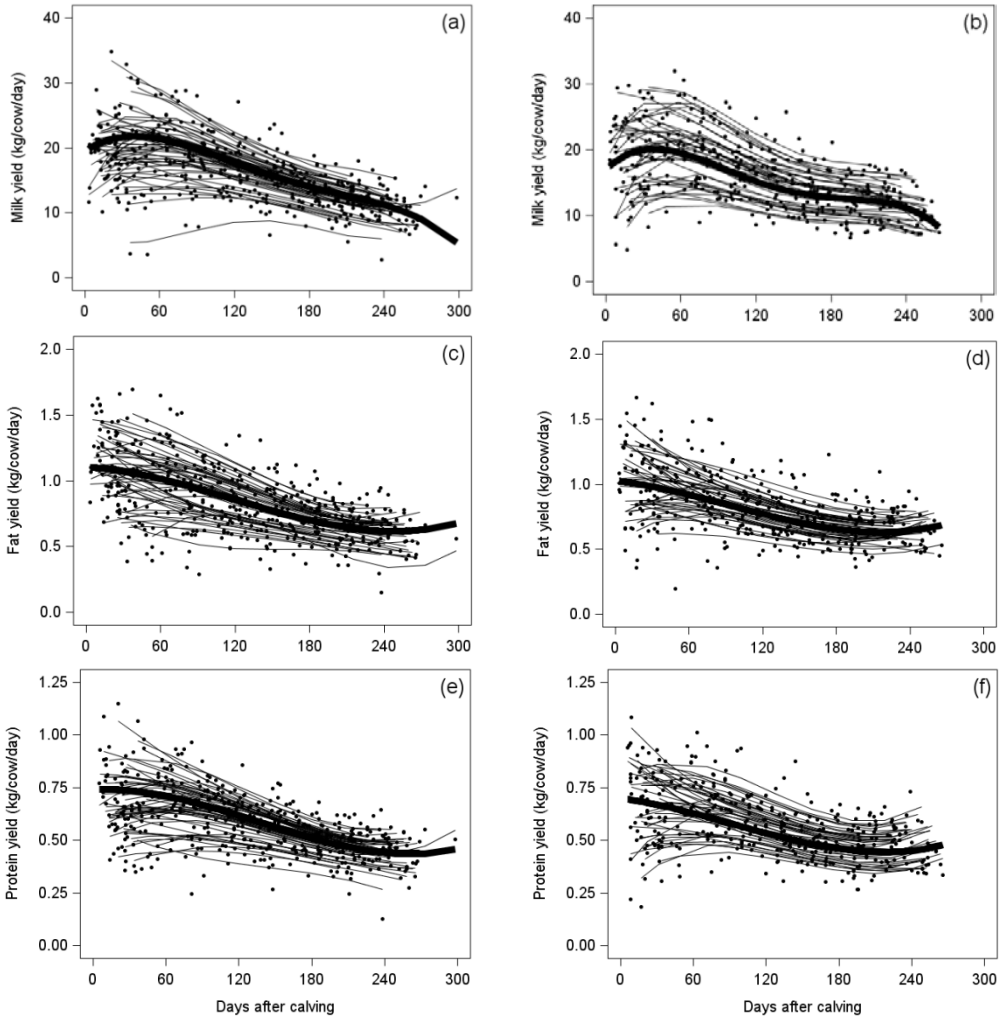


Figure 1. Scatter plots (solid black dots), overall fixed regression (thick black solid lines) and individual cow lactation curves (solid thin black lines) after fitting a 4th degree Legendre polynomial for MY (a: conventional; b: organic), and a 3rd degree Legendre polynomial for FY (c: conventional; d: organic) and PY (e: conventional; f: organic) of grazing dairy cows for the 2008-2009 production season.

REFERENCES

Brotherstone S., White I.M.S. and Meyer K. (2000) *Anim. Sci.* **70**: 407.
Hardeng F. and Edge V.L. (2001) *J. Dairy Sci.* **84:** 2673.
Kristensen T. and Kristensen S. (1998) *JordbrugsForskning* **2**: 6.
Lopez-Villalobos N., Harris B.L. and Pryce J.E. (2004) *Proc. N.Z. Soc. Anim. Prod.* **64:** 122.
McBride W.D., Greene C. (2007). *Agric. Econ. Assoc. Ann.Meeting*, Portland, Oregon, Aug, 2007.
Nauta W.J., Groen A.F., Veerkamp R.F., Roep D. and Baars T. (2005) *Neth. J. Anim. Sci*. **53:** 1.
SAS Institute. 2008. SAS Institute Inc., Cary, NC.
Sehested J., Kristensenb T. and Søegaard K. (2003) *Livest. Prod. Sci.* **80**: 153.
Tozer P.R. and Stokes J.R. (2001) *J. Dairy Sci.* **84**: 2782.

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:323-326

GENOTYPE BY ENVIRONMENT INTERACTIONS FOR AVERAGE DAILY GAIN USING MULTIPLE-TRAIT ANALYSES IN AUSTRALIAN PIGS

L. Li and S. Hermesch

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

SUMMARY

Data comprised of 265,103 records on pigs from nine herds collected from 2000 to 2010 were used to investigate whether genotype by environment interactions (GxE) existed for average daily gain (ADG) of pigs. Least squares means for herd by birth month from an animal model were used to quantify environmental conditions of contemporary groups. The environmental trajectory was divided into two, three or seven groups for alternative trait definitions of ADG considered to be a distinct trait for each environmental group. A multi-trait approach was used to investigate GxE. Heterogeneity of additive genetic variance and heritabilities were found for ADG between environmental groups when the environmental trajectory was divided into three or seven groups. Heritability estimates were highest for the intermediate environmental group (0.22±0.01) and reduced continuously to 0.15±0.02 for lower environmental groups. Estimated common litter effect did not differ significantly between trait definitions of ADG. Genetic correlations between ADG observed in different environments varied from 0.61±0.16 to 0.99±0.02. Genetic correlations were less than 0.80 when ADG was observed in two environments that differed by more than about 60 g/day indicating existence of significant GxE for ADG in pigs. At least 200 common sires were required to achieve statistical significance of these genetic correlations, demonstrating that large data sets with good data structures are required to detect GxE.

INTRODUCTION

Genotype by environment interactions (GxE) reduce the efficiency of a selection programme, as the ranking of animals differs between environments. Selecting the right genotypes for specific environments will increase genetic response across environments. Genotype by environment interactions can be analyzed using a multi-trait model in which traits records in different environments are considered separate traits (Falconer, 1952). Genetic correlations among separate traits quantify the extent of GxE, a value significantly less than unity demonstrates GxE. Further, a value of less than 0.8 was suggested to have biological importance (Robertson, 1959). This approach has been widely adopted to account for GxE in animal breeding.

Previous analyses (Li and Hermesch, 2012) showed that genotypes (breed or sire) had different sensitivities across the environmental trajectory defined by least squares means of herd by birth month contemporary groups (LSG). This study used multi-trait models to evaluate GxE for lifetime average daily gain (ADG) treated as a different trait for diverse environments classified according to LSG.

MATERIALS AND METHODS

Data. Records for 265,103 pigs from nine herds collected from 2000 to 2010 were available from the across-herd genetic evaluations of the National Pig Improvement Program database in Australia. Pigs were from three breeds: Large White (143,485), Landrace (87,946) and Duroc (33,672). Average daily gain was derived from live weight recorded shortly before slaughter on farms divided by age at recording. Mean (SD) for live weight, age at slaughter and ADG were 92.8 (13.6) kg, 143 (17.2) days and 649 (73.1) g/day, respectively. Based on previous analyses by Li

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

and Hermesch (2012), herd by birth month (HBM) was used as the contemporary group. There were 950 HBM groups with an average size of 279, ranging from 16 to 1071 pigs.

Analysis. Average daily gain based on all data was analyzed fitting the linear model: ADG = μ + sex + birth parity + breed + HBM + litter + animal + error, where μ is the overall mean. Fixed effects were sex, birth parity, breed and HBM contemporary group. Random effects were litter and animal effects. Genetic correlations were estimated using information from the numerator relationship matrix fitted in the animal model. The pedigree file contained 268,989 animals with 2,394 sires and 12,363 dams.

LSG from the model were used as an environmental descriptor to define environmental groups. LSG were normally distributed with mean (SD) of 644 (32.4) g/day and range of 534 to 738 g/day. Based on the distribution of LSG, three scenarios were considered to define ADG as a separate trait for different environments: 1) two environmental groups below and above the mean (644 g/day) of LSG; 2) three environmental groups for LSG <620 g/day, 620 to 660 g/day and >660 g/day; 3) seven environmental groups with increments of 20 g/day for LSG from <600 to >700 g/day. Genetic parameters were estimated fitting univariate and all pairs of bivariate analyses (3 and 21 pairs for three and seven trait definition respectively) using ASReml (Gilmour et al., 2009). Residual and phenotypic correlations were not estimated as each animal had only one observation.

RESULTS AND DISCUSSION

As expected, average phenotypic performance increased as environmental conditions improved as expected (Table 1). Coefficients of variation (CV) decreased from the inferior to superior environments for all trait definitions of growth rate. For the seven-trait analyses CV decreased from 11.3% to 8.7% indicating that pigs with higher growth rate had less observed variation relative to the mean. For all scenarios, all breeds had records across all traits.

Heritabilities. The heritability estimate for ADG defined as one trait across environments was 0.22±0.01 (Table 1). When ADG was treated as two traits, heritabilities did not differ significantly between these two traits. In contrast, heterogeneity of additive genetic and total variances as well as heritabilities existed for different environmental groups of the three- and seven-trait analyses. Highest estimates were found for the intermediate environmental group. In the seven-trait analyses, heritabilities and additive genetic variances decreased from 0.22±0.01 to 0.15±0.02 and from 964±61 to 622±88 g/day for ADG in the intermediate group (ADG4) to ADG in the lowest environmental group (ADG1). Zumbach *et al*. (2007) studied two purebred Duroc pig populations (P1 and P2) and their terminal crossbreds (C1 and C2) raised in different production environments and found a higher heritability estimate (0.32±0.01) in P1 raised in superior environments in comparison to the heritability estimate (0.16±0.01) obtained for C1 raised in inferior environments. However, no differences in heritability estimates between P2 and C2 was found. Common litter effect estimates did not differ significantly between trait definitions in our study.

Genetic correlations. Genetic correlations between ADG1 and ADG2 for two-trait definition and between ADG1 and ADG2 as well as ADG2 and ADG3 for three-trait definition were 0.98±0.01, 0.97±0.02 and 0.96±0.02, respectively. For three-trait definition, genetic correlation (0.78±0.06) between ADG1 and ADG3 differed significantly from unity with observed phenotypic mean difference of 76.8 g/day. The additive genetic (co)variance matrix among the seven traits was not positive definite, indicating that defining ADG as separate traits for less than seven environmental groups might be better for genetic evaluations. However, seven traits were defined in this study to see better the trend for change of genetic correlations along the environmental trajectory.

Table 1. Number of records (N), means and coefficients of variation (CV) along with additive genetic (σ^2_a), residual (σ^2_e) and phenotypic (σ^2_p) variances as well as heritability (h^2) and common litter effect (c^2) as a proportion of phenotypic variance for average daily gain (ADG) observed in inferior (i.e. ADG1) to superior (i.e. ADG7) environments

Scenario	Trait	N	Mean	CV(%)	σ^2_a	σ^2_e	σ^2_p	h^2	c^2
1 trait	ADG	265,103	650	11.3	955	2,845	4,314	0.22	0.12
2 traits	ADG1	136,641	625	10.9	834	2,885	4,268	0.20	0.13
	ADG2	128,462	675	10.3	977	2,851	4,317	0.23	0.11
3 traits	ADG1	63,269	610	10.9	714	2,908	4,182	0.17	0.13
	ADG2	122,081	645	10.4	936	2,909	4,378	0.21	0.12
	ADG3	79,753	687	9.9	840	2,853	4,179	0.20	0.12
7 traits	ADG1	19,118	593	11.3	622	2,999	4,206	0.15	0.14
	ADG2	44,151	618	10.6	727	2,879	4,165	0.17	0.13
	ADG3	56,459	635	10.4	815	2,925	4,278	0.19	0.13
	ADG4	65,622	654	10.2	964	2,932	4,434	0.22	0.12
	ADG5	44,518	672	9.9	737	3,111	4,368	0.17	0.12
	ADG6	19,245	695	9.3	648	2,864	4,066	0.16	0.14
	ADG7	15,990	721	8.7	731	2,403	3,593	0.20	0.13
Range of s.e.*		-	-	-	29-94	17-57	17-54	1-2	0-1

*Note: s.e. for h^2 and c^2 have been multiplied by 100.

Genetic correlations between ADG*i* and ADG*i+1* along with ADG*i* and ADG*i+2* ($0< i <6$) were not significantly different from unity indicating that no GxE existed for ADG expressed in similar environmental conditions (Table 2). Genetic correlations decreased as differences between environmental groups increased ranging from 0.61±0.16 to 0.99±0.02. Differences between phenotypic means of pairs (ADG*i* – ADG*j*, absolute value) ranged from 17 g/day (ADG2 versus ADG3) to 127 g/day (ADG1 versus ADG7) in the seven-trait definition. Genetic correlations were below 0.8 and of statistical significance when environmental groups differed by about 60 g/day (Figure 1a). Zumbach *et al*. (2007) found genetic correlations of 0.60±0.07 (P1 and C1) and 0.79±0.07 (P2 and C2) between growth rate recorded in purebred and crossbred populations that were raised in environments with different health status leading to superior performances of 60 and 100 g/day of the purebred populations in the two examples presented. Standard errors (s.e.) of genetic correlations were affected by the number of common sires shared between environmental groups (Figure 1b), decreasing from 0.27 to below 0.10. This indicates at least 200 common sires between two groups were required to detect GxE of biological significance. This threshold may vary for data sets with different data structures.

CONCLUSIONS

Genotype by environment interactions were found for growth rate based on variation in environmental conditions prevalent in herds with good health and management practices. Heritability estimates were highest for the intermediate environment and lowest for the most inferior environment. Genetic correlations decreased as differences between environmental groups increased. Estimates differed significantly from unity for ADG recorded in two environments that

varied in mean performance by about 60 g/day. This multi-trait methodology offers a practical approach to consider genotype by environment interactions for growth rate in pig breeding programs. However, large data sets with good data structures are required for genetic analyses.

Table 2. (a) Genetic correlations (above diagonal) and differences of phenotypic mean (below diagonal) of average daily gain (A); (b) Standard errors of genetic correlations (above diagonal) and number of common sires (below diagonal) observed between pairs of A for the seven-trait definition*

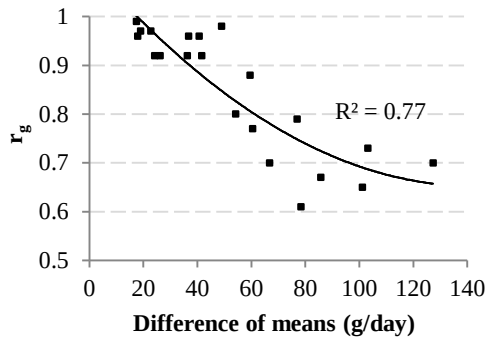
(a)

	A1	A2	A3	A4	A5	A6	A7
A1		92	92	77	61	65	70
A2	24		99	92	80	79	73
A3	42	17		97	96	88	67
A4	61	36	19		96	96	70
A5	78	54	37	18		97	98
A6	101	77	60	41	23		92
A7	127	103	86	67	49	26	

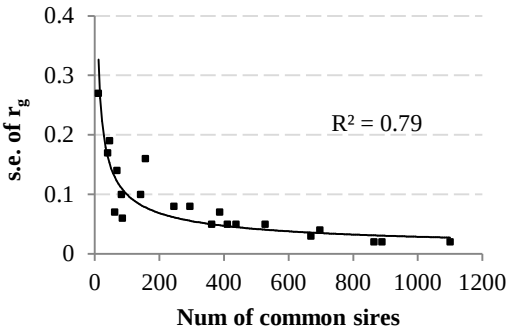
(b)

	A1	A2	A3	A4	A5	A6	A7
A1		5	5	8	16	19	27
A2	527		2	4	7	10	17
A3	437	889		2	3	8	14
A4	294	697	1100		2	5	10
A5	157	387	669	864		5	6
A6	46	142	245	362	410		7
A7	11	40	68	82	85	62	

* Note: Both above diagonal elements have been multiplied by 100; Estimates with underscore are significantly different from one ($p<0.05$).



(a)



(b)

Figure 1. Association between genetic correlations (r_g) and differences in means between two envrionments (a) and standard errors (s.e.) of r_g and number of common sires between two environments (b) based on seven-trait analyses.

ACKNOWLEDGMENTS

This project was funded by the Pork CRC under Project 2B-101.

REFERENCES

Falconer D. S. (1952) *American Naturalist* **86**: 293.

Gilmour A. R., Gogel B. J., Cullis B. R. and Thompson R. (2009) ASReml user guide. Release 3.0. VSN International Ltd., Hemel Hempstead, UK. 398 pp.

Li, L. and Hermesch, S. (2012) AGBU Pig Genetics Workshop – October 24-25 2012, http://agbu.une.edu.au/pig_genetics/workshop2012.html.

Robertson A. (1959) *Biometrics* **15**: 469.

Zumbach B., Misztal I., Tsuruta S., Holl J., Herring W. and T. Long (2007). *J. Anim. Sci*. **85:** 901.

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:327-330

INVESTIGATING THE GENETICS OF CULLING TIME AND THE EFFECTS OF FEEDING LEVEL ON OSTEOCHONDROSIS IN SOWS

E.M. van Grevenhof[1], H.C.M. Heuven[1], D.B. de Koning[2], W. Hazeleger[2] and B.J. Ducro[1]

[1]Animal Breeding and Genomics Centre, Wageningen University, P.B. 338, 6700 AH, The Netherlands. [2]Adaptation Physiology Group, Wageningen University, The Netherlands.

SUMMARY

Pig improvement schemes have traditionally aimed at improving growth rate and meat quality. More recently reproduction and longevity of sows and survival of piglets have been included in the selection objective. Improving longevity of sows is hampered by the lack of accurate and early recording of factors that contribute to reduced longevity. A study on growing pigs revealed large proportions of pigs showing signs of osteochondrosis (OC) (Van Grevenhof *et al.*, 2009). OC is a major cause of leg weakness in sows and hence an important economic and welfare issue (Kirk *et al.*, 2008). However, little is known about OC in sows and its impact on longevity. Our hypothesis is that including OC status in breeding schemes offers a good opportunity to more effectively select for improved leg quality and longevity.

The aims of this paper are to quantify and understand the mechanism of longevity by analysing OC and the time of culling after last insemination. OC is determined by genetic and environmental factors (housing, feeding) acting through biomechanical and metabolic pathways. A better understanding of these factors will enable the design of more effective management and breeding strategies. Experimental results, where different feeding levels were applied to 211 gilts, show that there are significant age dependent effects of feeding levels on the occurrence of OC. Switching to a higher feeding level after 10 weeks of age, increases OC prevalence compared to restricted feeding (OR: 1.8 - 8.5). In practise, gilts are regularly fed restricted early in life, after which feeding is switched to *ad libitum* for optimal growth. The results show that time of culling after last insemination is a heritable trait that might be used in selection in addition to longevity. By combining improvement of culling time with improved longevity, economics and welfare will even further be increased.

INTRODUCTION

Osteochondrosis. After fertility problems, leg weakness is the second most important reason for culling of sows. Osteochondrosis (OC) is a heritable disturbance of the endochondral ossification during skeletal growth and is a major cause of leg weakness in pigs and hence an important economic and welfare factor (Jorgensen and Anderson 2000; Kirk *et al.*, 2008). Feeding levels may be associated with osteochondrosis (OC) in the epiphyseal growth cartilage in gilts. As there is a short time frame of OC development in young growing animals, influencing OC may have different effects depending on the age. Little is known about the development of OC in gilts and sows. A study on growing pigs revealed that a large proportion of pigs show signs of OC (Van Grevenhof *et al.* 2009). The hypothesis for this study is that including OC at a young age as a selection trait offers a unique opportunity to more effectively select against leg weakness in sows and should ~~to~~ improve longevity of sows.

Culling time. Improving longevity of sows is hampered by the lack of accurate and early recording of factors that contribute to reduced longevity. It is known that leg weakness and fertility are major causes for culling of crossbred sows (Serenius and Stalder 2006). However, the correlation between leg weakness and longevity cannot be established as accurate recording of culling reason is often lacking, although the moment of culling is known. Time after last insemination until culling (culling time), is hypothesised to be a useful predictor of the culling

reason. In culling time, there appear to be two distinct peaks. The first peak after weaning is expected to be caused by detection of empty sows, while the second peak could be due to a mixture of culling reasons. Therefore, a genetic analysis was performed on culling time. Very little is known about the genetics of culling time, this study aims are to quantify and improve the understanding of the mechanisms of longevity by gaining insight into factors influencing time of culling after the last insemination.

MATERIAL AND METHODS

Osteochondrosis. This study will investigate age dependent effects of feeding levels, *ad libitum* versus restricted (80% of *ad libitum*), on the occurrence of OC in gilts at slaughter (26 wk of age). At weaning (4 wk of age), 211 gilts were subjected to 4 treatments of feeding levels. Gilts were administered either *ad libitum* feeding from weaning until slaughter (AA); restricted feeding from weaning until slaughter (RR); *ad libitum* feeding from weaning until 10 wk of age after which feeding levels were reduced to restricted feeding (AR); or restricted feeding from weaning until 10 wk of age after which feeding levels were increased to *ad libitum* feeding (RA) as often found in practice. At slaughter, the elbow joints, hock joints, and knee joints were collected. Joints were scored macroscopically for articular surface deformations indicative of OC. Analysis were done using PROC MIXED (SAS, 2010). The statistical model was $Y = \mu + treat + meas + (treat*meas) + pen + meas + e$, where Y represents the bodyweight observation of a gilt. The mean is represented by μ, treat represents the fixed class effect of treatments administered. The time points at which bodyweights were measured is represented by the fixed class effect meas. Interaction between treatments and measurements is represented by the fixed class effect (treat*meas) to assess differences between treatments for each measurement. *pen* represents the random effect of the experimental unit pen nested within treatment and is used as the error term for the treatment effect. The time points at which bodyweights were measured for each gilt, were also added as the repeated measures variable and is represented by *meas.*

Culling time. Records form a total of 111,987 F1 (reciprocal) crossbred sows were analysed, made available by the pig breeding company Topigs. The sows descended from 24,815 dams and 1,372 sires, kept on 189 farms in the Netherlands. All sows were born and culled between 2005 and 2012. The maximum parity reached (longevity) was measured by the parity of the last insemination and varied between 1 and 16 with a mean of 4.3 (SD 2.6). The animal model to estimate genetic parameters is $Y \sim mu + line + farm + YSbirth + sow + e$, where Y represents the culling time (categorical 1-5) or longevity (1-18), mu is the intercept of the model, line is the fixed genetic (reciprocal) cross of sows, farm represents the fixed farm effect at which the sows are kept, YSbirth accounts for a fixed year-season effect of birth of the sows and sow is the random effect of identification of the sow. A logistic regression model was used for the binary traits of culling category 1 to 5 (culled or not for each class separately). Analysis of the data were performed using SAS (SAS, 2010) and genetic parameters with a linear animal model using ASReml (Gilmour *et al.* 2009).

RESULTS AND DISCUSSION

Osteochondrosis. Figure 1 shows the impact of the feeding treatments on the change of bodyweight. The different age related feeding treatments resulted in varying OC prevalence. The OC prevalence was 77%, 60%, 57% and 43% for respectively the RA, AA, AR and RR treatment. Results show that gilts in the RA treatment have significantly higher odds to be affected with OC than gilts in the RR and AR treatments in the elbow joint (Table 1). Results show that there are age dependent effects of feeding levels on the prevalence of OC. Switching to a higher feeding level after 10 wk of age increases OC prevalence as opposed to a restricted feeding level. Age

dependent effects of feeding levels need to be taken into account to recognise its role in leg weakness and longevity of sows when a genetic analyses is conducted.

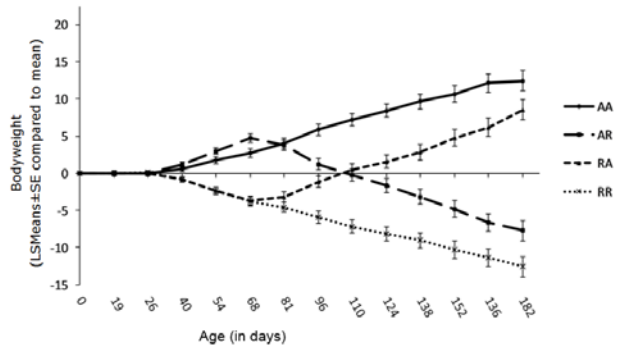


Figure 1. Bodyweight of treatment groups. Bodyweight (LSM +/- SE) profiles by treatment compared to the population bodyweight mean.

Table 1. Odds ratios (with P-values) of OC in elbow, hock and animal of different treatments compared.

	AR	RR
elbow RA	3.6 (0.04)	4.0 (0.03)
hock RA	3.3 (0.04)	8.5 (0.01)
animal RA	2.5 (0.01)	1.9 (0.01)
hock AA	5.3 (0.01)	

Culling time. Crossbred sows have the opportunity to express their full genetic potential for longevity. In contrast, purebred sows are replaced early to keep generation interval short to increase genetic response to selection. Results show that longevity, expressed as parity number of last insemination, has a heritability of 0.16 (SE 0.01) (Figure 2).

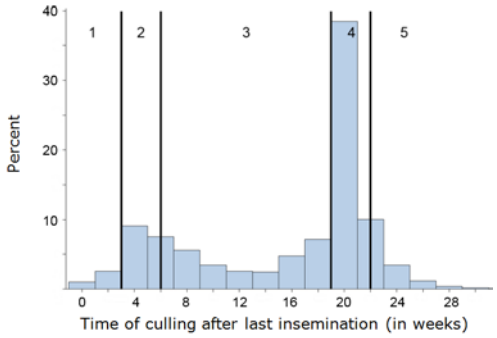


Figure 2. Time of culling after last insemination, expressed in weeks. The trait is divided into 5 culling time categories based on two distinct culling peaks in data, in categories 2 and 4.

Analysis revealed (Table 2) that culling time, expressed as time of culling after last insemination in 5 periods, has a heritability of 0.05. Combining longevity and time of culling in a bivariate analyses shows that the genetic correlation is 0.2 (not significantly different from 0), this suggests that these traits do not represent the same genetic mechanism, and that both traits can be combined in selection for improved sow performance. Heritabilities of the culling time categories 1 to 5 were found to be low, and varied between 0.03 and 0.10. The varying heritabilities possibly reflect the genetics of the main culling reasons at each time of culling. However the low heritabilities suggest that this threshold mechanism (due to censoring) by only taking a certain time period into account could cause bias and needs further analysis to fully understand the impact of these findings.

Table 2. Variances and heritabilities of traits related to time of culling after last insemination. The traits 'Culling time' and 'Longevity' are analysed using a linear animal model. The traits cullcat1-5 are expressed on a binary scale and analysed in a logistic regression model.

Trait	Scale	Var(a)	Var(e)	Var(tot)	h^2	SE
Culling time	1-5	0.035	0.737	0.772	0.046	0.004
Cullcat1	0-1	0.354	3.289	3.643	0.097	0.013
Cullcat2	0-1	0.158	3.289	3.447	0.046	0.006
Cullcat3	0-1	0.112	3.289	3.400	0.033	0.004
Cullcat4	0-1	0.100	3.289	3.389	0.029	0.005
Cullcat5	0-1	0.128	3.289	3.417	0.037	0.010
Longevity	1-18	0.730	3.827	4.557	0.160	0.010

CONCLUSIONS

The results show that time of culling after last insemination is a low heritable trait varying from 0.03 to 0.10 that might be used in selection in addition to longevity. Further analyses are needed to fully interpret the results and to relate the findings on culling time and longevity to observations of OC. By combining improvement of culling time with improved longevity, economics and welfare should benefit. Results showed that OC is influenced by feeding levels. In practise, gilts are regularly fed restricted early in life, after which feeding is increased to *ad libitum* for optimal growth. Age dependent effects of feeding levels need to be taken into account to recognise its role in leg weakness and longevity of sows when a genetic analyses is conducted.

REFERENCES

Gilmour A. R., Gogel B. J., Cullis B. R., Welham S. J., and Thompson R. (2009) ASReml User Guide, release 3.0, VSN, Hemel Hempstead, UK.

Jorgensen B. and Anderson S. (2000) *Anim. Sci.* **71**: 427.

Kirk R.K., Jorgensen B. and Jensen H.E. (2008) *Acta Vet. Scan.* **50**: 5.

SAS Institute Inc. (2010) SAS/STAT 9.22 User's Guide. Cary, NC: SAS Institute Inc.

Serenius T. and Stalder K.J. (2006) *J. Anim. Sci.* **84**: E166-E171.

Van Grevenhof E.M., Ott S., van Weeren P.R., Bijma P. and Kemp B. (2009) *Liv. Sci.* **135**: 53.

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:331-335

STAYABILITY TO CONSECUTIVE CALVINGS AS A MEASURE OF LONGEVITY IN CANADIAN SIMMENTALS

J. Jamrozik[1] and S.P. Miller[1,2]

[1]Centre for Genetic Improvement of Livestock, University of Guelph,
Guelph, ON, N1G 2W1 Canada
[2]Department of Agricultural, Food and Nutritional Science, University of Alberta,
Edmonton, AB, T6G 2P5 Canada

SUMMARY

Calving performance and culling data on Canadian Simmentals were used to determine whether a cow stayed in a herd for her 2nd and later (up to the 8th) calvings, given that she had calved as a 2 yr old. Estimates of heritability for stayability to consecutive calvings from linear random regression model were moderate (from 0.12 to 0.36) and they decreased in time. Variance due to cow's permanent environmental effect constituted the largest part of the total variance for all longitudinal points, followed by genetic and contemporary group components. Genetic effects of stayability to different calvings were highly correlated and the magnitude of correlation decreased with the increased time span between calvings. Stayability evaluations were favourably associated with estimated progeny differences for female fertility traits, direct and maternal growth and calving ease for bulls and cows. Scrotal circumference was not significantly correlated with estimates of genetic merit for stayability.

INTRODUCTION

Stayability is a measure of whether or not an animal remaines in the herd until a specified point in time. It is a trait with significant economic value. Stayability in beef cattle is traditionally defined as a probability that a cow remains in the herd until 6 yr of age given she has calved once (Brigham *et al.*, 2007). The main problem with such defined stayability is the lag between accurate prediction of stayability and the need for young replacement sires. Estimates of expected progeny differences (**EPD**) for stayability based on survival to 6 yr of age are currently being generated for Red Angus, Limousin and Simmental bulls in the USA (Garrick, 2011).

Martinez *et al.* (2005) considered 3 different approaches to defining beef cattle stayability: stayability to a specific age (whether a cow survived to a specific age given she was in the herd as 2 yr old), stayability to calving (whether a cow has a second, or later, calf given she had a calf as a 2 yr old), and stayability to weaning (whether a cow weans a second, or later, calf given she weaned the first calf). Brigham *et al.* (2007) estimated correlations among EPD for sires at different age definition for stayability (from 3 to 6 yr) ranging from 0.18 to 0.47, indicating that the expression of genetic merit of bulls changes depending on the age definition for stayability. Earlier measures of stayability, however, could serve as indicator traits of stayability to 6 yr (Martinez *et al.*, 2005).

Different statistical models have been used to analyze stayability. Random regression (**RR**) model is a longitudinal generalization of the multiple-trait model. Binary observations can be assigned to each discrete time in the cow's lifetime and EPD for stayability can be generated for

each point on the trajectory. Time dependent environmental effects are easy to implement in the RR model. Veerkamp *et al.* (2001) showed that the RR model is relatively robust to censoring.

The Canadian Simmental Association has been collecting calving performance data since the early 1970's. The Total Herd Reporting system providing culling information on cows has been in place since 2000. These two sources of data can easily be used to create stayability phenotypes for the population. The objective of this study was to estimate genetic parameters for stayability to consecutive calvings for Canadian Simmentals using a linear RR model and to compare stayability with other economically important traits in terms of associations among EPD for bulls and cows.

MATERIAL AND METHODS

Data. Stayability to calving was selected as a measure of animal's longevity. Each cow was assigned up to 7 binary records (**S2**, **S3**, …, **S8**) corresponding to stayability to 2nd, 3rd, and up to the 8th calving, given that she calved as a 2 yr old. Phenotypes for a given calving event were 1 (= cow calved, she was still present in the herd) or 0 (= no calving record, meaning that she was culled prior to that particular calving). Stayability records were generated by merging calving and culling data. The data file included 1,164,319 binary records on 188,579 cows. More than 62% cows had all 7 stayability records. Number of cows increased with a degree of completeness for a stayability history. Almost 70% of cows that calved as heifers stayed in the herd until their 2nd calving. Proportion of culled cows increased with the calving number: from 37% for S3 to 87% for S8.

Model. Random regression linear animal model was fitted to the stayability binary data. The model was:

$$y_{ijkmt} = YS_{it} + age_{jt} + \sum_{n=0}^{p} \beta_{ikn} z_{nt} + \sum_{n=0}^{p} \alpha_{mn} z_{nt} + \sum_{n=0}^{p} \rho_{mn} z_{nt} + e_{ijkmt} ,$$

where y_{ijkmt} is the stayability observation (S2, … , S8) on cow m for calving t; YS_{it} is fixed effect of year of birth by season of birth for calving t; age_{jt} is fixed effect of age at first calving class for calving t; β_{ikn} are random regression coefficients specific to k-th contemporary group (**CG**) defined as herd within i-th year by season; α_{mn} are random additive genetic coefficients specific to cow m; ρ_{mn} are random permanent environmental (**PE**) coefficients specific for cow m; e_{ijkmt} is the random residual effect for each observation, and z_{nt} are covariates. Orthogonal Legendre polynomials of order 3 were used for all fixed and random regressions.

Censored records were treated as missing data in the model. The CG effect comprised 72,986 levels. The pedigree file (5 generations back) included 282,775 animals.

In matrix notation the model can be written as:

$$\mathbf{y} = \mathbf{X}\,\mathbf{b} + \mathbf{U}\boldsymbol{\beta} + \mathbf{Z}\boldsymbol{\alpha} + \mathbf{W}\boldsymbol{\rho} + \mathbf{e},$$

where **y** was a vector of observations; **b** was a vector of fixed effects, **β** was a vector of random CG effects, **α** was a vector of animal additive effects, **ρ** was a vector of cow PE effects, **e** was a vector of residuals; **X**, **U**, **Z** and **W** denoted respective incidence matrices. Conditional distribution of the data (sorted by calving number) was assumed to be: $\mathbf{y} \mid \mathbf{b}, \boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\rho}, \mathbf{R} \sim N[\mathbf{X}\,\mathbf{b} + \mathbf{U}\boldsymbol{\beta} + \mathbf{Z}\boldsymbol{\alpha} + \mathbf{W}\boldsymbol{\rho}, \mathbf{R}]$; with $\mathbf{R} = \sum_{i=2}^{+8} \mathbf{I}_{n_i} \sigma^2_i$ and **I** denoting an identity matrix.

Methods. Bayesian methods with Gibbs sampling were used for fitting the model. Prior distributions for the parameters were: $\boldsymbol{\beta} \mid \mathbf{C} \sim N(\mathbf{0}, \mathbf{I} \otimes \mathbf{C})$, where **C** is the covariance matrix for the CG effect; $\boldsymbol{\alpha} \mid \mathbf{G} \sim N(\mathbf{0}, \mathbf{A} \otimes \mathbf{G})$, where **A** is an additive genetic relationship matrix between individuals, and **G** is the additive genetic covariance matrix between elements of **α**; $\boldsymbol{\rho} \mid \mathbf{P} \sim N(\mathbf{0}, \mathbf{I} \otimes \mathbf{P})$, where **P** is the covariance matrix for the PE effect; $p(\mathbf{b}) = N[0, p\mathbf{I}]$, with p = 10000 for all

levels of all fixed effects; $\sigma^2_k \mid \nu_k , s_k^2 \sim$ SIC $[\nu_k, \nu_k s_k^2]$, k = 2,…, 8; $\mathbf{C} \mid \nu_c, \mathbf{C}_0 \sim$ IW $[\nu_c, \nu_c \mathbf{C}_0]$; $\mathbf{G} \mid \nu_g, \mathbf{G}_0 \sim$ IW $[\nu_g, \nu_g \mathbf{G}_0]$; $\mathbf{P} \mid \nu_p, \mathbf{P}_0 \sim$ IW $[\nu_p, \nu_p \mathbf{P}_0]$; where ν_k and s_k^2 are parameters of independent inverted chi-square distributions, ν_c (ν_a, ν_p) and $\mathbf{C}_0$ (G_0, $\mathbf{P}_0$) are hyper-parameters of the inverted Wishart distributions. Minimal number of prior degrees of freedom was assumed for all co-variance components. Scale parameters for inverted Wishart were uncorrelated; prior values were equal to 0.01 and 0.0 for all variances and co-variances, respectively. All conditional distributions were of a closed form and Gibbs sampling scheme followed standard procedure for Gaussian linear models. Sampling was performed for 250,000 iterations where 50,000 constituted burn-in for each model. Convergence was monitored by visual inspection of trace-plots for selected co-variance components. Estimates of variance components, genetic parameters and EPD were calculated as posterior means of respective samples after burn-in.

Estimates of genetic merit for different stayability traits were subsequently correlated with EPD for other available traits in Canadian Simmentals. The traits were: growth (birth weight, weaning weight, post-weaning gain, yearling weight), calving ease, carcass (marbling, fat thickness, rib-eye area, carcass weight), female fertility (calving to first insemination, days to calving) and scrotal circumference. Correlations were estimated separately for bulls (with at least 10 daughters in the stayability model) and cows (with phenotypes for stayability).

RESULTS AND DISCUSSION

The largest contribution to the variance on the longitudinal scale was due to the PE effects, with an increasing impact along the time scale (from 40 to 80%), followed by genetic and CG components. Both genetic and CG relative contribution to the phenotypic variance decreased linearly with consecutive calvings. Residual variance constituted from 0.4% to 22% of the total variance, indicating relatively good fit of the model.

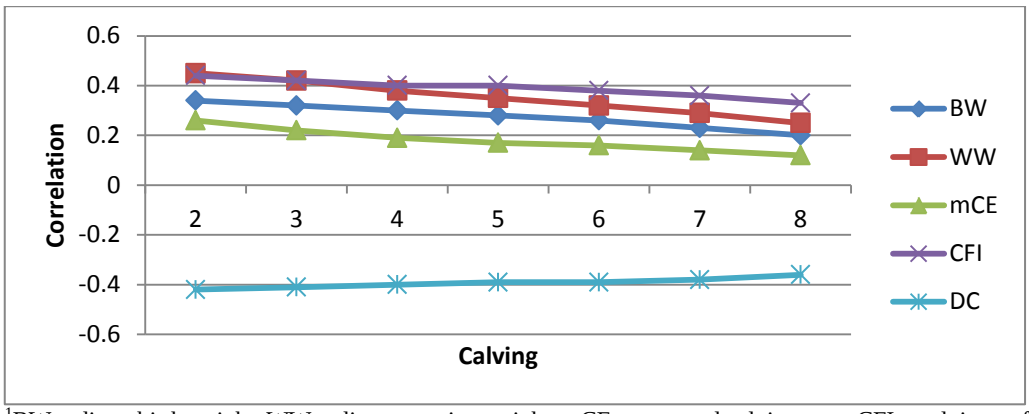
Table 1. Heritabilities (diagonal), genetic (above diagonal) and phenotypic (below diagonal) correlations for stayabilities to consecutive calvings (posterior SD are in brackets)

Calving number	2	3	4	5	6	7	8
2	**0.36**	0.96	0.89	0.83	0.77	0.71	0.61
	(0.006)	(0.002)	(0.006)	(0.008)	(0.010)	(0.011)	(0.013)
3	0.84	**0.23**	0.98	0.94	0.87	0.81	0.74
	(0.001)	**(0.004)**	(0.001)	(0.003)	(0.006)	(0.009)	(0.010)
4	0.67	0.96	**0.19**	0.98	0.94	0.89	0.81
	(0.002)	(0.001)	**(0.004)**	(0.001)	(0.003)	(0.006)	(0.008)
5	0.56	0.87	0.97	**0.16**	0.98	0.95	0.86
	(0.003)	(0.001)	(0.001)	**(0.004)**	(0.001)	(0.003)	(0.006)
6	0.47	0.71	0.85	0.96	**0.15**	0.99	0.89
	(0.003)	(0.002)	(0.001)	(0.001)	**(0.004)**	(0.001)	(0.005)
7	0.40	0.55	0.69	0.85	0.96	**0.13**	0.94
	(0.003)	(0.002)	(0.002)	(0.001)	(0.001)	**(0.004)**	(0.003)
8	0.31	0.44	0.53	0.64	0.75	0.88	**0.12**
	(0.003)	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)	**(0.003)**

Heritabilities of stayability to different calvings (Table 1) showed decreasing trend in time; S2 had the largest value of heritability. Stayability to the last calving (S8) still showed a reasonable

level of heritability. Heritabiliy estimates are comparable with results from other studies. Martinez *et al*. (2005) reported haritability for stayability to calvings (from second to sixth in Hereford cows) between 0.18 to 0.25 from linear models. Estimates of phenotypic and genetic correlations among different stayability to calving traits are in Table 1. Not totally perfect correlations indicated, in general, that stayabilities to different calvings are different traits from both a phenotypic and genetic perspective. The magnitude of correlations decreased with the increasing distance between calving events on the longitudinal scale. Stayability to 2nd calving (S2) would still, however, be a relatively good indicator of stayability to later calvings. Estimates of correlations among CG (PE) effects for different stayability traits exhibited, in general, similar patterns as phenotypic and genetic correlations.

Estimates of correlations between sires EPD for stayability and other traits available for Canadian Simmentals were smaller in magnitude than corresponding values for cows. Cows' correlations for selected trait are in Figure 1. Stayability was favourably associated (P<0.001) with female fertility traits (calving to first insemination and days to calving) for bulls and cows. Cows with better genetic potential for direct and maternal growth, calving ease, carcass marbling and adult cow weight tended to exhibit better stayability. Only maternal calving ease in sires influenced stayability, and the strength of associations decreased in time. Growth traits in bulls also showed positive correlations with stayability, although the magnitude of correlation coefficients was smaller than for cow's EPD. Sire's EPD for carcass traits did not show significant association with stayability. Similarly, scrotal circumference EPD were not significantly correlated with EPD for stayability for bulls and cows



[1]BW = direct birth weight, WW = direct weaning weight, mCE = maternal calving ease, CFI = calving to first insemination (heifers), DC = days to calving (mature cows)

Figure 1: Correlations among EPD for stayability and other selected traits[1] for cows with phenotypes for stayability

ACKNOWLEDGEMENTS

Funding for this research has been provided by Agriculture & Agri-Food Canada through the Canadian Agricultural Adaptation Program and by the Canadian Simmental Association. Data was kindly provided by the Canadian Simmental Association.

REFERENCES

Brigham B.W., Speidel S.E., Enns R.M. and Garrick D.J. (2007) *Proc. Western Section, ASAS,* **58**:27.

Garrick D.J. (2011) *Gen. Sel. Evol*. **43**:17.

Martinez, G.E., Koch R.M., Cundiff L.V., Gregory K.E., Kachman S.D. and L.D. Van Vleck (2005) *J. Anim. Sci.* **83**:2033.

Veerkamp R.F., Brotherstone S., Enge B. and T.H.E. Meuwissen (2001) *Anim. Sci*. **72**:1.

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:336-339

'DEER IMPROVEMENT' – GENETIC SELECTION IN A RECENTLY DOMESTICATED LIVESTOCK SPECIES.

B. Gudex[1], D. Johnson[1], D. Ford[2], T. Benton[2] and J. Chardon[1,2]

[1] LIC, Private Bag 3016, Hamilton 3240, New Zealand
[2] Deer Improvement, P.O. Box 1593, Queenstown 9300, New Zealand

SUMMARY

'Deer Improvement' is a commercial red deer breeding program based in the South Island of New Zealand. The breeding program utilises progeny testing, AI, MOET, foetal aging, DNA parentage testing, computer tomography (CT) and ultrasound carcass scanning, to maximise genetic progress towards the objective of improving the profitability of venison production. Annual genetic trends (based on the industry DEERSelect breeding values) of +0.8 kg/yr weaning weight, +1.1kg/yr yearling weight, +0.6kg/yr carcass weight and -0.01 days/yr conception date have been realised and 'Deer Improvement' has bred 14 of the top 20 stags on the 'DEERSelect' replacement index (July 2013 across herd evaluation). This paper discusses the structure of the breeding program and estimates genetic parameters for growth and eye muscle area traits.

INTRODUCTION

First farmed in New Zealand in 1969, there are currently ~1.1 million farmed deer in New Zealand from which ~23 thousand tonnes of venison (~$200 million) and 434 tonnes of deer velvet (soft antler, ~$26 million) was exported in the 2010-11 financial year (DINZ 2011). With the short history of deer domestication, farming and breeding deer has its challenges. Firstly, there is little prior knowledge/research on the species that can be utilised; secondly, deer exhibit marked seasonality in growth with little or no growth during winter and in reproduction; and thirdly, they retain a lot of wild behaviours that can adversely affect production and/or farmer safety. The 'Deer Improvement' breeding program began in 2004 with the intention of breeding superior venison producing red deer for distribution to the industry both on the hoof and via artificial breeding. Due to the relatively low value of the velvet exports, no emphasis has been placed upon selection for these traits. The objective of this paper is to profile the 'Deer Improvement' breeding program, thus illustrating the difficulties encountered in the genetic improvement of this recently domesticated livestock species and how these difficulties were overcome.

MATERIALS AND METHODS

The 'Deer Improvement' breeding program started with the purchase of 20 stags in 2004 and 15 in 2005 from a range of bloodlines. These stags were selected on the basis of within herd growth breeding values and, as a consequence, were predominantly of the recently imported Eastern European subtype which is larger/faster growing than the English subtypes that were originally captured from the wild for farming purposes in NZ. The selected stags were then progeny tested (1000 AI over 4 farms) and the top 2% of the resulting yearling stags and 6% of the yearling hinds were selected to form a nucleus herd at Deer Improvement's farm at Balfour (Southland). Currently, the farm has approximately 1000 hinds, of which the top 3% are used in an MOET program, the remaining 1st and 2nd fawners (2 & 3 year old hinds) are naturally mated and the mixed age older hinds are mated via a single round of artificial insemination (AI) and back up natural matings. The natural mating of the 1st and 2nd fawners allows the recording of conception date and avoids the lower AI conception rates that occur in these hind age groups. By selecting on conception date, the reproductive seasonality of red deer can be altered to allow more time for fawns to reach target weights before the venison price premiums expire in late spring and the next

cohort of fawns are born (Archer and Amer 2009). This is currently the only reproductive trait for which breeding values are estimated in deer.

Physically matching red deer fawns to their mothers and accurately determining birth dates is difficult due to the 'wild' origins of the species. Hinds typically hide their new born offspring (Morris and Archer 2007) and human interaction at this time adversely affects fawn survival (Asher and Pearce 2002). As a consequence, the parentage of each live fawn is established via DNA parent matching (GeneMark, LIC, Hamilton, NZ) in conjunction with mating, foetal age and mob information. Each fawn was DNA sampled at approximately 3 months of age. Originally the parent matching utilised microsatellite markers, but these were upgraded to a SNP panel for the 2010 born and subsequent cohorts (for more detail, see Gudex *et al.* 2013). Knowledge of birth dates is required for the accurate evaluation of growth (Amer *et al.* 1999) and is determined from the date of artificial breeding and/or foetal ages determined by ultrasonic pregnancy scanning, plus the gestation length of red deer (232 days). Prior to foetal aging, the conception date of naturally born fawns was determined by rotating stag teams between mobs of hinds so that each possible mating could only have occurred in a specific 2 week period. Unfortunately, this process adversely affected conception rates and stags are difficult to handle during the mating season.

In addition to conception date, weights (up to 8 are recorded during the first year), lean meat yield, hind fertility, conformation and behaviour traits are also recorded/observed. The collection of lean meat measurements started in 2007 and involves yearling stags that have been identified as potential sires undergoing a computed tomography (CT) scan prior to reaching 100 kg live weight (limit of scanner) and since 2010, all fawns undergoing ultrasound eye muscle scanning in October. The ultrasound scans cannot be carried out before the CT scanning as the winter coat of deer is comprised of hollow hair which interferes with the ultrasound waves (Ward *et al.* 2010). Fertility is assessed through ultrasound pregnancy scanning and both conformation and behaviour are observed subjectively by the farm manager. To date, no objective measures of deer behaviour have been found that adequately describe temperament with sufficient variation and heritability to be utilised in a breed program (Archer *et al.* 2009).

Breeding values are estimated primarily by the national deer genetic evaluation system – DEERSelect (Archer 2005, Archer and Amer 2009), though a separate growth breeding value is also estimated internally. The internal breeding values are estimated from all weights collected prior to 1 year of age using a bespoke random regression program (only direct genetic effects fitted - D. Johnson *unpublished* 2006). These are used to assist the selection of 16 month old stags in mid to late February for semen collection before a DEERSelect evaluation including the latest fawn cohort and their weaning weights is available. The genetic parameters for growth up to 1 year of age and for eye muscle area (ultrasound) were estimated via a multivariate animal model (no maternal effects) fitted in ASReml (Gilmour *et al.* 2009). The model included dam age, age at measurement, contemporary group (mob & birth year) and sex as fixed effects and covariates. Live weight at the time of measurement was also included as a covariate for the eye muscle area.

Avoiding inbreeding and maximising genetic diversity is a challenge due to the small number of stags that comprised the founder population of the Eastern European subtype in NZ and also the extensive use of AI and MOET. Currently, multiple lines are maintained to allow crossing where necessary and outcrosses are actively sought and progeny tested. The average herd inbreeding coefficients published in this paper were calculated using the pedigree viewer software (Kinghorn 2011) and the mate selection function of this software was used for the first time in 2013.

RESULTS AND DISSCUSSION

The age of the semen donor and farm were found to influence the conception rate to AI. On the Balfour farm in 2012, the average conception rate achieved using semen obtained from yearling stags was 49%, compared with the 77% obtained using semen from older stags. Variation

in AI conception rate between farms was observed by Deer Improvement's commercial AI service in 2012 to be between 63 and 83%. The flushing of hinds for embryo transfer revealed a hind age effect, with maiden hinds producing an average of 6 embryos per flush and older hinds (3 or 4 year old) an average of 12. Of the 247 embryos implanted in 2012, 70% were identified as being alive after tagging and DNA matching. The 2010 fawn cohort had parentage assigned using both the then new (in 2011) SNP marker panel and the existing microsatellite marker panel. The SNP panel was able to resolve both parents for 92% of the fawns and the microsatellite panel 68%. Utilising mating / mob / foetal age data allowed a further 2% to be resolved by the SNP panel and 4% by the microsatellite panel (Gudex *et al.* 2013). With 90 to 95% of fawns now being identified to their dams, it is possible to cull hinds that do not have progeny matched to them knowing that it is unlikely that the hind actually reared a fawn but the DNA failed to match them.

Table 1. **Fawn traits recorded by 'Deer Improvement' & their heritability (± standard error)**

Trait(s)	Units	First recorded	Number of animals recorded	Phenotypic standard deviation	Heritability
Weaning Weight	kg	2004	10700	7.58	0.392 ±0.025
Autumn Weight	kg	2004	9258	8.24	0.396 ±0.025
Spring Weight	kg	2004	8245	9.43	0.308 ±0.023
Yearling Weight	kg	2007	515	n/a - too few records	
Ultrasound Eye Muscle Scan*	mm & cm^3	2009	1227	3.07	0.246 ±0.061
Computer Tomography Scan	kg, mm & cm^2	2007	59	n/a - too few records	

* genetic parameters estimated from eye muscle area (cm^3), other measurements recorded include eye muscle depth and width (both mm).

Since 2004, 'Deer Improvement' has collected over 90000 progeny, weight, foetal age and lean meat records from approximately 13500 animals on 7 farms (6 progeny test and 1 breeding farm). No culling is carried out prior to the collection of the spring weight and although DEERSelect estimates breeding values for yearling weight, spring weights are submitted as they are within the age range permissible for yearling weight and the timing allows the weights to be included in October DEERSelect across herd evaluation. This evaluation is crucial for the business as the breeding values are used in each year's sale catalogue. 'Deer Improvement' also considers that spring weight is a better breeding objective than yearling weight given that there are venison price premiums available in the spring, feed is more abundant and the new cohort of fawns has not yet been born (Archer and Amer 2009). Genetic parameters for and between some of the traits recorded in the fawns are displayed in tables 1 and 2. It is important to note that these genetic parameters are estimated only from 'Deer Improvement' data and that the estimates used by DEERSelect (not publically available for comparison) are obtained separately from a larger and more diverse dataset.

Table 2. Genetic (below diagonal) and phenotypic (above diagonal) correlations (± standard error) between the fawn traits recorded by 'Deer Improvement'

Weight	Wean	Autumn	Spring	EMA
Weaning		0.918 ± 0.002	0.783 ± 0.005	0.512 ± 0.023
Autumn	0.984 ± 0.004		0.843 ± 0.004	0.507 ± 0.107
Spring	0.906 ± 0.015	0.924 ± 0.012		0.443 ± 0.027
EMA	0.573 ± 0.105	0.605 ± 0.107	0.444 ± 0.126	

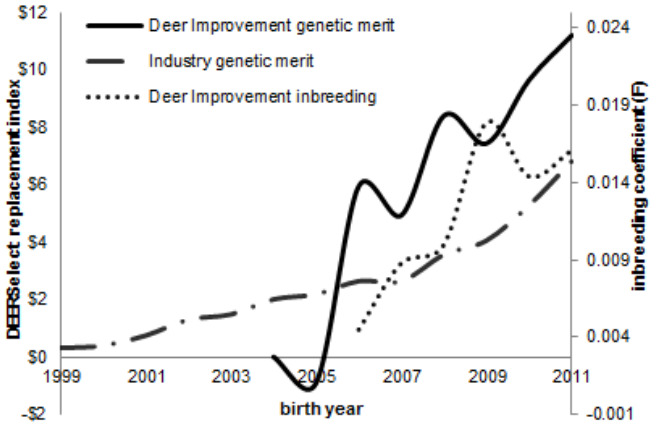


Figure 1. The genetic trend for the DEERSelect Replacement index for 'Deer Improvement' & across the whole industry. The inbreeding trend for 'Deer Improvement' is also displayed.

Since 'Deer Improvement' began in 2004, its' genetic improvement in the DEERSelect replacement index has been over twice that of the industry average ($1.60 vs $0.69 - figure 1). While most of the extra gain was achieved during the initial screening and selection step (2004 to 2006), the genetic trend after 2006 has remained ahead of the industry average ($1.05 vs. $0.84 per year) but the difference is less pronounced. This gain is reflected in the July 2013 DEERSelect across herd evaluation sire list where 14 of the top 20 stags ranked on the replacement index were bred by 'Deer Improvement'. Underpinning the increase in the index, weaning weights have risen by 0.8 kg per year, yearling weight up by 1.1kg per year (with 'Deer Improvement' 15 out of the top 20 sires), carcass weight up by 0.6kg per year (16 out of the top 20 sires) and conception date down by 0.01 days per year (9 out of the top 20 sires). The benefit to commercial farmers from this is that the increased growth and earlier conception date will make it easier to target finishing in the early spring where price premiums exist, there is greater feed availability and before the next cohort of fawns are born (Archer and Amer 2009). Balancing the genetic gains made has been a small increase in inbreeding within the herd, though it is impossible to determine how much of this increase is correlated with selection and how much is due to the more accurate and deeper pedigrees now available for the younger animals.

REFERENCES

Amer P., Nicoll G. and Garrick D. (1999) *Proc. NZ. Soc. Ani. Prod.* **59** : 129

Archer J., Asher G. and Ward J. (2009) *Proc. Deer Branch NZ. Vet. Assoc* **26** : 115

Archer J. (2005) *Proc. Deer Branch NZ. Vet. Assoc.* **22** : 89

Archer J. and Amer P. (2009) *Proc. Assoc. Advmt. Anim. Breed Genet.* **18** : 207

Asher G. and Pearce A. (2002) *Proc. Wld. Deer Farm. Cong.* **3** : 99

DINZ (2011). http://www.deernz.org/sites/dinz/files/DINZ-AnnualReport-FinalWeb% 20201011. pdf . Viewed 19 November 2012,

Gilmour A.R., Gogel B.J., Cullis B.R., and Thompson R. 2009 ASReml User Guide Release 3.0 VSN International Ltd, Hemel Hempstead, HP1 1ES, UK

Gudex B., Walker M., Fisher P. & Spelman R. (2013) *Anim. Genet.* (accepted for publication)

Kinghorn B.P (2011) *Genetics Selection Evolution* **43** : 4

Morris C. and Archer J. (2007) NZ J. Ag. Res. **50** : 163

Ward J., Archer J., Farmer R., Sweetman M., Nicoll G. and Yuan J. (2010) *Proc. NZ. Soc. Ani. Prod.* **70** : 275

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:340-343

PREDICTION OF GENOMIC BREEDING VALUES ACROSS GENETIC GROUPS

J.H.J. van der Werf[1,2], D.J. Brown[1,3] and A.A. Swan[1,3]

[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 Genetics and Breeding Unit, Armidale, NSW, 2351

SUMMARY

An estimate of breeding value is generally made up of two components; a prediction of line or breed effect (genetic group) and a prediction of the deviation within genetic group. When merging conventional and genomic breeding values, bias can easily occur when these two components are not correctly identified, estimated and weighted. More work is needed to determine the best way to combine information from pedigree based groups with genomic information.

INTRODUCTION

Genomic selection has been introduced in various livestock industries. Genotypic data on a relatively small number of animals are usually merged with phenotypic and pedigree information from many animals that were not genotyped. Methods that are used to achieve this vary from the *ad-hoc* "blending" methods to the so–called "single step" (SS) method (Misztal *et al.* 2009). In most genetic evaluations the origins of the current cohort of selection candidates can be traced back to a number of different base populations. These can represent different breeds or strains within breed. We will refer to these different genetic origins as "genetic groups". This paper provides a discussion on handling genetic groups in genomic evaluation.

The variation that exists across genetic groups can be large and grouping strategies can have significant effects on the ranking of selection candidates. This is the case especially in sheep breeding programs where across breed evaluations are common, the pedigree is often not very deep, and seedstock flocks are sometimes not sufficiently linked. When genomic information is merged with phenotypic information, handling group effects appropriately can be a challenge. Genotypic data provides information about population substructures, and this could be utilized to estimate differences across groups in genetic evaluation procedures. For example, how well can we rank an animal on genetic merit, when it is genotyped but otherwise of unknown origin? This paper explores the procedures used to estimate genetic merit of animals across genetic groups, both with and without genomic selection, and proposes strategies that can be used in genetic evaluation. We will use the term EBV for an estimated breeding value based on pedigree and phenotypes, GBV for estimated breeding values based on genotypes and phenotypes and GEBV for combinations of those.

THEORY

Across-group EBVs based on pedigree and phenotypes. Best Linear Unbiased Prediction (BLUP) procedures are generally used for the prediction of breeding values. Quaas (1988) described the theory of using genetic groups. A mixed model containing genetic groups is

$$\mathbf{y} = \boldsymbol{Xb} + \boldsymbol{ZQg} + \mathbf{Za} + \boldsymbol{e} \qquad [1]$$

where the vector $\boldsymbol{y}$ contains the phenotypes, $\boldsymbol{b}$ contains fixed effects, $\boldsymbol{g}$ refers to group effects, $\boldsymbol{a}$ refers to animals' additive genetic effects within genetic groups and $\boldsymbol{e}$ are residual effects. $\boldsymbol{X}$ and $\boldsymbol{Z}$ are incidence matrices relating data to fixed effects and animals, respectively. The matrix $\boldsymbol{Q}$ relates animals to groups and $\boldsymbol{ZQ}$ relates records to groups. We will consider both animal and group

effects as random. The across-group estimated breeding value $\hat{\boldsymbol{u}} = \boldsymbol{Q}\hat{\boldsymbol{g}} + \hat{\boldsymbol{a}}$. Genetic group effects can be estimated when sufficient phenotypes exist within groups, and if there is sufficient linkage between groups, i.e. animals from different groups with records in the same contemporary group. Quaas (1988) used mixed model equations based on Eqn.[1] but modified to

$$\begin{bmatrix} \boldsymbol{X'X} & \boldsymbol{0} & \boldsymbol{X'Z} \\ \boldsymbol{0} & \alpha\boldsymbol{Q'A^{-1}Q} + \lambda\boldsymbol{I} & -\alpha\boldsymbol{Q'A^{-1}} \\ \boldsymbol{Z'X} & -\alpha\boldsymbol{A^{-1}Q} & \boldsymbol{Z'Z} + \alpha\boldsymbol{A^{-1}} \end{bmatrix} \begin{bmatrix} \widehat{\boldsymbol{b}} \\ \widehat{\boldsymbol{g}} \\ \widehat{\boldsymbol{u}} \end{bmatrix} = \begin{bmatrix} \boldsymbol{X'y} \\ \boldsymbol{0} \\ \boldsymbol{Z'y} \end{bmatrix} \qquad [2]$$

where $\boldsymbol{A}$ is the NRM among animals, α = var(e)/var(a) and λ= var(e)/var(g). The modified equations provide solutions for across-group EBVs ($\hat{\boldsymbol{u}}$), and the part of the equations relating the genetic groups can be seen as an augmentation of the inverse of the matrix $\boldsymbol{A}$. In fact, we can factor out the group equations by substitution. Since off-diagonals blocks with fixed effects are zero, this equates to absorbing group equations into animal equations, giving

$$\begin{bmatrix} \boldsymbol{X'X} & \boldsymbol{X'Z} \\ \boldsymbol{Z'X} & \boldsymbol{Z'Z} + \alpha[\boldsymbol{A^{-1}} - \boldsymbol{A^{-1}Q}(\boldsymbol{Q'A^{-1}}\,\boldsymbol{Q} + \lambda\boldsymbol{I})^{-1}\boldsymbol{Q'A^{-1}}\,] \end{bmatrix} \begin{bmatrix} \widehat{\boldsymbol{b}} \\ \widehat{\boldsymbol{u}} \end{bmatrix} = \begin{bmatrix} \boldsymbol{X'y} \\ \boldsymbol{Z'y} \end{bmatrix} \qquad [3]$$

Eqn.[3] will give the same solutions as Eqn.[2] hence across-group EBVs are estimated in $\hat{\boldsymbol{u}}$ via regular mixed model equations without groups, but by using: $\boldsymbol{A^{-1}} - \boldsymbol{A^{-1}Q}(\boldsymbol{Q'A^{-1}Q} + \lambda\boldsymbol{I})^{-1}\boldsymbol{Q'A^{-1}}$ rather than $\boldsymbol{A}^{-1}$.
Therefore, the inverse of this matrix can be seen as an 'across-group numerator relationships matrix (*NRM*)' which is:

$$\boldsymbol{G} = [\boldsymbol{A^{-1}} - \boldsymbol{A^{-1}Q}(\boldsymbol{Q'A^{-1}Q} + \lambda\boldsymbol{I})^{-1}\boldsymbol{Q'A^{-1}}]^{-1}.$$

Consider a simple example where we have phenotypes on 4 unrelated animals, two from each of two genetic groups. The matrix $\boldsymbol{G}$ will then have diagonals 1+ k, within-group off-diagonals k and across group off-diagonals 0, where k= var(g)/var(a). When var(g) is large in comparison to var(a), i.e. for large k, the group differences will mainly determine ranking on across-group EBVs whereas with small k, $\boldsymbol{G}$ is close to $\boldsymbol{A}$ and groups can be practically ignored. Hence, Eqn.[3] shows that across-group EBVs can be calculated when using the appropriate across-group NRM. The latter involves knowledge of the group structure (as defined in the matrix $\boldsymbol{Q}$), and knowledge of variance ratio k.

Across-group GBVs based on genotypes and phenotypes. Genomic relationship matrices (***GRM***) can be constructed, e.g. using VanRaden (2008). When a ***GRM*** is formed for multi-breed populations the diagonal elements will be larger and there will be larger off-diagonal elements within breed compared with a breed-specific *GRM*. Also, off-diagonals within breeds will be larger than across breeds (or groups). This is similar to $\boldsymbol{G}$ described in the previous paragraph, and the ***GRM*** can be considered as an 'across-group' relationships matrix if across breed allele frequencies are used. However, there are two differences. Firstly, $\boldsymbol{G}$ is trait specific, because k varies between traits. The ***GRM*** could be trait specific if genomic regions were differentially weighted according to their significance in explaining genetic variance, but not otherwise. The second difference between $\boldsymbol{G}$ and ***GRM*** is that the grouping structure in $\boldsymbol{G}$, as defined in $\boldsymbol{Q}$ and based on pedigree, is not necessarily the same as the grouping structure implied by the ***GRM***. Principal component analysis has been proposed to reveal population structure from genomic information (Price *et al*. 2006). The variance among "groups" in a genomic evaluation can be estimated by replacing $\boldsymbol{Q}$ with the eigenvectors relating to the most significant eigenvalues of the ***GRM*** and fitting these to the data. Brown *et al.* (2013) observed strong relationships between the principal components of the GRM and average flock EBVs in Merino sheep. On the other hand, Daetwyler *et al*. (2010) observed that large single sire families could explain more variation in the ***GRM*** than lowly represented breeds. Moreover, genetic distances between individuals do not

always reflect phenotypic differences. Nonetheless, partitioning of variance based on the ***GRM*** structure is likely to be an improvement.

Combining EBV and GBV. In a method where information from phenotypes and pedigree is combined with genomic information, we have:

$$GEBV = w_1.EBV + w_2.GBV,$$

and the weights (w_i) are derived from their respective *within group* accuracies. If both EBV and GBV were across-group estimated breeding values we can rewrite this blending formula as:

$$GEBV_{across} = w_1.[PGroup + EBV_{within}] + w_2.[GGroup + GBV_{within}],$$

where PGroup is the group solution for pedigree defined groups and GGroup is the solution for groups derived from the ***GRM***. The GGroup term can represent breed differences for the animals that were genotyped, or genetic groups within breeds. Often, breed differences are fitted in genomic analysis, but it may be harder to fit a more subtle group structure within breeds. For example, in the Australian genetic evaluation, Merino rams are grouped by flock of origin, but within the cohort of genotyped animals there may be limited information per flock to estimate these differences reliably. Hence, flock-groups are not fitted in the genomic analysis and the GBV_{within} term will contain genetic differences between flocks and within breed because the ***GRM*** is derived across those flock groups. This is illustrated in Fig. 1 where we plot flock averages of GBV and EBV for 1610 young Merino rams in 34 Australian flocks that were genotyped using the Illumina 50K ovine SNP chip. The EBVs were calculated based on the full *MERINOSELECT* genetic evaluation based on pedigree and phenotypes, including those of animals that were genotyped. In this analysis, genetic groups are allocated to groups on a flock basis. The GBVs were estimated based on an analysis of genotyped animals only, including these young rams and ~10,000 animals in the CRC multi-breed reference population (Moghaddar et al., 2013). In the latter model, no genetic groups were defined at the flock level but Fig. 1 demonstrates that there is a good concordance between average flock solutions for EBV and GBV, hence flock differences are estimated implicitly using an across flock GRM. This is not a surprise, given that the reference population is an important link between animals from the different groups in both types of analysis, with most flocks having a significant genomic relationship with the reference population. This also confirms the observation in this paper, that genomic analysis can accommodate a between group component, even if it is not explicitly fitted.

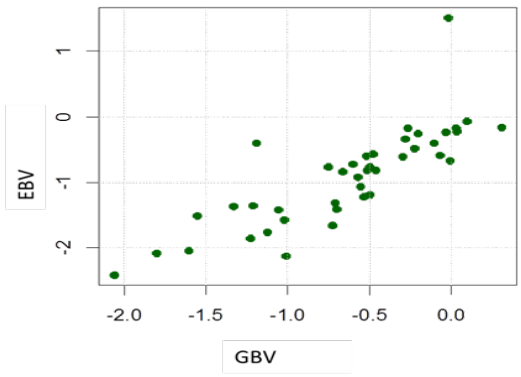


Figure 1: Average GBV for fibre diameter versus average EBV by flock (genetic group) based on 1610 Merino rams in 34 flocks in *MERINOSELECT*.

Although the example in Fig.1 suggests a good agreement between across group ranking of EBV and GBV, there are several pitfalls when blending these sources of information. If EBV and GBV contain between group differences, it would be appropriate to derive weights based on across group accuracy, or, better, to use different weights for 'between' and 'within'-group components of breeding value, because the reliability of estimating within-group differences might be very different from the reliability of estimating groups effects. With relatively few animals genotyped, the accuracy of estimating GGROUP effects is likely small. A second problem could occur if

genetic evaluation was across breed and breed differences maybe included in EBV but not in GBV. Breed effects or principal components are often fitted in genomic analysis, but not added back into GBV, hence only estimating $GBV_{within.}$ This would give:

$$GEBV_{across} = w_1.\,[PGroup + EBV_{within}] + w_2.GBV_{within},$$

which is biased as the between group differences are given insufficient weight ($w_1 < 1$). A solution could be to only blend the within group components and use the more reliable estimate of PGROUP to compare across groups:

$$GEBV_{across} = PGroup + w_1.EBV_{within} + w_2.GBV_{within}.$$

In this approach it is required that implicit group differences in GBV_{within} are not also included in the variation in PGROUP, but this may be difficult to avoid as illustrated in Figure 1. Finally, an obvious problem would occur in the blending procedure if there is a significant overlap in the data used to estimate EBV and GBV, respectively. A more coherent way to combine information is proposed in the SS method where information on genotype, pedigree and phenotypes are combined into one model (Misztal *et al*. 2009, Swan *et al*. 2011). Misztal *et al.* (2013) discussed how this procedure could account for genetic groups, and he proposes to use Eqn.[2] but with $\boldsymbol{A}^{-1}$ replaced by $\boldsymbol{H}^{-1}$, where $\boldsymbol{H}^{-1}$ is based on pedigree as well as genomic relationships. This procedure avoids double counting of information and the weighting of the various 'between' and 'within'-group components will be more likely correct. The main challenge is the genetic groups defined based on pedigree are not necessarily interpretable the same way as the genetic groups derived from the genomic data. More work is needed on how to best handle this problem in SS methods.

CONCLUSION

Both pedigree and genomic approaches can be used to retrieve information about an animal's breeding value from information across genetic groups. However, group definitions may differ and blending the information via ad-hoc methods can easily lead to bias. A SS approach should be able to handle this correctly, but methods to combine genomic and pedigree genetic groupings need more investigation.

REFERENCES

Brown D.J., Swan A.A. Gill J.S. and Banks R.B. (2013) *Proc. AAABG* **20**: (this publ.)
Daetwyler H.D., Kemper K., van der Werf J.H.J and Hayes B.J. (2012) *J. Anim. Sci*. **90**: 3375.
Misztal I., Legarra A., Aguilar I. (2009) J. Dairy Sci. **92**: 4648.
Misztal I., Vitezica Z.G., Legarra A., Aguilar I., and Swan A.A. (2013) *J. Anim. Breed. Genet.* doi: 10.1111/jbg.12025.
Moghaddar N., Swan A.A., and van der Werf J.H.J. (2013) *Anim. Prod. Sci.* (In Press)
Price A.L., Patterson N.J., Plenge R.M. *et al.* (2006) *Nature Genetics* **38**: 904.
Quaas R.L. (1988) *J. Dairy Sci.* **71**: 1338.
Swan A.A., Brown D.J., Tier B., and van der Werf J.H.J. (2011) *Proc. AAABG* **19**: 331.
VanRaden P.M. (2008) *J. Dairy Sci.* **91**: 4414.

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:344-347

SEQUENCING AND GENOTYPING FOR THE WHOLE GENOME SELECTION IN CANADIAN BEEF POPULATIONS

K. Stachowicz[1,2], S. Larmer[1], J. Jamrozik[1], S.S. Moore[3] and S.P. Miller[1,4]

[1]Centre for the Genetic Improvement of Livestock, University of Guelph, Canada
[2] AbacusBio Limited, Dunedin, New Zealand
[3]Queensland Alliance for Agriculture and Food Innovation, University of Queensland, Australia
[4]Department of Agricultural, Food and Nutritional Science, University of Alberta, Canada

SUMMARY

The project "Whole Genome Selection through Genome Wide Imputation in Beef Cattle" is a research initiative with a goal to develop low cost genome wide selection methodologies for Canada's beef industry. Ten cattle populations were included: 6 purebred beef breeds (Angus, Charolais, Gelbvieh, Hereford, Limousin, and Simmental), Canadian Holsteins, and 3 composite beef populations. The first step was to use pedigree analysis to identify the key animals to be sequenced and genotyped. For each population, 30 animals will be sequenced, 480 genotyped with HD SNP panel and 560 genotyped with 50K SNP panel. Pedigree analysis revealed good data quality, i.e. pedigree completeness and depth. Ancestors with the highest genetic and inbreeding contributions to the reference population were identified. From the top animals, 30 were chosen for sequencing based on their relationships with each other, to avoid sequencing closely related animals. The top 30 identified ancestors explained from 41% to 63% of the population gene pool, depending on the breed. Younger bulls with high number of progeny were considered for genotyping in addition to the top ranking ancestors, in order to ensure sufficient links to the phenotypic data. Genotyping the top 1,000 animals will cover over 90% of the genetic base of those breeds and should allow for highly accurate genomic prediction.

INTRODUCTION

The Canadian Cattle Genome Project, formally entitled "Whole Genome Selection through Genome Wide Imputation in Beef Cattle" (www.canadacow.ca), is focused on delivering genomic technology to Canada's beef industry. The project will include research to define the social and economic benefits and costs of using genomic technology in livestock improvement; develop tools for low-cost, accurate genome wide selection methodologies for breeders; and complete research so that genome wide selection can be used in Canadian herds for particularly difficult to measure yet valuable traits. Genotypes from a wide range of cattle populations will be used to develop accurate and robust genomic predictions.

Described is the method of identification of the key animals in the Canadian cattle populations to be sequenced and genotyped using the HD or 50K SNP panels.

MATERIALS AND METHODS

Pedigrees of purebred beef populations were obtained for Angus, Charolais, Gelbvieh, Hereford, Limousin, and Simmental breeds from respective breed associations. Holsteins were included as they make a significant contribution to global beef production and pedigree was provided by the Canadian Dairy Network. Analysis of each of the pedigrees was performed in order to assess data quality and pedigree structure using CFC (Sargolzaei *et al.* 2006) and Pedig (Boichard 2002) software. Completeness and depth of the pedigree are very important factors, which may affect the estimates of inbreeding coefficients, relationships among animals and also founder and ancestor contributions. Three different measures were used to assess the quality of pedigrees: percentage of animals with both parents known, discrete generation equivalent and

pedigree completeness index. The average number of discrete generations (DGE) provides an indication of how many complete discrete generations were present in a given pedigree (Sölkner *et al*. 1998). Pedigree completeness index (PCI) as a harmonic mean of parental contributions, is always zero when either parent is unknown regardless of the depth and completeness of the pedigree of the other parent. Inbreeding can also only be estimated if information on both parents' ancestors is available. Therefore, PCI is an important measure of pedigree quality for inbreeding and relationship estimation (MacCluer *et al*. 1983).

An inbreeding coefficient of each individual in the pedigree was calculated and averaged for each year of birth. However, as the absolute values of inbreeding are relative to the quality and depth of pedigree, the rate of increase of inbreeding (ΔF) per year (or per generation) should be used when comparing between different populations or assessing how inbreeding is accumulating in the population. It was also used to find effective population size for each breed, calculated as: $N_e=1/2\Delta FL$, where L was the average generation interval. The reference population included animals born between 2006 and 2011, which represented the last generation.

Effective number of founders is a measure of founders' contribution to the current population and reflects the unequal contributions of founders due to selection rates and variation of family size (Lacy 1989). Effective number of founder genomes is the number of equally contributing founders with no loss of founder alleles that would give the same amount of genetic diversity as is present in the reference population. It accounts for the loss of genetic diversity that occurred in the population due to genetic drift and bottlenecks. Effective number of ancestors and their genetic contributions (Boichard *et al*. 1997) were calculated to identify ancestors with highest marginal and total genetic contributions to the reference population. Additionally, the decomposition of inbreeding into ancestral components was performed, which allowed the identification of ancestors with the highest contribution to inbreeding in the reference population.

In order to choose animals for sequencing, top 100 ancestors with the highest marginal genetic contributions, top 100 with the highest total genetic contributions, and top 100 with highest contribution to inbreeding were chosen, which resulted in less than 200 top influential animals to be considered. Females were removed, as accessing DNA was highly unlikely. Remaining bulls were ranked based on their relationships with each other, to make sure that closely related animals will not be sequenced. As the approach described above does not identify younger bulls a secondary list of "young bulls" was created including bulls born after 2000 ranked based on number of progeny and descendants with phenotypic records. Top animals for each birth year till 2009 were chosen. The animals from the "young bulls" list (100-150 animals) were ranked based on relationships to make sure that they were not too closely related with each other and with older bulls chosen as described above. This resulted in the top 25 ancestors and top 5 young bulls with DNA available selected for sequencing from each breed.

In order to identify animals for genotyping, the top 3,000 animals with the highest genetic contributions and top 3,000 with the highest contributions to inbreeding were considered. For each breed, 400 ancestors (including 25 chosen for sequencing) will be genotyped with high density (649K) SNP panel and 560 with 50K SNP panel. They were chosen based on their contributions rankings and DNA availability. Additionally, for each breed 80 younger bulls (including 5 chosen for sequencing) will be genotyped with the HD panel.

A different approach was implemented when choosing Holstein animals, as at the time of analysis over 200,000 cows and bulls were already genotyped and over 40 were sequenced. To select the top 30 candidates for sequencing, an imputation analysis was carried out with a reference population comprised of 2,000 randomly selected animals genotyped with 50K panel. This ensured a large enough population to not be biased by a small number of sires, while still being computationally manageable for multiple imputations. Reference genotypes were filtered, randomly removing 5,000 SNP to mimic the imputation from a higher density panel to sequence

more closely (45K to 50K). SNP were not removed for minor allele frequency (MAF), as the imputation of alleles with minor allele frequency will be critical in imputation to full sequence. Imputation of rare variants will be of the utmost importance to many sequence studies, as these variants have been linked to disease traits in other species (Cirulli and Goldstein 2010).

The top 200 bulls genotyped with 50K were selected as candidates based on their genetic contributions. Using genotyped animals only for this study helped to ensure that DNA was available for all animals chosen to be sequenced. Imputation was carried out iteratively, using FImpute 2.2 (Sargolzaei *et al.* 2011). First, a reference population of 35 already sequenced bulls, whose genotypes were available, was established. At this point, any bull who had a sire or maternal grand-sire sequenced was removed from the potential candidate group. Based on relatedness to the population, candidate bulls were added 10 at a time, starting with the animals with the highest relationship coefficients with the entire population. Accuracy of imputation for all SNP and for SNP with MAF <5% was calculated, then each of the 10 bulls was individually removed and accuracies were once again calculated. Any bull, when removed, who affected the accuracy of imputation, either for all SNP or for low MAF SNP by greater than 0.5% was included in the reference population, and was indicated to be sequenced. As the iteration was processed, groups of 10 animals were continually assembled with the remaining animals until all sires had been considered. Once all animals had been considered, the group with the greatest contributions to imputation accuracy were selected to complete the group of animals to be sequenced.

For genotyping, Holstein ancestors with the highest genetic and inbreeding contributions that have not already been genotyped and have DNA available were selected. Additionally, a high degree of relatedness to the entire population, and more importantly to the group of sequenced animals, was thought to be ideal. This will ensure high imputation accuracy from HD panel to sequence and help to accurately grow the database of sequenced individuals.

For synthetic populations, pedigree quality was not sufficient to perform analysis described above for beef breeds. Most influential animals were chosen based on the number of progeny with phenotypic records for the traits of interest.

RESULTS AND DISCUSSION

The results presented are for Angus (AN), Hereford (HE), Limousin (LM), Simmental (SM), and Holstein (HO) breeds. The data analysis revealed good pedigree quality for all breeds. Percentage of animals with both parents known varied from 85% for HO to 96% for AN. Discrete generation equivalent for animals born in 2011 was 11 for LM and SM, 12 for AN and HE, and 14 for HO. Pedigree completeness index was considered for 5 generations back and reached 99% for LM, 97% for HE and SM, 96% for AN, and 90% for HO. These results imply that choosing animals for sequencing and genotyping based on pedigree records was a reasonable approach.

The summary of the results obtained for the four breeds is presented in Table 1. The level of inbreeding for the reference population was considerably lower for beef breeds when compared to HO. However, similar rates of increase of inbreeding were observed for LM, SM and HO, which resulted in similar effective population size for those three breeds. The effective population size for AN and HE was significantly higher. Effective number of founders was the lowest for LM and the highest for SM, while effective number of founder genomes and effective number of ancestors were lowest for HO. This indicates that HO has a lower level of genetic variability when compared with beef breeds. This is further visible when looking at the number of ancestors needed to explain given percentage of gene pool. Six ancestors were needed to explain 50% of the gene pool in HO while 48 for HE. The 30 top contributing ancestors accounted for 41% of gene pool for HE, 43% for AN, 53% for SM, 61% for LM, and 83% for HO. The top 1,000 contributing ancestors explained 94% of gene pool for AN, 95% for HE and HO, 97% for SM, and 99% for LM. Therefore, genotyping those animals will provide very good coverage of the populations' gene

pool and will help to ensure good quality imputation for use in developing genomic predictions.

Future developments in genetic evaluation methodology will capitalize on genomic sequence data to provide more accurate estimates of breeding values for selection. Imputation makes it possible to provide sequence data on many animals at a reasonable cost. Although the accuracy of imputation to sequence in the individual breeds is not known, the methods presented provide a means to prioritize animals for sequencing to ensure maximum coverage of the unique genome segments in each breed, which will maximize the imputation accuracy for a given level of investment.

Table 1. Summary of the results for the reference population

	AN	HE	LM	SM	HO
Total number of animals in pedigree	1,566,899	1,087,982	423,639	1,168,127	10,530,778
Number of animals in reference population	444,832	107,236	44,852	140,657	1,753,375
Pedigree completeness index (%)	96	97	99	97	90
Average inbreeding (%)	2.2	2.8	2.9	2.1	5.8
Average rate of increase of inbreeding (%/year)	0.02	0.02	0.11	0.10	0.11
Generation interval (years)	4.87	4.67	4.96	4.88	5.09
Effective population size	545	429	91	107	88
Effective number of founders	611	463	171	681	309
Effective number of founder genomes	52	48	22	35	8
Effective number of ancestors	103	101	47	69	16
No. of ancestors explaining 25% of gene pool	11	11	5	7	2
No. of ancestors explaining 50% of gene pool	46	48	18	26	6
No. of ancestors explaining 75% of gene pool	184	178	63	96	18
No. of ancestors explaining 90% of gene pool	596	543	190	293	63
No. of ancestors explaining 95% of gene pool	1,136	1,029	341	621	932
No. of ancestors explaining 100% of gene pool	8,730	7,645	2,607	8,179	>200,000
% of gene pool explained by 30 ancestors	43	41	61	53	83
% of gene pool explained by 1,000 ancestors	94	95	99	97	95

ACKNOWLEDGEMENTS

Research was made possible through funding from Genome Canada, Genome Alberta, Alberta Livestock and Meat Agency, the Canadian Simmental Association and Agriculture and Agri-Food Canada's Growing Forward program. The cooperation of the Canadian Angus, Hereford, Limousin and Simmental associations as well as Holstein Canada and the Canadian Dairy Network is gratefully acknowledged.

REFERENCES

Boichard D., Maignel L. and Verrier E. (1997) *Genet. Sel. Evol.* **29**:5.

Boichard D. (2002) *Proc. 7th World Cong. Genet. Appl. Livest. Prod*. **28**:13.

Cirulli E.T. and Goldstein D.B. (2010) *Nat. Rev. Genet.* **11**:415.

Lacy R.C. (1989) *Zoo. Biol.* **14**:565.

MacCluer J.W., Boyce A.J., Dyke B., Weitkamp L.R., Pfennig D.W. and Parsons C.J. (1983) *J. Hered.* **74**:394.

Sargolzaei M., Iwaisaki H. and Colleau J.J. (2006) *Proc. 8th World Congr. Genet. Appl. Livest. Prod.* **27**:28.

Sargolzaei M., Chesnais J.P. and Schenkel F.S. (2011) *J. Dairy Sci. 94, E-Suppl.* **1**:421.

Sölkner J., Filipcic L. and Hampshire N. (1998) *Anim. Sci.* **67**:249.

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:348-351

USE OF HIGH DENSITY GENOTYPING AND TRAIT-DEPENDENT METHODS IN GENOME-ASSISTED EVALUATIONS

J.A. Jiménez-Montero[1], D. Gianola[2,3], K. Weigel[2], R. Alenda[1] and O. González-Recio[4]

[1]Dpt Producción Animal. ETSI Agrónomos. Universidad Politécnica de Madrid. 28040 Madrid, Spain
[2]Department of Dairy Science, University of Wisconsin-Madison, 1675 Observatory Dr., WI 53076,USA
[3]Department of Animal Sciences, University of Wisconsin-Madison, 1675 Observatory Dr., WI 53076,USA
[4]Dpto Mejora Genética Animal. INIA. Ctra. La Coruña km 7,5. 28040 Madrid

Presenting author: gonzalez.oscar@inia.es

SUMMARY

This study presents the predictive ability in genome-assisted evaluations using different density arrays and two different statistical methods. Predictive ability from the Genomic BLUP (**G-BLUP**) using 1632 progeny tested sires genotypes with the BovineSNP50v2 BeadChip (**50K**) were considered as benchmark. First, genotypes from the BovineLD BeadChip (**LD6K**) were imputed to a BovineSNP50v2 BeadChip (50K). Second, genotypes were imputed to the BovineHD BeadChip (**HD**). The Random Boosting (**R-Boost**) was evaluated as an alternative method to G-BLUP. Four traits were analyzed: milk yield (**MY**), fat percentage (**FP**), somatic cell count (**SCC**) and days open (**DO**).

In general, R-Boost and G-BLUP showed similar results with overlapping confidence interval. Low density genotypes imputed to 50K achieved a similar predictive ability than native 50K genotypes. However, an increase in Pearson correlation and lower predictive mean square error were found across traits when genotypes were imputed to HD. The larger improvements were found for DO when using imputed HD genotypes (up to 0.06 greater Pearson correlation units) and for FP using R-Boost (up to 0.20 greater Pearson correlation units).

These results showed that the predictive ability of certain traits may be improved either imputing genotypes to HD or utilizing a method that get adapted to the genetic architecture of trait.

INTRODUCTION

The next key objective in genomic selection programs is to translate the huge, and increasing, amount of genomic information in a useful tool to breeders (Pryce and Daetwyler 2012). Low density SNP panels and posterior imputation is a promising way to reduce genotyping costs while maintaining a large predictive reliability. There is a need to integrate different density SNP panels in genomic breeding programs. Further, there may be an interaction between the density of the original genotype and the statistical method for DGV prediction, and these may be trait dependent as well. It is known that methods based on marker regression have better predictive ability than methods based on genomic relationship matrices in traits that are regulated by major genes. Higher-density arrays are expected to capture a larger amount of genetic variance because LD between markers and causal mutations is supposed to be higher. However, previous studies have shown only a slight increase in predictive accuracy using arrays of up to 700K SNPs (VanRaden *et al*. 2013).

The objective of this study was to compare imputation accuracy, predictive ability, and selection efficiency for selection candidates genotyped at different densities using the Random Boosting (R-Boost) and G-BLUP algorithms.

MATERIAL AND METHODS

Genotypes and phenotypes. A total of 2658 genotyped bulls were used in this study, using the BovineSNP50.v2 Beadchip for 2226 bulls and the BovineSNP50.v1 Beadchip (Illumina Inc,) for 240 bulls. These 2658 bulls build up the 50K Spanish Holstein population (50K). Additionally, 192 were genotyped using the BovineHD BeadChip (HD). Editing and filtering processes of genotypes led to 39,714 and 540,501 SNPs for the 50K and HD evaluations, respectively.

A total of 1632 progeny tested bulls born before 2006 were used as the reference set (1412 for DO), labeled as **TRAIN50K**. Bulls born between 2006 and 2010 were used as the validation sets (382 for MY, FP and SCC and 216 for DO), labeled as **TEST50K**.

Four complex traits were examined: milk yield (**MY**), fat percentage (**FP**), somatic cell count (**SCC**) and days open (**DO**). These traits were selected to show differences regarding heritability of the trait and amount of phenotypic information available. Deregressed MACE progeny proofs (**DRP**) from January 2009 Interbull evaluation (Jairath *et al*. 1998) were used to estimate marker coefficients in the reference set.

Imputation. Low density genotypes in the testing set were created *in silico*, masking SNPs included in the 50K assay that were not included in the Bovine LD (LD6K) (Illumina Inc.) assay. Thereafter, phased haplotypes from 1632 animals in TRAIN50K were used as reference set for imputing the LD6K validation set using Beagle (Browning and Browning 2009). The outcomes were referred as 6K50K. Then, imputation from 50K (6K50K, TEST50K and TRAIN50K) to HD (6KHD, 50KHD and TRAIN50KHD) was implemented using the original HD population as reference.

Genomic evaluation models. Two different genomic evaluation models were used: Random boosting (R-Boost) (Gonzalez-Recio *et al*., 2013) and G-BLUP (VanRaden, 2008).

Predictive ability. *Accuracy and predictive MSE.* The prediction accuracy of genomic evaluations was computed as the Pearson correlation between the predicted DGVs and the December 2011 DRPs. The PMSE of predictions was also estimated. Means and confidence intervals were estimated using bootstrapping for each evaluation output (Efron, 1986), although these results are not shown in this work.

RESULTS AND DISCUSSION

Imputation Performance The allele error rate at imputing genotypes from customized LD6K to 50K was 1.3%. Those results are in accordance to previous studies using similar population sizes (Zhang and Druet, 2010; Berry and Kearney, 2011; Dassonneville *et al*. 2012). The LD6K array is an important tool for candidates preselection and genotyping females. Allele error rate of imputation from 50K to HD was 0.9% when a small reference population of 192 bulls genotyped in HD was used.

Predictive ability Pearson correlations obtained from the two methods are shown in Table 1. Both methods resulted in similar accuracy, although R-Boost was the preferred method for FP and G-BLUP for MY and DO, while for SCC results were case dependant.

Table 1. Accuracy for the genomic estimation of two evaluation methods for four traits of economic interest in dairy cattle after imputation from 6K and 50K to 50K and HD.

Trait	Method	6K50K	TEST50K	6KHD	50KHD
Milk yield (MY)	G-BLUP	**0.59**	**0.59**	0.54	0.55
	R-Boost	0.55	0.57	0.54	0.54
Fat percentage (FP)	G-BLUP	0.60	0.60	0.55	0.55
	R-Boost	0.78	0.78	0.79	**0.80**
Somatic cell count (SCC)	G-BLUP	0.49	0.48	**0.50**	0.47
	R-Boost	0.45	0.46	**0.50**	0.49
Days open (DO)	G-BLUP	0.29	0.19	**0.32**	0.31
	R-Boost	0.19	0.22	0.20	0.28

In bold: The preferred method within trait and set criteria

Prediction accuracy slightly increased for all traits after imputation to HD, excepting for MY. This increment was more relevant (up to 6 points in accuracy) for DO, which was the analyzed trait with smaller heritability. These results were in accordance with results previously reported for other Holstein populations, where estimates from HD were slightly better than those from 50K (Erbe *et al*. 2012, VanRaden *et al*. 2013).

Confidence intervals estimated by bootstrapping showed that distributions regarding prediction accuracy widely overlapped across methods and sets for MY and SCC (results not shown). However, R-Boost estimates were more accurate than G-BLUP for FP. As expected, large bootstrapped confidence intervals were found for DO (results not shown).

The MSE of prediction showed notable differences between evaluation methods (Table 2). R-Boost showed smaller PMSE in all four traits (12%, 54%, 12%, and 5% smaller for MY, FP, SCC, and DO, respectively). Note that R-Boost aims to minimize the MSE, as this is assumed as the loss function. Those results, and the aforementioned accuracies were in agreement with (Jiménez-Montero *et al*., 2013), although bootstrap analyses showed no significant differences. As before, lower PMSE were obtained after imputation to HD, excepting for MY.

CONCLUSIONS

Imputation using Beagle software was efficient for the reconstruction of 50K genotypes from low-density chips. Genomic evaluation methods (R-Boost and G-BLUP) resulted in similar prediction ability for the traits and genotypes included in this study. R-Boost showed clearly better performance for FP, and in terms of PMSE for all traits. However, no clear differences were found in terms of accuracy.

In general, some improvement in the predictive accuracy was obtained after imputation to HD. Genetic and genomic evaluation units should consider using different methods regarding the trait evaluated, and imputation to HD might be interesting to increase the predictive ability of some traits, especially those of low heritability or those regulated by major genes (e.g. FP, SCC, DO).

Table 2. Mean Squared Errors in the validation set regarding genomic evaluation method and imputation strategy (from 6K and 50K to 50K and HD) for milk yield, fat percentage, somatic cell count and days open.

Trait	Method	6K50K	TEST50K	6KHD	50KHD
Milk yield (MY)	G-BLUP	255	258	276	278
	R-Boost	236	**229**	241	240
Fat percentage (FP)	G-BLUP	0.044	0.044	0.048	0.047
	R-Boost	0.030	0.030	0.028	**0.027**
Somatic cell count (SCC)	G-BLUP	155.1	154.5	152.0	143.3
	R-Boost	138.3	137.4	**131.7**	133.4
Days open (DO)	G-BLUP	548.6	636.3	530.2	535.1
	R-Boost	541.9	523.4	546.4	**519.6**

In bold: The preferred method within trait and set criteria

ACKNOWLEDGMENTS

The authors acknowledge funds from the project CDTI-P080250866 UPM and the agreement INIA-CC10-046, to CONAFE, EUROGENOMICS consortium, ASCOL, ABEREKIN, XENETICA FONTAO and GENETICAL for providing biological samples and phenotypes used in this study and to "Dirección General de Producciones y Mercados Agrarios" ,"Laboratorio Central de Veterinaria del Ministerio de Agricultura, Alimentación y Medio Ambiente" for support of genotyping process and specially Dr. Dassonneville for helpful suggestions and comments.

REFERENCES

Berry D.P. and Kearney J.F. (2011) *Animal* **5:**1162

Browning B.L. and Browning S.R. (2009) *The American Journal of Human Genetics* **84:** 210

Dassonneville R., Fritz S., Ducrocq V. and Boichard, D. (2012) *Journal of Dairy Science* **95**: 4136

Efron, B., (1986) Statist. Sci. **1:** 54

Erbe M., Hayes B.J., Matukumalli L.K., Goswami S., Bowman P.J., Reich C.M., Mason B.A., and Goddard M.E. (2012) *Journal of Dairy Science* **95**: 4114

González-Recio O., Jiménez-Montero J.A. and Alenda R. (2013) *Journal of Dairy Science* **96**: 614

Jairath L., Dekkers J.C.M., Schaeffer L.R., Liu Z., Burnside E.B. and Kolstad B. (1998) *Journal of Dairy Science* **81**: 550

Jiménez-Montero J.A., González-Recio O. and Alenda R. (2013) *Journal of Dairy Science* **96**: 625

Mäntysaari E., Liu Z. and VanRaden P. (2010) *Interbull Bull* **41**: 17

Pryce J.E. and Daetwyler H.D. (2012) *Anim. Prod. Sci*. **52**: 107

VanRaden P.M., (2008) *Journal of Dairy Science* **91**: 4414

VanRaden P.M., Null D.J., Sargolzaei M., Wiggans G.R., Tooker M.E., Cole J.B., Sonstegard T.S., Connor E.E., Winters M., van Kaam J.B.C.H.M., Valentini A., Van Doormaal B.J., Faust M.A. and Doak G.A. (2013) *Journal of Dairy Science* **96**: 668

Zhang Z. and Druet T. (2010) *Journal of Dairy Science* **93**: 5487

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:352-355

APPLICATION OF WHOLE GENOME SEQUENCE TECHNOLOGY TO DAIRY CATTLE BREEDING BY LIC.

M. Keehan, A. Scott, T. Lopdell, T. Johnson and R. Spelman

LIC, Private Bag 3016, Hamilton 3240, New Zealand

SUMMARY

Whole genome sequence (WGS) technology has become affordable to animal breeding companies. This paper will describe how LIC is applying WGS to dairy cattle breeding, show results LIC has currently obtained from WGS and the expected benefits for dairy farmers. The benefits to farmers are improved reliability of genomic evaluation and the ability to detect and control low frequency recessive variations of large effect within the NZ dairy population.

INTRODUCTION

LIC is an artificial breeding and herd recording dairy cooperative owned by farmer customers that has approximately 80% of the NZ dairy market. LIC has been marketing bulls selected by genomic evaluation as well as the traditional method of evaluation based on daughter performance. Teams of genomically selected bulls have lower reliability than daughter proven bulls but they can improve the overall rate of genetic gain due to the lower generation interval. A genomically selected bull can be used at one or two years of age whereas a daughter proven bull is widely used at five years of age. This increase in the overall rate of genetic gain has potentially huge benefits for New Zealand. Despite these benefits there is huge customer demand to increase the accuracy and reduce the bias of genomic evaluations.

LIC has an extensive resource of genotypes from commercial Single Nucleotide Polymorphism (SNP) chips for gene discovery and genomic evaluation. They include 60,000 animals genotyped on the 50 thousand (K) SNP Illumina Bovine SNP50 (Illumina Inc., San Diego CA), 3276 animals genotyped on the 700K SNP Illumina Bovine HD and 12,000 on the "GGP" 9K SNP Geneseek Genomic Profiler (Neogen Corp., Lincoln, NE). Recently it has become cost effective to augment a standard SNP chip with additional custom SNP content.

The cost of whole genome sequence technology has declined from billions of dollars for the first human genome sequence to less than US$5000 per sequence. This tremendous reduction has been achieved from major improvements in all areas of genome technology. The decrease has enabled the 1000 Genomes Project Consortium (2010) to pioneer widespread low coverage sequencing to discover low to rare frequency variants in the human population. The 1000 Genomes project improved many bioinformatics tools e.g. alignment algorithms (Li and Durbin 2010), Samtools (Li *et al.* 2009) and variant detection and calling - GATK (Depristo *et al.* 2011). Phasing of genotypes from low coverage sequence data to form a reference panel which can then be used to impute sequence data from a low cost SNP chip is another important advance.

Whole genome sequence (WGS) has advantages over commercial SNP chips. Most variants in the population are found including structural changes, insertion, deletions and variants excluded by SNP chip chemistry. In some circumstances the maternal or paternal phase of the SNPs can be determined from WGS. Disadvantages of WGS include data size and comparatively low accuracy at low sequence coverage compared to SNP chip genotypes. Levels of sequence coverage can also vary throughout the genome.

LIC has been funded to embark on a seven year research project to apply WGS technology to dairy cattle breeding within NZ. This paper reports how LIC has been utilising the WGS data for

both genomic selection and gene discovery and gives preliminary and ongoing results from analysis of the sequence data and the development of a bovine mammary transcriptome resource.

MATERIALS AND METHODS

Sampling. Tissue samples have been taken from 502 animals. Phase 1 consisted of 25 animals sequenced in 2011 and Phase 2 includes an additional 477 sequenced in late December 2012. Samples came from the LIC quantitative trait loci gene discovery herd, LIC bull semen stocks and the Vialactia phenotypic outlier discovery program. Key ancestral bulls from the national pedigree were chosen by the program ExomePicks (Abecasis 2010).

Library construction. DNA was extracted and sent to Illumina FastTrack for library construction. Most libraries generated were 100 base pair paired end reads. 2K, 5K and 10K insert size libraries were constructed from one high coverage animal intended for deNovo assembly.

Sequencing. All libraries were sequenced on Illumina HiSeq. Twenty three phase1 animals sequenced in 2011 ran on the V2 chemistry and were sequenced to an average 28X coverage. The 477 remaining animals ran on the newer IlluminaV3 chemistry.

Mapping. Reads were aligned using the RTG version 2.7.2 (Real Time Genomics, Inc. San Bruno, CA) against version 3.1 of the University of Maryland Bovine genome reference assembly. (Zimin *et al* 2009) . Median mapping coverage (X) is 6X, mean is 10X and maximum is 138X.

Genotype Calling. Genotypes were called using RTG, Samtools and GATK pipelines. Phase 1 genotypes were called using the consensus from the three pipelines. Phase 2 genotypes were called using RTG v 2.7.2 in pedigree aware mode and samtools in multiple sample mode with default parameters. Phase 2 consensus genotypes have not been called at time of writing.

Imputation. An early trial assessment of imputing sequence has been done using Beagle version 3.3.2 (Browning 2006). The first 4 Mb of chromosome 3 from the RTG pipeline was selected as it had completed genotype calling first. SNP Variants to impute were selected by requiring a variant call format QUAL field > 30. The sequence genotypes were sub sampled to either the 50K or GGP set of markers. The subsampled genotypes were then imputed back to 50K, HD and then Sequence. The 50K and HD reference were the LIC standard production reference populations of 15,000 animals genotyped with 50K and 3000 with HD SNP chip.

Deletorious recessive detection. The program SnpSift from snpEff (Cingolani et al 2012) was used to identify variants that lack observed homozygotes of one allele. These variants were further filtered by snpEff and manual curation. Fourty nine candidates were then sent for validation genotyping in a sample of 1350 animals.

Genomic Selection. The Phase1 dataset was used to generate a list of SNPs within 32 important dairy production genes chosen from the scientific literature. This list of 500 SNPs was added as LIC custom content to a Neogen GGP panel. 13,650 animals from the LIC Sire Proving Scheme will be genotyped on this custom list. Results have not been analysed at the time of writing.

RNA-Seq. Bovine mammary samples were extracted from 29 cows. RNA was converted to DNA at the University of Auckland. NZ Genomics performed sequencing on Illumina HighSeq. Results were analysed using a standard Tophat and Cufflinks pipeline. (Trapnell *et al.* 2009)

RESULTS AND DISCUSSION

SNP Discovery. 29,362,664 SNP were called from the Phase 2 RTG call set.

Genotype calling. 62 of the 502 sequenced animals have independent HD genotypes and 393 have independent 50K genotypes. Concordance has been computed between the sequence and HD genotypes and varied depending on the average level of mapped coverage per animal. For sequence coverage of 4X to 6X concordance was around 97%, at 10X concordance rose to 99% and at >20X concordance was greater than 99.5%

Imputation. Imputing a 4 megabase test region of the genome from GGP and 50K to sequence gave a genotypic concordance of 0.93 and 0.95. The sequence reference was then filtered to use only SNPs with a reference phasing beagle $R^2 > 0.995$. This filtering removed 75% of the SNPs and gave genotypic concordances of 0.94 and 0.95. The "round trip" validation method (subsampling and using self as a reference) represents a perfect haplotype distribution match between the reference and the imputing population and gives an upper bound imputation accuracy Johnson *et al.* (2011) reported concordances of 0.98 to 0.99 for imputing from 50K to HD. The initial low imputation accuracy of sequence can be attributed to errors in heterozygotes from low coverage genotypes, and enrichment of sequence errors within low allele frequency SNPs. Future work will improve imputation accuracy by better SNP and genotype filtering strategies, using additional phase information from reads and pedigree and genotype likelihood recalibration techniques. Currently the whole genome call set is 120 GB for the 500 animals. If the 50,000 animals genotyped on the 50K SNP chip were to be imputed to sequence it would generate a 12 TB file. The exploitation of a 12 TB file is computationally and statistically challenging.

Deletorious recessive detection. Analysis of the validation dataset has not yet been completed. It was observed that one of these predicted recessives was also in the centre of an association peak for a currently unpublished recessive mutation and explained the phenotype perfectly. Hence it is likely this will be a powerful technique to detect recessive acting genes of large deleterious effect and low to moderate frequency in the population. This approach is similar to the method used by Van Raden *et al.* (2011) to discover three Holstein fertility haplotypes in the North American population. Farmers should expect further improvements in fertility through the reduction in frequency of deleterious alleles.

Genomic Selection. Genomic selection theory is based on using a panel of anonymous evenly spaced markers that are individually in linkage disequilibrium with causative variants. Whole genome sequence technology enables the discovery of almost all variants within the population of interest. Thus in theory it should be possible to identify all causative variants or a sufficiently large subset of markers with the greatest LD to the causative variant Unpublished LIC internal experiments with 2 major milk protein markers as fixed effects have shown improvements to genomic evaluation which suggests the benefit from using causal markers.

RNA-seq. The first application of the RNA-seq data has been used to correct annotation within a genomic interval that is being investigated as a causative gene. An additional exon was discovered in that region and the presence of a 3 base pair in-del confirmed.

CONCLUSION

The WGS data is being utilised in many ways. The first use of the sequence data has been as a NZ population specific SNP discovery platform. These SNPs will be used in genomic evaluation and to search for additional recessive alleles reducing fertility. RNA-seq will be used indirectly to help explain and validate the actions of causative variants. Imputation of WGS will be restricted to small regions of the genome or subsets of animals until imputation accuracy has been improved.

ACKNOWLEDGEMENTS

This research was funded by the New Zealand Ministry for Primary Industries through the Primary Growth Partnership and LIC.

REFERENCES

Abecasis G. (2010) http://genome.sph.umich.edu/wiki/ExomePicks
Browning S. (2006) *Am J Hum Genet* **78**:903.
Cingolani P., Platts A., Wang le L., Coon M., Nguyen T., Wang L., Land S.J., Lu X., Ruden D.M.

(2012) *Fly* **6:** 80.
DePristo M.A., Banks E., Poplin R., Garimella K.V., Maguire J.R., Hartl C., Philippakis A.A., del Angel G., Rivas M.A., Hanna M., McKenna A., Fennell T.J., Kernytsky A.M., Sivachenko A.Y., Cibulskis K., Gabriel S.B., Altshuler D. and., Daly M.J. (2011*) Nat Genet* **43**: 491.
Johnson D., Spelman R., Hayr M. and Keehan M. (2011) *Proc. Assoc. Advmt. Anim. Breed. Genet*. **19**:379.
Li H., Handsaker B., Wysoker A., Fennel T., Ruan J., Homer N., Marth G., Abecasis G., Durbin R. (2009) *Bioinformatics* **25**: 2078.
Li H. and Durbin R. (2010) *Bioinformatics* **26**: 589.
The 1000 Genomes Project Consortium (2010) *Nature* **467**:1061.
Trapnell C., Pachter L., Salzberg S.L. (2009) *Bioinformatics* doi:**10**:1093/bioinformatics/btp120
VanRaden P., Olson K., Null D. and Hutchison J. (2011) *J Dairy Sci* **94**: 6153.
Zimin A., Delcher A., Florea L., Kelley D., Schatz M., Puiu D., Hanrahan .F, Pertea G., Van Tassell C., Sonstegard T., Marcais G., Roberts M., Subramanian P., Yorke J. and Salzberg S. (2009) *Genome Biology* **10**:R42

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:356-359

PRELIMINARY ANALYSIS OF INTENSITY SIGNALS FROM SNP DATA BASED ON POOLED DNA SAMPLES IN BEEF AND POULTRY

A. Reverter[1], J. Henshall[2], R. McCulloch[1], R. Hawken[3] and S.A. Lehnert[1]

[1]CSIRO Food Futures Flagship, 306 Carmody Rd., Brisbane, QLD 4067, Australia
[2]CSIRO Animal, Food and Health Sciences, Chiswick, Armidale, NSW 2350, Australia
[3]Cobb-Vantress Inc., 4703 U.S. Highway 412 East, Siloam Springs, Arkansas 72761-1030, USA

SUMMARY

Pooled genomic DNA has been proposed as a cost-effective means of conducting genome-wide association studies (GWAS) as they reduce the number of genotyping assays required. However, algorithms for genotype calling of biallelic SNP are not adequate with pooled DNA samples because they assume the presence of two fluorescent signals, one for each allele, and operate under the expectation that, at most, two copies of the variant allele can be found for any given SNP and DNA sample. We adapted analytical methodology developed originally for two-channel gene expression microarray technology and applied it to SNP genotyping of pooled DNA samples in three datasets. We show that both differential hybridization (green minus red intensity signals) and abundance (average of both signals) provide useful information in the prediction of SNP allele frequencies. This is particularly true when making inference about extreme SNP that are either nearly fixed or highly polymorphic. We demonstrate the use of a model-based clustering method via mixtures of bi-variate normal distributions to capture the relationship between hybridization intensity values and SNP allele frequencies from pooled DNA samples. We further show that when the SNP allele frequencies are known, either because the individuals in the pools or from a closely related population are themselves genotyped, a polynomial regression model with linear and quadratic components can be developed with high prediction accuracy.

INTRODUCTION

According to Craig *et al.* (2005), SNP allelic frequencies are approximated using a correction factor for the ratio of the intensity of A and B probes corresponding to the two alleles. The authors proposed a pooling-test statistic which is a function of the number of individuals in the pool and the number and standard deviation of the replicates. The approach was successfully employed by Pearson *et al.* (2007) and general issues regarding the feasibility of GWAS using pooled DNA samples was recently and comprehensively reviewed by the same authors (Szelinger *et al.* 2011).

Brohede *et al.* (2005) proposed a so-called polynomial-based probe-specific (PPC) algorithm:

$$f(A_j) = \beta_0 + \beta_1 x_j + \beta_2 x_j^2$$

where $x_j = A_j/(A_j+B_j)$ and A_j and B_j are the observed signal intensity values for A and B alleles, respectively. Using the PPC approach in pooled DNA samples, Anantharaman and Chew (2009) concluded that the algorithm is highly accurate and reproducible especially when a suitable reference sample set is used to estimate the beta values for PPC.

Recently, Henshall *et al.* (2012) explored the value of logistic regression of genotype on phenotype to estimate the effect of SNP genotype from pooled DNA samples. Various pooling strategies were explored and pooled genotypes generated *in silico* as the frequencies of alleles in animals in the pool. The authors concluded that pooling DNA from individuals within groups was superior to pooling DNA across groups.

The aim of this paper is to conduct an initial examination of the value of analysing intensity signals from SNP data based on pooled DNA samples in beef and poultry. Analytical approaches include a model-based clustering method and a polynomial regression of signal intensities.

Table 1. Description of pooled DNA samples employed in this study

Dataset	Species	Chips	Description
DATA1	Bovine	3	Proof of Concept: One, two and five DNA samples are pooled and genotyped to explore the resulting signals.
DATA2	Bovine	24	Bovine Stature: 24 pools each with seven DNA samples from a genotyped population of 1,193 Santa Gertrudis cows.
DATA3	Chicken	12	Chicken Pools: 35 individually genotyped chickens are pooled in groups of 5, 10 or 20 and the pools genotyped.

MATERIALS AND METHODS

Data and edits. Three datasets were employed with varying number of SNP chips from 3 to 24 (Table 1). All chips contained ~50K SNP designed for bovine and chicken DNA.

DATA1 – Proof of Concept. In order to explore the pattern of clusters in the intensity signals that emerge from SNP data using pooled DNA samples, we designed a simple experiment comprising three 50K SNP chips and bovine samples. For this initial proof of concept, a single DNA sample as well as DNA samples from pooling two and five whole blood samples were analysed.

DATA2 – Bovine Stature. Blood samples from 76 cows where used to create 11 pools. Each pool contained equal amounts of whole blood from 7 individuals pooled according to their stature so that individuals with similar height were pooled together. To allow for the measurement of technical variation, one of the pools was replicated. These 12 pools were assembled from whole blood frozen and thawed once. The same pool design was then replicated, but with whole blood frozen and thawed twice. These 76 individuals were part of a larger population of 1,193 cows previously individually genotyped with the same SNP chip.

DATA3 – Chicken. The blood from 35 chickens individually genotyped using the Illlumina Chicken 60K SNP chip was pooled in groups of 5, 10 or 20, and DNA extracted. Technical replicates were performed to obtain a total of 12 chips.

MA-Plots in SNP genotype data. In MA-plots, the y-axis containing "M" (for Minus) or difference between green and red intensity signals was plotted against "A" (for Average) in the x-axis. The base-2 logarithmic scale was used throughout. These plots are often employed in the context of gene expression data to check for the need for further normalization of the raw intensity signals and, most importantly, to identify genes differentially expressed. In the context of SNP data from truly biallelic SNPs and individual samples, the intensity signals are supposed to be either perfect green (eg. genotype AA), or perfect red (eg. genotype BB) or perfect yellow (eg. genotype AB). However, when pooled samples are used deviations from "perfect" green, red or yellow are expected from any given SNP due to possible genotype differences among the samples.

Model-based clustering. Model-based clustering via mixture of distributions has been proposed by a number of authors to analyse microarray gene expression data (see for instance Reverter *et al.* (2006) and references therein). In the present study, the MA-paired values of each SNP were subjected to model-based clustering via a mixture of an *n*-component mixture of bi-variate normal densities. Parameters of the mixture were estimated using EMMIX (McLachlan *et al.* 2002).

RESULTS AND DISCUSSION

Figure 1 shows the MA-Plots resulting from the analysis of DATA1 (Proof of Concept) along with the estimated distributions of the mixture models. When only the DNA of one individual is

genotyped (Figure 1A) the MA-Plot shows three distinct clusters corresponding to the three possible genotypes: upper cluster for homozygous AA with positive M-values, middle cluster for heterozygous AB with intermediate M-values, and lower cluster for homozygous BB with negative M-values. When the DNA of two individuals is pooled and genotyped, the resulting MA-Plot (Figure 1B) shows five distinct clusters for 0 to 4 copies of the variant allele, B. Finally, when 5 DNA samples are pooled the clusters get diffuse with monomorphic SNPs occupying the extremes in the scale of M-values (Figure 1C). Importantly, in all three cases, one novel finding is that the clusters with intermediate M-values are associated with higher A-values and this is reflected in the estimated means for the distributions of the mixture models.

Figure 2 shows the MA-Plots resulting from the analyses of DATA2 (Bovine Stature; Figure 2A) and DATA3 (Chicken Pools; Figure 2B). Overlaid in these plots are the SNP first allele frequencies (FAF) estimated from genotyping the individual DNA samples and colour-coded from red to yellow to green for low, intermediate and high FAF, respectively. These plots anticipate the strong relationship between the FAF and the MA-values resulting from genotyping pools.

In particular, when the FAF was analysed as a function of the MA-values, the following second-degree polynomial was obtained (R^2 = 86%):

$$FAF = -0.655 - 0.154M + 0.211A + 0.0015M^2 - 0.0091A^2$$

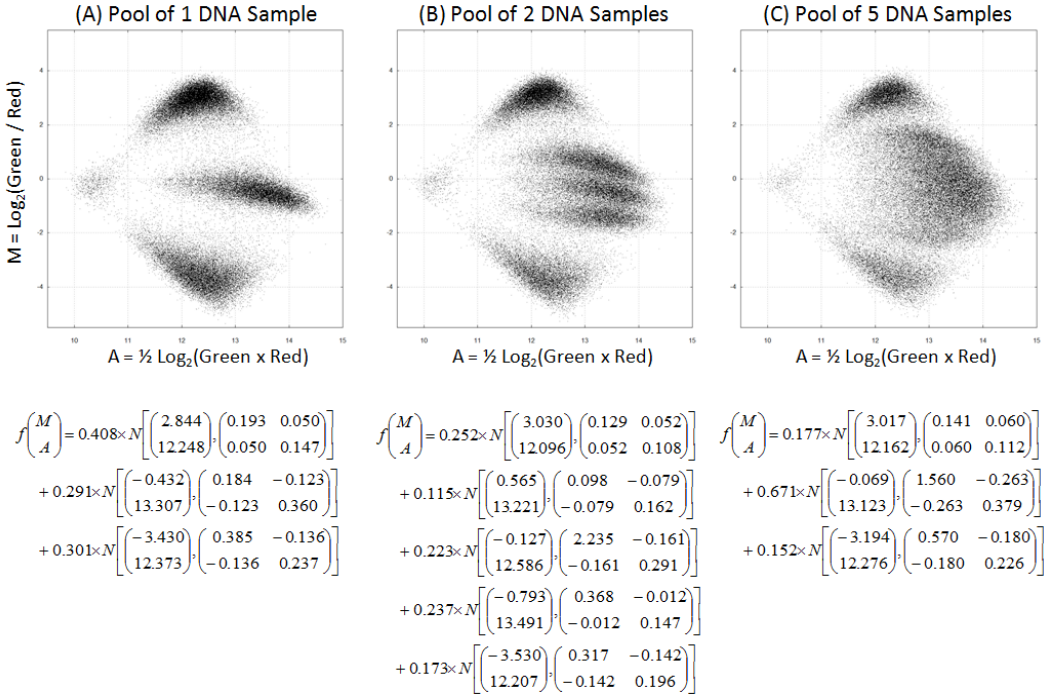


Figure 1. MA-Plots and model-based clustering via mixtures of distributions for the three chips of DATA1 – Proof of Concept: A: a single DNA sample; B: A pool of two DNA; C: A pool of five DNA samples.

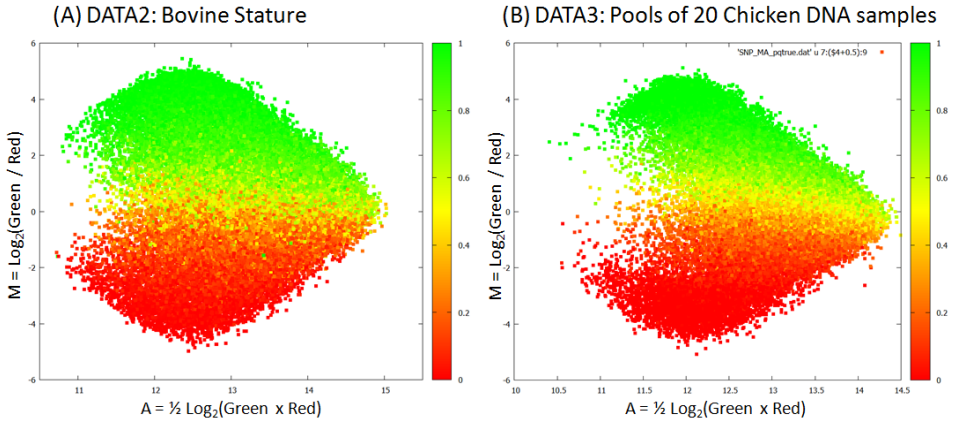


Figure 2. MA-Plots for DATA2 and DATA3 with overlaid estimates of first allele frequency from red (low frequency) to green (high frequency) based on genotypes of individual DNA samples.

CONCLUSIONS

The present study represents a first attempt to explore the numerical attributes of the intensity signals that should be considered when the intention is to genotype pools of DNA. We conclude that a strong relationship exists between the relative signal intensity of the two channels (red and green) and the SNP allele frequencies and show how this relationship can be formally ascertained by means of mixtures of distributions and polynomial equations. Further research is required to ascertain the extent to which model-based clustering and polynomial equations are suited to the use of pooled DNA samples to the development of application tools including estimation of family contributions to pools, SNP association to phenotypes and accurate genomic predictions.

ACKNOWLEDGMENTS

The work described under DATA2 was funded by a CSIRO Julius Career Award to Rachel J Hawken. The authors are grateful to Beef CRC project staff for collecting blood samples.

REFERENCES

Anantharaman R. and Chew F.T. (2009) *BMC Genet*. **10**:82.

Brohede J., Dunne R., McKay J.D. and Hannan G.N. (2005) *Nucl. Acids Res*. **33**:e142.

Craig D.W., Huentelman M.J., Hu-Lince D., Zismann V.L., Kruer M.C., Lee A.M., Puffenberger E.G., Pearson J.M. and Stephan D.A. (2005) *BMC Genomics* **6**:138.

Henshall J.M., Hawken R.J., Dominik S. and Barendse W. (2012) *Genet. Sel. Evol*. **44**:12.

McLachlan G.J., Bean R.W. and Peel D. (2002) *Bioinformatics* **18**:413.

Pearson J.V., Huentelman M.J., Halperin R.F., Tembe W.D., Melquist S., Homer N., Brun M., Szelinger S., Coon K.D., Zismann V.L., Webster J.A., Beach T., Sando S.B., Aasly J.O., Heun R., Jessen F., Kolsch H., Tsolaki M., Daniilidou M., Reiman E.M., Papassotiropoulos A., Hutton M.L., Stephan D.A. and Craig D.W. (2007) *Am. J. Hum. Genet*. **80**:126.

Reverter A., Ingham A., Lehnert S.A., Tan S.H., Wang Y., Ratnakumar A. and Dalrymple B.P. (2006) *Bioinformatics* **22**:2396.

Szelinger S., Pearson J.V. and Craig D.W. (2011) *Methods Mol. Biol*. **700**:49.

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:360-363

A GENOMIC PREDICTION CROSS-VALIDATION APPROACH COMBINING EWE REPEATED PHENOTYPES AND RAM DAUGHTER TRAIT DEVIATIONS

H.D. Daetwyler[1,2], S. Bolormaa[1,2], D. J. Brown[2,3], J.H.J. van der Werf[2,4] and B. J. Hayes[1,2,5]

[1]Biosciences Research Division, Department of Primary Industries, Bundoora 3083, Victoria, Australia, [2]Cooperative Research Centre for Sheep Industry Innovation, Armidale 2351, NSW, Australia, [3]Animal Genetics and Breeding Unit (AGBU), Armidale 2351, NSW, Australia
[4]School of Environmental and Rural Science, University of New England, Armidale 2351, NSW, Australia,[5]La Trobe University, Bundoora 3086, Victoria, Australia

SUMMARY

Reproduction traits generally are expressed later in life and have high economic value, which makes them good candidates for improvement through genomic selection. We derive genomic prediction equations for number of lambs weaned (NLW) from genotyped rams' daughter records and repeated ewe records for genotyped ewes to make the best use of all available information. In order to assess the accuracy of the genomic estimated breeding values (GEBV) for this trait, an across sire family cross-validation design is proposed, and was compared with fully random cross-validation. The accuracy of genomic BLUP (GBLUP) using both ram and ewe records is compared with GBLUP using only ewe records, as well as BLUP with no genomic information using both ram and ewe records. The combined approach resulted in higher accuracies than both GBLUP only using ewe records, and BLUP. The approach provides a way to make use of all data available to maximise accuracy of GEBVs for traits such as NLW.

INTRODUCTION

Although reproduction traits, such as number of lambs weaned (NLW) have a high economic value, they are difficult to genetically improve, due to low heritability (h^2) and because they are expressed later in life. Genomic prediction may increase genetic gain in these traits because it can predict a ram or ewe's performance early in life with, potentially, a higher accuracy than Best Linear Unbiased Prediction (BLUP).

Genomic prediction makes use of a reference population in which animals are both genotyped and phenotyped to predict genomic estimated breeding values (GEBV) of selection candidates (e.g. young rams) based only on genotype. The individuals available with both genotypes and phenotypes are primarily ewes with repeated NLW records in large research projects. In addition, genotyped industry rams have daughters with NLW records. Combining these two sources of information should make optimal use of the data available. However, each source needs to be properly weighted to account for the differences in phenotype accuracy.

The aim of this paper was to compare strategies for combining daughter information on genotyped rams with individual repeated observations on genotyped ewes in a reference population for NLW genomic predictions. Accuracy of the resulting GEBV was assessed with a number of different cross-validation strategies.

METHODS

Phenotypes, genotypes and estimates of heritability and repeatability. Proper weighting of ewe and ram records in this combined genomic prediction approach required the estimation of h^2 and repeatability (t). Data on NLW from years 1992 to 2012 were retrieved from the Australian Sheep Genetics database, giving 290,636 records on 244,672 ewes, where 111,572 ewes had a known sire (8,036 sires). The number of progeny per sire ranged from 1 to 339. The phenotype data included 53 breeds and each animal's breed proportions were calculated from the pedigree.

The main breeds were Merino (MER), Border Leicester (BL), Coopworth, Polled Dorset and White Suffolk. Values for h^2 and t were estimated using $\mathbf{y} = \mathbf{Xb} + \mathbf{Z_1Qq} + \mathbf{Z_1a} + \mathbf{Z_2pe} + \mathbf{e}$, where **y** is a vector of phenotypes, **X**, $\mathbf{Z_1}$, and $\mathbf{Z_2}$ are design matrices, **b** is vector of fixed effects, **q** is a vector breed effects**, a** is a vector of animal effects, **pe** is a vector of permanent environmental effects, and **e** is the vector of random errors. The following distributions were assumed: **a** ~ N (0, σ_a^2 **A**), **q** ~ N (0, σ_q^2 **I**), and **e** ~ N (0, σ_e^2 **I**), where **A** is the numerator relationship matrix, σ_a^2 is the genetic variance, σ_q^2 is the variance of breed effects, and σ_e^2 is the residual variance. Fixed effects included the mean, conception site, lambing site, year of lambing, age at lambing, conception method, and an indicator of whether a ewe lambed as a yearling.

A subset of the animals were genotyped using the Ovine50K SNP chip comprising a total of 54,977 single nucleotide polymorphism (SNP). The genotype quality control and imputation of sporadic missing genotypes is described in Daetwyler *et al.* (2012) and reduced the number of SNP to 48,599. The final multi-breed reference population included 4114 animals (of which 317 were rams), including 2,205 MER, 788 BL, 486 from terminal breeds, 185 other maternal breeds, 185 other MER type breeds, and 283 of unknown breed. The genotyped ewes originated from the Sheep CRC information nucleus and the SheepGENOMICS project (van der Werf et al. 2010; White et al. 2012). Results are reported for the MER and BL breeds and their crosses, as they made up the majority of genotyped animals.

Trait and daughter trait deviations. The phenotypes for NLW used in GBLUP analyses were trait deviations (TD) for ewes and Daughter Trait Deviations (DTD) for rams. The phenotype was corrected for fixed effects using the same model applied to calculate heritability and repeatability but excluding the animal effect. The residuals (corrected phenotype) were used to calculate TD and DTD. Trait deviations are calculated as TD = Σ(residual)/N, where N is the number of records. Rams DTD are calculated only from their ungenotyped daughters' TD: DTD =Σ(TD)/p, where p is the number of progeny per ram. Genotyped ewes are included in reference population and not used in their ram's DTD to avoid double counting. Rams with less than 3 progeny were removed. The number of records contributing towards genotyped ewe TD and ram DTD was 6066 and 9213, respectively, demonstrating that adding the sire DTD more than doubled the reference population. DTD contain only the genetic merit of the sire, thus DTD was doubled for analysis. The TD and DTD were weighted in the model to account for the differential accuracy of phenotypes using $\left(1-h^2\right)\left[ch^2+\left(\left(1+(n_i-1)t\left[n_i\right]^{-1}\right)-h^2\right)\right]^{-1}$ and $\left(1-h^2\right)\left[ch^2+(4-h^2)\left[p_i\right]^{-1}\right]^{-1}$, respectively, where n_i is the number of records for animal *i*, *c* is the proportion of the genetic variance *not captured* by the markers, and *p* is the number of progeny for ram *i* (Garrick et al. 2009). Five different values of the c (0.25, 0.35, 0.50, 0.60 and 0.75) were used for calculating weights of TD and DTD.

Genomic prediction analysis. Genomic BLUP (GBLUP) was used to predict GEBVs and BLUP was used for predict estimated breeding values (EBV). GEBVs were calculated based on the following model: $\mathbf{y^*} = \mathbf{1\mu} + \mathbf{Xb} + \mathbf{Zg} + \mathbf{e}$, where $\mathbf{y}^*$ is a vector of TD and DTD, **1** is a vector of ones, **μ** is the mean, **b** is vector of fixed effects and included sex, **X** and **Z** are design matrices, **e** is the vector of random errors, and **g** is a vector of either GEBVs or EBVs. In GBLUP, **g** it is distributed as N (0, σ_g^2 **G**), where σ_g^2 is the genetic variance explained by the markers and **G** is the genomic relationship matrix (Yang et al. 2010). All models were run in ASReml.

Measuring accuracy with cross-validation. Cross-validation, where the data is divided into a number of subsets and each subset is predicted once from the other subsets, was used to estimate accuracy of GEBVs. The six subsets for cross-validations were chosen either completely random or by random whole sire family. In random sire family cross-validation, sires were randomly

allocated to subsets. All progeny of a sire (both genotyped and ungenotyped) was then allocated to the same subset to ensure prediction was across sire families resulting in a conservative estimate of prediction accuracy. The division of subsets were the same for GBLUP and BLUP. The values presented are the mean of five replicated cross-validations, where new random subsets were chosen in each replicate.

Accuracies were calculated as the correlation of GEBV with **y*** for MER and BL breeds, where animals were assigned to breed groups according to sire breed. The accuracy of the true breeding value was approximated by dividing this correlation by the mean accuracy of the EBVs of sires and ewes in the reference population. The sires' EBV accuracy was calculated using only their non-genotyped daughters and ewe accuracies were from the BLUP model used to calculate h^2 and calculated as $\sqrt{\left(1-SE^2\right)/\sigma_a^2}$, where SE^2 is the standard error of prediction for the EBV. Potential bias of GEBV or EBV was investigated by regression of **y*** on GEBV or EBV.

RESULTS AND DISCUSSION

The h^2 and t estimated with full BLUP model for NLW was 0.06 and 0.08, respectively. The estimate for h^2 was consistent and the estimate for t was slightly lower than literature estimates (Safari et al. 2005). In cross validations, the accuracy of GBLUP was always higher than BLUP for both breeds (Table 1). BLUP accuracy was higher when cross-validation was random, due to prediction within families, but deteriorated significantly when cross-validation was across sire families, as expected. Adding the ram DTD into the analysis increased the accuracy in all scenarios showing a clear benefit of making use of all available data. The increase in GBLUP accuracy of combined data was greater when cross-validation was across sire families, indicating that combining data had a greater effect on animals less related to the reference. The average size of validation sets was 278 for MER and 99 for BL.

Table 1. Accuracy and slope of regression(y*, GEBV) of BLUP and GBLUP, when using only trait deviations of ewes (TD) or TD plus daughter trait deviations of rams (DTD)

		Accuracy			Slope		
Breed	Cross Validation	BLUP TD+DTD	GBLUP TD	GBLUP TD+DTD	BLUP TD+DTD	GBLUP TD	GBLUP TD+DTD
MER	Random	0.19	0.21	0.24	1.12	1.37	1.34
	Sire Fam	-0.05	0.06	0.12	-1.65	1.35	1.55
BL	Random	0.15	0.19	0.22	1.10	2.47	1.61
	Sire Fam	0.10	0.20	0.27	2.23	3.88	2.83

The slopes of the regression of **y*** on GEBV were variable. A slope > 1 (i.e. the GEBV under-estimated the phenotype) was observed in most analyses. Sire family cross-validation resulted in more bias than random subsets. BL slopes exhibited more upward bias than MER. It is of note that the slopes in BL based on TD alone showed a stronger upward bias than the combined data. One reason for this could be that the genotyped research ewes were actually BL/MER crosses exhibiting heterosis, whereas the daughters of the industry sires are expected to be more purebred BL. The bias in BL could also be due to the scaling of the multi-breed **G**, which can affect variance components. Further work will investigate fitting heterosis and scaling **G** for breed specific inbreeding and allele frequencies (Erbe et al. 2012). Using the current approach, the GEBVs would have to be blended with BLUP breeding values to be distributed to sheep breeders. A one-step approach (e.g. Aguilar et al. 2010) would also make use of all data and would reduce

issues related to blending. However, the combination of **A** and **G** matrices also needs to account for the proportion of the genetic variance not captured by the markers, just like *c* in our approach.

The weights on TD and DTD require an assumption on *c*, defined as the proportion of the genetic variance *not captured* by the markers (Garrick et al. 2009). The true value of *c* depends on the interplay between marker density and the effective population size (N_e) of the breed. A breed with a higher N_e (e.g. MER N_e 800, (Kijas et al. 2012)) would need a higher marker density than a breed with a lower N_e (e.g. BL N_e 150) to achieve the same *c*. The range of GEBV accuracy from different *c* values was small for both MER and BL (maximum 3%). This lack of a clear signal could be due to the limited size and multi-breed nature of the reference population and the low h^2 of NLW.

The accuracies achieved using GBLUP are encouraging and more genetic gain would be achieved for NLW through genomic selection than BLUP. Furthermore, making use of all data on both ewes and rams substantially increased the accuracy of prediction.

ACKNOWLEDGMENTS

The authors acknowledge funding from the CRC for Sheep Industry Innovation, MLA, and AWI Ltd. We thank K. Gore and K. Geenty for managing the CRC information nucleus database, A. Swan for comments and the many staff involved at the CRC sites across Australia.

REFERENCES

Aguilar I., Misztal I., Johnson D.L., Legarra A., Tsuruta S. and Lawlor T.J. (2010) *J. Dairy Sci.* **93**:743.

Daetwyler H.D., Swan A.A, van der Werf J.H.J., and Hayes B.J. (2012) *Gen. Sel. Evol.* **44**:33.

Erbe M., Hayes B.J., Matukumalli L.K., Goswami S., Bowman P.J., Reich C.M., et al. (2012) *J. Dairy Sci.* **95**: 4114.

Garrick D., Taylor J. and Fernando R. (2009) *Gen. Sel. Evol.* **41**:55.

Kijas J.W., Lenstra J.A., Hayes B., Boitard S., Porto Neto L.R., San Cristobal M., et al. (2012) *PLoS Biol.* **10**:e1001258.

Safari E., Fogarty N.M. and Gilmour A.R. (2005) *Livest. Prod. Sci.* **92**: 271.

van der Werf J.H.J., Kinghorn B.P. and Banks R.G. (2010) *Anim. Prod. Sci.* **50**: 998.

White J.D., Allingham P.G., Gorman C.M., Emery D.L., Hynd P., Owens J., et al. (2012) *Anim. Prod. Sci.* **52**: 157.

Yang J., Benyamin B., McEvoy B.P., Gordon S., Henders A.K., Nyholt D.R., et al. (2010) *Nat. Genet.* **42**: 565.

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:364-367

THE EXTENT AND DISTRIBUTION OF LINKAGE DISEQUILIBRIUM IN EXTENSIVELY RAISED CHICKEN POPULATIONS OF SOUTHERN AFRICA

K.S. Khanyile[1,2], E. F. Dzomba[2] and F. C. Muchadeyi[1*]

[1]Biotechnology Platform, Agricultural Research Council, P Bag X5, Onderstepoort, 0110, South Africa
[2]University of KwaZulu-Natal, Discipline Genetics, School of Life Sciences, Private Bag X01, Scottsville 3209, Pietermaritzburg, South Africa
Correspondence: MuchadeyiF@arc.agric.za

SUMMARY

The amount of linkage disequilibrium (LD) is an important source of information about historical events of recombination and allows inferences about genetic diversity and genomic regions that have undergone selection. Linkage disequilibrium is equally important in studying effective population size and rate of inbreeding particularly in extensively raised and wild animal populations where pedigree records are scarce. The objective of this study was to investigate LD in village chicken populations of Southern Africa. These chickens are raised under scavenging systems of production characterized by uncontrolled breeding and frequent population bottlenecks due to disease outbreaks and fluctuations in feed supplies. DNA samples from 312 extensively raised chickens from South Africa, Malawi and Zimbabwe were genotyped using the Illumina iSelect chicken SNP60K BeadChip. A panel of 43,157 out of the total 57,636 (74.8%) SNPs was used in the final analysis after screening for those that had a minor allele frequency of less than 5%, were out of Hardy-Weinberg equilibrium ($P<0.01$) and had a call rate of less that 95%. Results indicated that LD averaged between 0.45 and 0.58 for SNPs that had a pairwise distance of less than 20 kb. LD dropped to 0.34 for SNPs between 20 and 100 kb after which it remained constant. LD was further analyzed for its decay over marker distance and differences between populations from different geographic locations. Results are discussed in terms of historical changes in effective population size and resultant recombination rates. The utility of the iSelect chicken SNP60K beadchip in investigating free-range chicken population genetics is demonstrated.

INTRODUCTION

Linkage disequilibrium (LD) is defined as a non-random association of alleles at two or more loci (Hendrik 2005; Qanbari *et al*. 2010). The importance of LD is in providing information about historical events of recombination thereby explaining genetic diversity in genomic regions undergoing selection. LD also allows estimation of effective population size and rate of inbreeding in extensively raised and wild animal populations without pedigree records (Wragg *et al*. 2012).

The village chicken production system in Africa is mainly based on scavenging village chickens (Kitalyi 1998), that are used to meet the multiple household social, economic and cultural needs and are crucial to biodiversity (Delany 2003). However, very little is known about the genetic composition of village chickens in developing regions like Southern Africa. Diversity studies using autosomal microsatellite (Muchadeyi *et al*. 2007) and mtDNA sequences (Mtileni *et al*. 2011) have not defined the genetic stability of these populations. Demographic population parameters such as effective population size and inbreeding levels, that influence the risk to extinction of these populations, remain uncharacterized due to the absence of pedigree and other population census records in these village chicken production systems. The availability of large-scale sequence data in chickens has resulted in an increase in the marker density and achieved a

comprehensive SNP coverage of the chicken genome. The chicken 60K SNP genotyping chip has the potential to unravel the genetic information in extensively raised chicken populations. Applying LD analysis will permit estimation of demographic and evolutionary parameters of these populations. The aim of this study was to investigate the extent and distribution of LD in extensively raised chicken populations of South Africa, Zimbabwe and Malawi using the Illumina iSelect chicken SNP60K BeadChip.

MATERIALS AND METHODS

Chicken populations, blood collection and DNA isolation. A total of 312 village chicken samples were collected from South Africa (n =147), Malawi (n = 30) and Zimbabwe (n = 135). In South Africa, village chickens representing Limpopo (n = 15), Eastern Cape (n = 26) and Northern Cape (n = 35) populations, and four conservation flocks of the Naked Neck (n = 20); Potchefstroom Koekoe (n = 20); Ovambo (n = 10) and Venda (n = 20) chickens kept at Agriculture Research Council Poultry Breeding Resource, were sampled as described in Mtileni et al. (2011). The sampling of the village chickens from Zimbabwe (n = 135) and Malawi (n = 30) populations is described in Muchadeyi et al. (2007). Blood was collected from the selected chickens onto FTA Micro Cards (Whatman Bio Science, UK) and DNA was isolated using a modified protocol of the Qiagen® DNA blood and tissue kit.

SNP genotypes and quality control. The chicken DNA samples were genotyped using the iSelect chicken SNP60K bead chip produced by Illumina Inc. SNP quality control was done using Plink (1.07) software to remove SNPs that were either out of Hardy-Weinberg equilibrium (HWE) *(P< 0.01)*, showing a minor allele frequency (MAF) of at least 5%, had low call rate (< 95%) and with missing genotypes (> 5%). SNPs that were on unknown chromosomes, mtDNA, linkage groups and/or sex chromosomes were excluded from further analyses. After filtering, 45676, 44667,46905 and 43157 SNPs on 28 autosomal chromosomes were used for each of the Malawi, South Africa, Zimbabwe and combined populations, respectively.

Linkage Disequilibrium analysis. A pair-wise LD (r^2) was estimated using PLINK (1.07) software for SNPs on chromosome 1 to 28 for the individuals belonging to the three populations using the following formula:

$$r^2 = \frac{(f11f22 - f12f21)^2}{fA1fA2fB1fB2}$$

A Generalized Linear Model procedure (Proc GLM) in the Statistical Analysis System (SAS) was used to determine the effects of SNP marker interval (bp), chromosome, and population group and interaction of chromosome-by- population on the decay of LD using the following model:

$$r^2_{ij} = \mu + Pop_i + Gga_j + (Pop \times Gga)_{ij} + b\text{SNPint} + e_{ik},$$

Where: Pop_i was the effect of *i*th chicken population of either, Malawi, Zimbabwe or South Africa; Gga_j was the effect of the *j*th chromosome 1-28; and SNPint the effect of SNP interval fit as a covariate with *b* the regression coefficient.

RESULTS AND DISCUSSION

Effects of chicken population, chromosome and distance between SNPs on LD. LD was calculated on 28 of the 38 chicken autosomes. The chromosome size, SNP interval distance and number of SNPs per chromosome support the differences between macrochromosome 1-5 that had high number of SNPs and large intervals between SNPs and micro-chromosomes 16-28, which are smaller and had less SNPs that were relatively close together (Megens *et al*. 2009). Linkage disequilibrium (r^2±SD) averaged 0.38 ± 0.20 and ranged from 0.34 ± 0.14 -0.45 ± 0.24 in Malawi,

0.34 ± 0.15 - 0.52 ± 0.27 in Zimbabwe and 0.34 ± 0.14 - 0.50 ± 0.27 in South African chicken populations. Overall, there was no significant difference in r^2 values ($P<0.05$) between populations indicating similarities between the Malawian, Zimbabwean and South African village chicken populations. However, LD varied significantly between chromosomes ($P<0.001$) with chromosome 8 having the highest LD of 0.52 ± 0.26 followed by chromosome 22 with an r^2 ± SD value of 0.49 ± 0.28. The high LD might be an indication of selection at genes on these chromosomes (Hendrick 2005) particularly natural selection pressures as these chicken populations are raised under extensive systems of production where human selection pressures are minimal (Mtileni *et al*. 2010). Although population did not influence genome-wide LD, a population by chromosome interaction was observed whereby the Zimbabwean chicken population had the highest LD on chromosome 8 (0.52 ± 0.267) and the South African chicken population was highest on chromosome 22 (0.49 ± 0.29). Such interactions need to be further investigated as they might indicate different selection pressures in different populations (Wragg *et al*. 2012).

Another factor that influenced LD was the SNP interval. To further understand this, LD was computed at different distance interval of 0-1 kb, 1-10 kb, 10-20 kb, 20-40 kb, 40-60 kb, 60-100 and 100kb plus using SNP data from chromosomes 1-28 (Fig 1a) and from chromosomes 8; 22 and 13 as indicated in Figures 2b, c and d respectively.

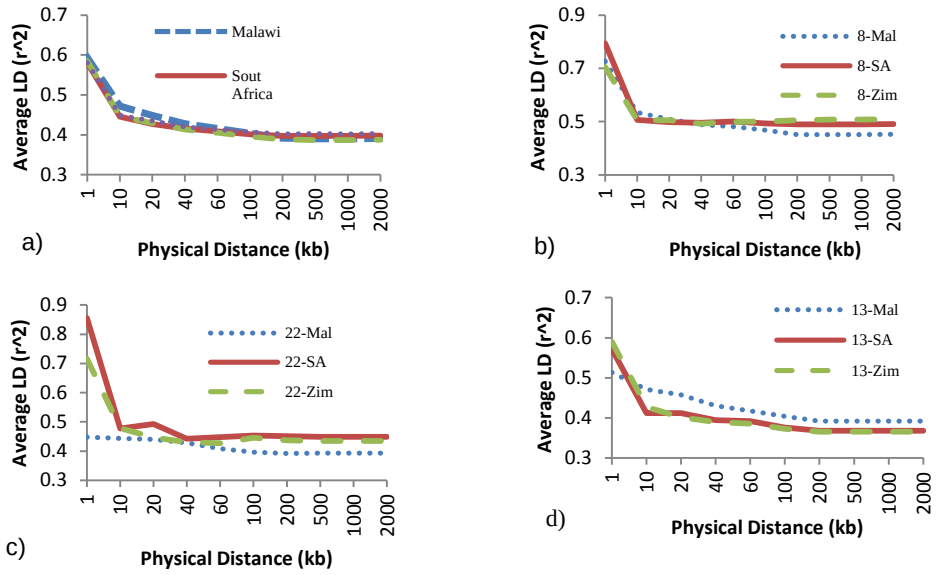


Figure 1. Average LD decay with an increase in physical distance between SNPs for a) chromosomes 1-28, b) chromosome 8; c) chromosome 22; and d) chromosome 18.

The LD averaged 0.58 for SNPs within a 10 kb interval and decayed to 0.45 -0.47 for SNPs between 10-30 kb after which they remained constant. The LD decay at chromosome 8 of the Malawi chickens continued to decline after 40kb. In Zimbabwe and South African chickens, LD at chromosome 22 made a sharp decay from 0.7 (Zimbabwe) and 0.85 (South Africa) to an r^2 below 0.5 at 10kb after which it stayed constant. On the same chromosome LD was maintained around 0.45 over all sliding windows in the Malawi chicken population.

Overall, a higher LD was observed in the Southern African chicken populations compared to

other chicken populations observed in other studies (Qanbari *et al*. 2010; Wragg *et al*. 2012). For example, in a commercial egg laying flock, r^2 averaged 0.32 ± 0.33 with a minimum 0.21 ± 0.26 (Qanbari *et al*. 2010) whereas it was maintained around 0.38 in this study.

CONCLUSION

A relatively high LD that persisted over long SNP intervals was observed in the South African, Zimbabwean and Malawian chicken populations. This LD pattern seems to be consistent with low and steady effective population sizes. The study recommends for a further investigation on the role of selection and population bottlenecks on chromosomes 8 and 22 that had significantly high LD.

ACKNOWLEDGEMENTS

Genotyping of samples was funded by the Agricultural Research Council-Biotechnology Platform (ARC-BTP) while Mr KS Kanyile held fellowships from the ARC-Professional Development Program and the University of KwaZulu-Natal postgraduate program.

REFERENCES

Delany M.E., (2004) *Mec. of Dev*. **121**: 1169.

Hendrick P. W. (2005) Genetics of populations 3rd. Jones and Bartlett publishers, Canada

Illumina Technical notes: DNA analysis, Analyzing GoldenGate Genotyping Data

Kitalyi A. J. (1998). Village chicken production systems in rural Africa: Household food security and gender issues. FAO Animal Production and Health Paper 142, Rome, Italy.

Megens H., Crooijmans R.P.M.A., Bastiaansen J. V. M., Kersten H. H. D., Coster A., Jalving R., Vereijken A., Silva P., Muir W. M., Chen H. H., Hanotte O. and Groene M. A. M. (2009) *BMC Genetics*. **10**: 86.

Mtileni B. J., Muchadeyi F. C., Maiwashe A., Groeneveld E., Groeneveld L. F., Dzama K. and Weigend S. (2011) *Journal of Animal Breeding and Genetics*. **128**: 209.

Mtileni B. J., Muchadeyi F. C., Maiwashe A., Phitsane P. M., Halimani T. E., Chimonyo M. and Dzama, K. (2010) *Applied Animal Husbandry and Rural Develop*. **2**:18.

Muchadeyi F. C., Eding H., Wollny C. B. A., Groeneveld E., Makuza S. M., Shamseldin R., Simianer H. and Weigend S. (2007) *Animal Genetics*. **38**: 332.

Qanbari S., Hansen M., Weigend S., Rudolf P. and Simianer H. (2010) *BMC Genetics*. **11**: 471.

Wragg D., Mwacharo J., Alcalde J., Hocking P. and Hannotte O. (2012) *Heredity*. **109**: 6.

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:368-371

MICRORNA PROFILING IN CATTLE DIVERGENTLY SELECTED FOR RESIDUAL FEED INTAKE

W. AL-Husseini[1,4], C. Gondro[1], R.M. Herd[2], J.P. Gibson[1], and P.F. Arthur[3] and Y. Chen[3]

[1]The Centre for Genetic Analysis and Applications, School of Environmental and Rural Science, University of New England, Armidale, NSW 2351
[2]NSW Department of Primary Industries, Beef Industry Centre, Armidale, NSW 2351
[3]NSW Department of Primary Industries, Elizabeth Macarthur Agricultural Institute, Menangle, NSW 2568
[4]Department of Biology, College of Science, University of Babylon, Babylon, Iraq

SUMMARY

MicroRNAs (miRNAs) are short non-coding RNAs that post-transcriptionally regulate expression of mRNAs in many biological pathways. Here we report comprehensive miRNA profiles by deep sequencing in Angus cattle divergently selected for residual feed intake (RFI). Two miRNA libraries were constructed from pooled RNA extracted from livers of low and high RFI cattle, and sequenced with the Illumina Genome Analyser. We identified 305 known bovine miRNAs. bta-miR-143, bta-miR-30, bta-miR-122, bta-miR -378 and bta-let-7 were the top 5 most abundant miRNA families expressed in liver, representing more than 63% of expressed miRNAs. Mir-143 is the most expressed bovine miRNA in liver, and is up-regulated in high RFI cattle. Mir-122 is the second most expressed miRNA in liver and is down regulated in high RFI animals. The differentially expressed miRNAs may play important roles in the regulation of the bioprocesses responsible for variation in RFI in cattle.

INTRODUCTION

MicroRNAs (miRNAs) are small (~ 22 nucleotides) non-coding RNA that regulate gene expression by targeting mRNA in a sequence-specific manner, leading to either translational repression or degradation of the targeted transcript. MicroRNAs are now known to repress thousands of target genes and regulate cellular processes, including cellular proliferation, differentiation and apoptosis. The aberrant expression or alteration of miRNAs also contributes to a range of human pathologies, including diabetes and cancer (Lu *et al*. 2005).

Residual feed intake (RFI) is a measure of feed efficiency in beef cattle. It is the difference between an animal's actual feed intake recorded over a test period and its expected feed intake based on its size and growth rate (Koch *et al*. 1963). Genome wide association studies have been used to identify gene markers associated with RFI in beef cattle. More than a hundred single nucleotide polymorphisms (SNP) have been reported as being associated with RFI (Barendse *et al*. 2007). However, a large proportion of these SNP are not located in annotated genic regions of the bovine genome. Some of the most significant SNP for RFI were in or close to miRNA motifs which suggests that these miRNAs could play an important role in RFI variation (Barendse *et al*. 2007).

Considerable progress has been made in the characterization of miRNAs in livestock genomes over the last decade, and a wide and diverse range of conserved and species-specific miRNAs have been identified. However, little is known about their role in regulation of key cellular and physiological pathways involved in feed efficiency and RFI. Liver is a central controller of metabolism and a major driver of whole animal oxygen consumption in mammals. In this study we profiled miRNAs abundance in liver tissue of Angus bulls from high and low RFI selection lines using a deep sequencing approach. Here we report the first liver miRNA profile of known and

putative novel bovine miRNAs. Differentially expressed miRNAs between high and low RFI selection lines are also discussed.

MATERIALS AND METHODS

Animals. Young Angus bulls resulting from approximately three generations of divergent selection for RFI were used in this study. The selection lines were established in 1993 at the Agricultural Research Centre, Trangie, NSW, Australia (Arthur *et al*. 2001). Bulls were born in 2005 and, at approximately one year-of-age, their growth and feed intake were measured. Postweaning RFI for each animal in the test group was calculated using a linear regression model of feed intake on mean metabolic live weight and average daily gain. Based on the RFI test results, liver biopsies were taken from 30 animals with the lowest RFI and 30 animals with the highest RFI as described by Chen *et al.* (2011). Total RNA from liver tissue was isolated using TRI reagent (Ambion, Applied Biosystems) according to the manufacturer's instructions.

Small RNA library construction and analysis of small RNA sequencing data. Based on availability and the quality of RNA, two pools of total RNA were constructed from 13 high RFI animals and 13 low RFI animals with equal quantities (1 µg) from each animal. Libraries of small RNAs were prepared using a Small RNA Sequencing kit (Illumina) and sequenced by Illumina Genome Analyser. Sequencing data were analysed using miRanalyzer (Hackenberg *et al*. 2011). In brief, known bovine miRNAs were identified by mapping all sequence reads to known bovine miRNAs in miRBase (version19), and reads that matched known bovine miRNAs were grouped and removed from the dataset. Reads that mapped to known miRNAs in other species were grouped as homologue miRNAs. The remaining reads were aligned to libraries of known transcripts. To identify bovine-specific novel miRNAs, the remaining sequence reads were mapped to *Bos taurus* genome (bostau6, UMD_3.1) using Bowtie. Mapped reads were first clustered into putative mature miRNAs and pre-miRNAs. The putative candidate miRNAs were reported based on at least three out of five different Random Forest models (Hackenberg *et al*. 2009). To compare the differentially expressed miRNAs between the two libraries (low and high RFI), the expression of each specific miRNA (read counts) were normalised to percentage of million mapped reads (PMMR) and fold-changes were calculated between the high and low RFI pools.

RESULTS

There were 10,820,087 and 12,808,022 high quality sequence reads for the high RFI pool and the low RFI pool, respectively. Approximately half of these reads were an exact match to known bovine mature miRNAs (Figure 1). A total of 304 known miRNAs were detected as being expressed in bovine liver. Table 1 lists the differentially expressed miRNAs between high and low RFI animals along with their means for phenotypic traits. We defined the most abundant miRNAs as those expressed in more than 1% of the mapped miRNAs. Medium abundance miRNAs were those between 0.1-1%. Two-fold expression changes between high and low RFI was considered differentially expressed. Generally, the reads of the most abundant miRNAs were more than 100,000 fold higher than those of the scarce miRNAs. The 5 most expressed miRNA families were bta-miR-143, bta-miR-30, bta-miR-122, bta-miR-378 and bta-let-7, which constituted more than 63% of the total sequence reads, suggesting that they are the most abundantly expressed miRNAs in bovine liver tissue. In total 52 miRNAs homologous with other species and not listed in the bovine miRBase (version 19) was identified based on the precursor sequence.

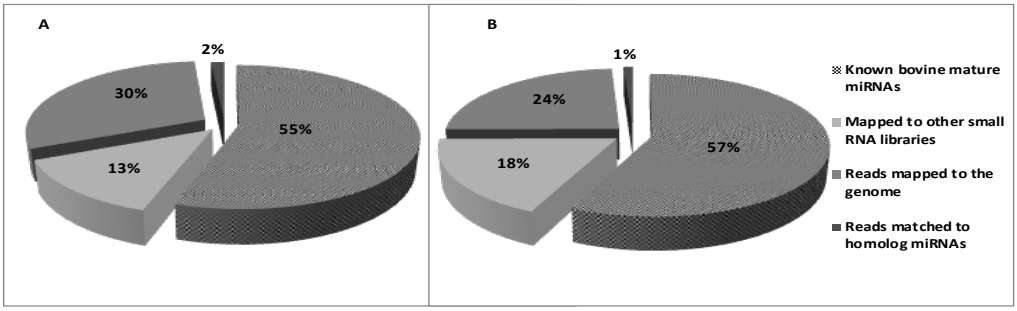
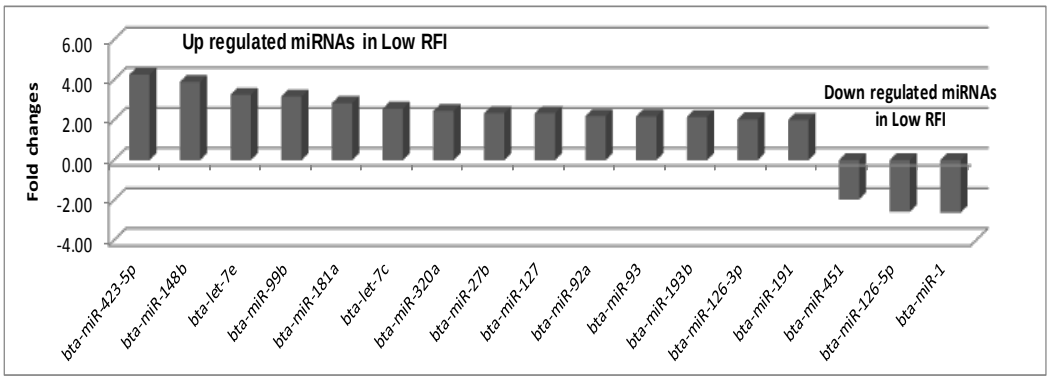


Figure 1. Distribution of mapped high quality reads in both pools: low RFI line (A), and high RFI line (B).

Table 1: Means for animal traits and differences in the top most abundant known miRNAs between low RFI pool and high RFI pool

	RFI (kg/day)	Feed intake(kg/day)	Average daily gain (kg/day)	Rib fat depth (mm)
Low RFI	-1.07	10.49	2.12	7.68
High RFI	0.99	12.63	1.98	10.73
Top 4 up regulated miRNAs in low RFI				
	bta-let-7b	bta-miR-122	bta-miR-30d	bta-let-7a-5p
Low RFI	2.25	12.29	2.25	5.96
High RFI	0.81	4.82	1.06	2.90
Fold change	2.79	2.55	2.12	2.06
Top 4 down regulated miRNAs in low RFI				
	bta-miR-143	bta-miR-192	bta-miR-21-5p	bta-miR-101
Low RFI	15.9	2.43	0.90	1.27
High RFI	31.6	5.19	2.33	4.1
Fold change	-1.99	-2.13	-2.59	-3.22

Figure 2. Fold changes of the medium abundant bovine miRNAs differentially expressed between high and low RFI cattle.



Ten putative novel bovine-specific miRNAs, based on precursor sequence and secondary structure were found. Those putative novel bovine-specific miRNAs candidates were expressed by 6,534 read counts (ranging from 2102 to 22 reads) in both RFI pools.

DISCUSSION

Accumulating evidence shows miRNAs play important regulatory roles in physiological and developmental processes in many tissues, including liver. We found that the most expressed miRNA in bovine liver is bta-mir-143 which constituted of 20% of total expressed miRNAs. This is a clear difference from the miRNAs expression pattern in human and mouse liver where the most expressed miRNA is mir-122 (Rottiers and Näär, 2012). Besides the discovery of novel miRNAs in bovine liver, we have identified 18 miRNAs up-regulated in low RFI and 7 miRNAs down-regulated in low RFI cattle. This is consistent with previous mRNA expression with microarrays, in which there were more genes up-regulated in high RFI animals than low RFI animals (Chen *et al.* 2011). These differentially expressed miRNAs may play important roles in regulation of the physiological processes involved in RFI in beef cattle. For example, bta-mir-143 is up-regulated in high RFI animals. Mir-143 was up-regulated in liver of genetic and dietary mouse models for obesity. Overexpression of miR-143 impairs insulin-stimulated AKT activation and glucose homeostasis (Jordan *et al*. 2011). The knockout mir-143 mice do not develop obesity-associated insulin resistance. MiR-122 was up-regulated in low RFI animals. Mir-122 was the first miRNA to be linked to metabolic control and it is the most expressed miRNA in human and mouse liver and affects hepatic cholesterol and lipid metabolism. Suppression of mir-122 by antisense reduced plasma cholesterol levels by 25–30% in mice. It also reduced the genes involved in lipid synthesis in liver and decreased hepatic cholesterol and fatty acids (Rottiers & Näär, 2012).

In conclusion, we have identified 305 known bovine miRNA in bovine liver. Mir-143 is the most expressed bovine miRNA in liver, and is up-regulated in high RFI cattle. Mir-122 is the second most expressed miRNA in liver and is down regulated in high RFI animals. The differentially expressed miRNAs may play important roles in the regulation of the bioprocesses responsible for the variation in RFI in cattle.

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.

Chen Y., Gondro C., Quinn K., Herd R.M., Parnell P.F. and Vanselow B. (2011) *Anim. Genet.* **42**: 475.

Hackenberg M., Sturm M., Langenberger D., Falcon-Perez J.M. and Aransay A.M. (2009) *Nucl. Aci. Rese.* **37**: W68.

Hackenberg M., Rodrı´guez-Ezpeleta N. and Aransay A.M. (2011) *Nucleic Acids Research* **39**: W132.

Jordan S.D., Kruger M., Willmes D.M., Redemann N., Wunderlich F.T., Bronneke, H.S., Merkwirth C., Kashkar H., Olkkonen V.M., BottgerT., Braun T., Seibler J. and Bruning J.C. (2011) *Nat. Cell. Biol.* **13**:434.

Koch R.M., Swiger L.A., Chambers D. and Gregory K.E. (1963) *J. Anim. Sci.* **22**: 486.

Lu j., Miska E. A., Alvarez-Saavedra E., & et al. (2005). *Nature*, **435**: 834.

Rottiers V. and Näär, A.M. (2012) *Nat. Rev. Mol. Cell Bio.l* **13**: 239.

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:372-375

GENOME-WIDE EPISTASIS ASSOCIATION OF ULTRASOUND-SCANNED CARCASS TRAITS IN BEEF CATTLE: TWO-STAGE MODELS

A. A. Ali[1,2], M. S. Khatkar[1] , H. N. Kadarmideen[3] and P. C. Thomson[1]

[1] ReproGen - Animal Bioscience Group, Faculty of Veterinary Science, University of Sydney, Camden NSW 2570, Australia, [2] CSIRO Animal, Food and Health Sciences, PMB PO Aitkenvale, Townsville, QLD 4814, Australia, [3] Faculty of Health and Medical Sciences, University of Copenhagen, Grønnegardsvej 7, 1870 Frederiksberg C, Denmark

SUMMARY

Most genome-wide association studies (GWAS) on complex traits in livestock have focused on identifying single-locus effects ignoring any epistatic interaction. Here we consider analytical methods that explicitly look for statistical interactions between two loci. Two-stage models were used to ease multiple testing problems and computational demand using beef cattle data as an example. The results suggest that fitting epistasis models for GWAS using two-stage models is a useful strategy for detecting significant interactions between genetic loci and may help in searching for candidate genes and polymorphisms influencing phenotypic variation.

INTRODUCTION

Genome-wide association studies (GWAS) are now routinely used to identify genomic regions associated with traits of interest; however, this ignores an important class of genomic associations, that of epistatic interactions (Carlborg and Haley 2004; Hemani *et al*. 2013). Identifying genome-wide interactions among SNPs (single nucleotide polymorphisms) using high-density SNP chip genotypes is a difficult task due to statistical complexity (e.g. multiple testing) and computational burden (Marchini *et al*. 2005; De Lobel *et al*. 2010). For example, consider the current Bovine SNP Chip which comprises of more than 50,000 SNPs, the number of SNP combinations would be 1.25×10^9 for testing two SNPs at a time; this analysis could take several days or even weeks on a standard workstation. The number of possible interactions involving more than two loci will be exponentially higher. For these reasons, epistasis is not yet a standard tool in complex trait studies. Rather than testing all possible pair-wise comparisons, a more practical strategy might be to examine a subset of SNPs which could have influence on a trait of interest. Here we show a two-stage approach for analysing genome-wide epistasis association (GWEA) using real ultrasound scan measures for carcass traits on beef cattle.

MATERILS AND METHODS

Animals and data: Animals used were part of a northern Australian breeding project of the Co-operative Research Centre for Beef Genetic Technologies. A total of 583 heifers of Brahman breed were ultrasound-scanned for eye muscle area (EMA, cm^2), rump fat depth at P8 site (P8, mm) and rib fat depth measured the between 12^{th} and 13^{th} ribs (RIB, mm). The fixed environmental effects recorded were age (in days), month of calving, herd, and cohort (combination of experimental location and heifer's year of birth). The details of the resource population and the data were described previously (Barwick *et al*. 2009; Bolormaa *et al*. 2011). Animals were genotyped using 10,000 ParaAllele/Affymetrix SNP chips (Khatkar *et al.* 2007). After quality control, 565 animals representing 51 sire families and 6,715 SNPs on bovine autosomal chromosomes (BTA) were analysed. The SNP positions were mapped to the genome

assembly UMD 3 (http://www.cbcb.umd.edu/research/bos_taurus_assembly.shtml). Missing SNP genotypes were imputed using Beagle (Browning and Browning 2007). Genotypes were coded as 0 for the homozygote for allele (A), 1 for the heterozygote (AB), and 2 for the homozygote for allele (B).

Additive association model: SNP effects were estimated by a single-trait-single-SNP association analysis. The additive effect of a SNP on each trait was estimated from its associated regression coefficient. The traits and SNP were fitted using the following linear mixed:

$$\text{Trait} = \mu + \text{fixed\_effect} + \beta N_{SNP} + \text{Animal} + \varepsilon$$

where Trait is the phenotypic measurement for EMA, P8 or RIB trait, μ is the overall mean, fixed_effects were significant fixed effects specific to each trait (described in Results), N_{SNP} is the number of copies of the allele (0, 1, and 2) at the SNP as a covariate, Animal is the polygenic effect of animals to account for the effect of relatedness, and ε is the random error. Random animal and residual effects were assumed to be normally distributed with zero mean, and additive genetic variance σ_A^2 and residual variance σ_ε^2.

Epistasis association model: To ease computational burden, a subset of significant SNPs at a lenient threshold of *p*-value ≤ 0.01 was selected from the additive association model, At this threshold, 82 SNPs for EMA, 86 SNPs for P8, and 92 SNPs for RIB were available and analysed for two loci (e.g. $SNP_1 \times SNP_2$) epistasis using the following linear mixed model:

$$\text{Trait} = \mu + \text{fixed\_effect} + SNP_1 + SNP_2 + SNP_1 \times SNP_2 + \text{Animal} + \varepsilon$$

where SNP_1 and SNP_2 are three-level factors for genotypes (e.g. AA, AB, and BB) at SNP_1 and SNP_2, where $SNP_1 \times SNP_2$ in the interaction between SNP_1 and SNP_2 genotypes as an indicator for epistasis effect. Note that while the SNP effect for the additive association model was treated as a covariate (to maximise power of association detection), for estimating interactions it is necessary to treat the effects of SNPs as factors. A separate model was fitted for each pair of SNPs for all pair-wise combination of selected SNPs. Other terms in the model were the same as the additive association model. To account for multiple testing, false discovery rates was estimated for GWEA results using the *q*-value package in R 2.15 version (Storey and Tibshirani 2003). All analyses were performed using a REML procedure in ASReml-R package (Butler *et al*. 2009).

RESULTS AND DISCUSSIONS

Summary statistics and estimates of heritability for ultrasound scan measures of eye muscle area and fat depths traits are presented in Table 1. The effects of age, month of calving, herd and cohort were significantly associated with phenotypes considered.

Table 1. Summary of ultrasound scan measures for eye muscle area (EMA, cm^2), rump fat depth at P8 site (P8, mm) and fat depth between 12^{th} and 13^{th} ribs (RIB, mm) in Brahman cattle

Trait	No. animals	Mean	SD	Min	Max	Heritability
EMA	564	43.8	6.6	28.0	64.0	0.42 ± 0.15
P8	565	3.7	3.7	1.0	12.0	0.48 ± 0.15
RIB	565	1.9	1.0	1.0	6.0	0.65 ± 0.17

SNP association: Results of the GWAS analyses for ultrasound scan carcass related traits are presented in Figure 1. Using a *q*-value threshold of 0.05 which corresponds to *p*-value $\leq 1 \times 10^{-6}$, a total of six SNPs on BTA 5 and 14 were identified as being significantly associated with the ultrasound scan measures of fat depths traits, most of these SNP were located on BTA 14 at 22 to 24 Mb. The most significant SNP ($p = 4.9 \times 10^{-11}$; $q = 3.3 \times 10^{-7}$; MAF = 0.498) was rs29020688 on BTA 14 at 24 Mb for fat depth at P8. This SNP was located in the intronic region of *Bos taurus* XK, Kell blood group complex subunit-related family, member 4 (XKR4) gene. Recently, the XKR4 was identified as a candidate gene affecting rump fat depth in Australian tropical cattle (Bolormaa *et al*. 2011). Other significant SNPs for fat depth traits were rs29010515, rs29010516 on BTA 14 and rs29010471 and rs29026420 on BTA 5.

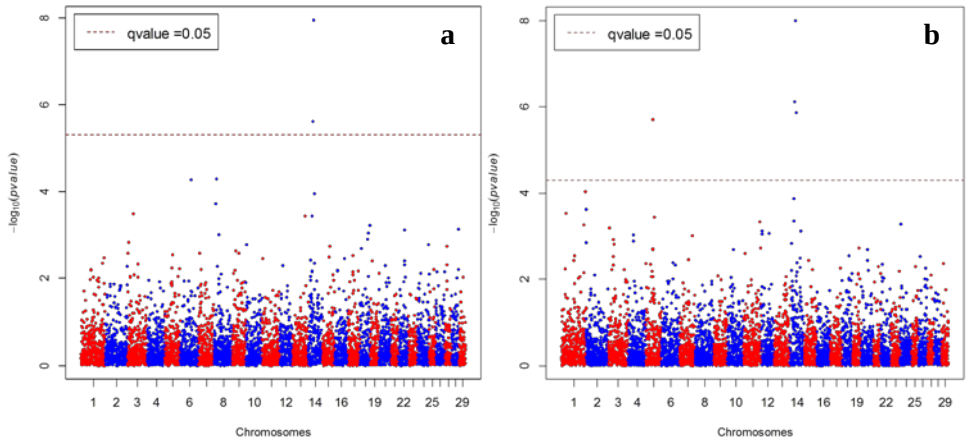


Figure 1. Manhattan plots of GWAS scan, a significant threshold (dashed red line) is drawn corresponding to an FDR of 0.05. a: rump fat depth at P8 site. Note a significant association on chromsome 14. b: rib fat depth. Note the significant associations on chromosome 5 and 14.

Epistasis association: Image plots in Figure 2 depict graphical representation of GWEA results for ultrasound scanned fat depths at P8 site and RIB traits. These heat maps show epistatic signals or so called 'hot spot' (red colour spots) where significant epistatic association were detected at genome-wide level. The colour gradient in the panel on the right side of the plot represents $-\log_{10}$ (*p*-values). Evidence of epistatic association was detected on several chromosomes; however, the strongest epistatic signals were for pairs of SNPs on BTA 8 and 12, BTA 8 and 14, and BTA 8 and 15 for fat depth at P8 site trait. Using a *q*-value of 0.10 which corresponds to a *p*-value of 10^{-4}, significant epistatic interactions between SNPs in DDX56 (BTA 4) and EFHD2 (BTA 16), MAP3K5 (BTA 9) and TMEM132D (BTA 12) genes were detected to be associated with scanned RIB fat depth, whereas an interaction between SNPs in ROBO2 (BTA 1) and DZANK1 (BTA 13) genes was associated with P8 site fat depth. These epistatic genes could be detected only after fitting epistasis models. Once significant gene-gene epistatic interactions are identified, this will facilitate to define networks of interacting genes that can be incorporated into existing functional annotation and molecular pathways, and hence provide the genomic basis of improving carcass traits efficiently in beef cattle.

CONCLUSIONS

This study demonstrated the usefulness of modelling epistasis in the analysis of complex traits, and identified potential candidate genes affecting carcass traits using field data from an Australian Brahman cattle population. Information about epistasis can add to our understanding of the complex genomic networks that form the fundamental basis of biological systems.

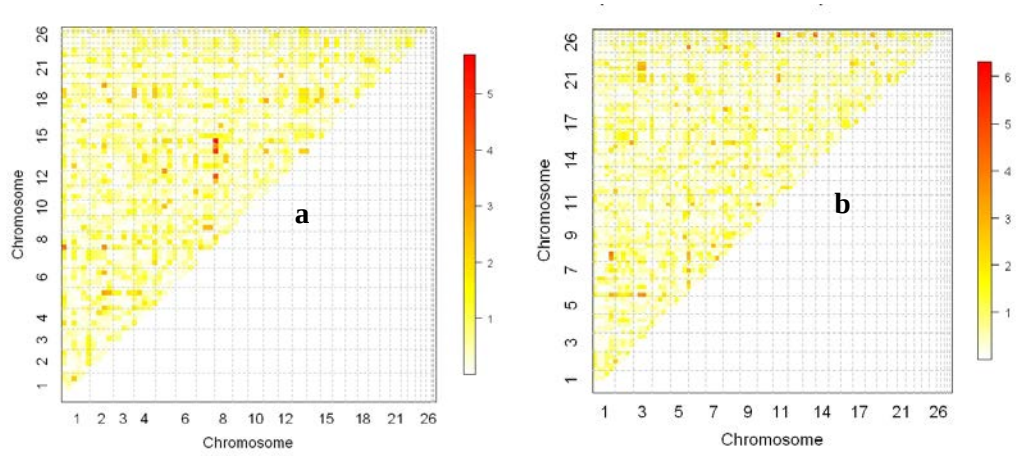


Figure 2. Heatmap image of genome-wide epistasis association. The heatmap legend scale (right side) is on $-\log_{10}$(*p*-value) scale. a: fat depth at P8 site. b: fat depth at RIB. Note red spots indicate epistatic signals.

ACKNOWLEDGEMTNS

We acknowledge the CRC for Beef Genetic Technologies for providing the data. A. A. Ali was the recipient of a CSIRO Animal, Food and Health Sciences PhD scholarship.

REFERENCES

Barwick S.A., Johnston D.J., Burrow H.M., Holroyd R.G., Fordyce G., Wolcott M.L., Sim W.D. and Sullivan M.T. (2009) *Anim. Produc. Sci.* **49**: 367.

Bolormaa S., Neto L.R.P., Zhang Y.D., Bunch R.J., Harrison B.E., Goddard M.E. and Barendse W. (2011) *J. Anim. Sci.* **89**: 2297.

Browning S.R. and Browning B.L. (2007) *Am. J. Hum. Genet.* **81**: 1084.

Butler D.G., Cullis B.R., Gilmour A.R. and Gogel B.J. (2009) Tech. Report QLD DPI.

Carlborg O. and Haley C.S. (2004) *Nat. Rev. Genet.* **5**: 618.

De Lobel L., Geurts P., Baele G., Castro-Giner F., Kogevinas M. and Van Steen K. (2010) *Eur. J. Hum. Genet.* **18**: 1127 .

Hemani G, Knott S, Haley C. (2013) *Plos Genet.* **9**: e1003295.

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., Raadsma H.W. (2007). *Genet.* **176**:763

Marchini J., Donnelly P. and Cardon L.R. (2005) *Nat. Genet.* **37**: 413.

Storey J.D. and Tibshirani R. (2003) *Proc. Nat. Acad. Sci.* **100**: 9440 .

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:376

RUMEN DIFFERENCES BETWEEN SHEEP IDENTIFIED AS BEING LOW OR HIGH EMITTERS OF GREENHOUSE GAS

W.E. Bain[1], L Bezuidenhout[1], N.B. Jopson[2], C.S. Pinares-Patino[3] and J.C. McEwan[1]

[1]AgResearch Invermay, Private Bag 50034, Mosgiel, 9053 New Zealand
[2]AbacusBio, PO Box 5585, Dunedin, 9058, New Zealand
[3]AgResearch Grasslands, Tennent Drive, Private Bag 11008, Palmerston North, New Zealand

ABSTRACT

Methane is known to be one of the main greenhouse gases contributing to climate change. A primary contributor to methane emissions are ruminants. It is estimated that over 50% of greenhouse gases in New Zealand are produced by the agricultural sector. Research using chambers has given the ability to collect accurate measurements on emissions, and animals have been identified as high and low emitters. Forty five ewes born in 2009, consisting of the top and bottom 10% of methane emitters (gCH4/kgDMI) were computer tomography (CT) scanned in June 2012 (24 high and 21 low emitters). Animals were CT scanned using Cavalieri's theorem, with images collected at 15 mm intervals and a total of 30-32 images collected from each animal. The following rumen compartments; reticulum, rumen and atrium, ventral sac of rumen (rostral), dorsal sac of rumen and ventral sac of rumen were measured for volume, surface area, raft, liquid and gas volume, and weight. Data were analysed using R, and parsimonious models selected for the different response variables. Measurements between the two groups differed in range from 12% to 28% with high methane emitters having larger total surface area (12.35%), total volume (20.24%), and total raft volume and weight (26.8, 27.8 respectively). The results indicate that rumens from high methane emitters have larger volumes compared to low methane emitting sheep. In conclusion it would appear that selection for reduced methane emissions has resulted in a correlated change in rumen size in ewes.

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:377-380

GENETIC TRENDS IN A MERINO LINE SELECTED FOR A REDUCED FIBRE DIAMETER RELATIVE TO AN UNSELECTED CONTROL FLOCK

S.W.P. Cloete[1,2], J.J. Olivier[2] and E. du Toit[3]
[1]Department of Animal Sciences, University of Stellenbosch, Matieland 7602, South Africa
[2]Directorate Animal Sciences: Elsenburg, Elsenburg 7607, South Africa
[3]Directorate Animal Sciences: Tygerhoek, Riviersonderend 7250, South Africa

SUMMARY

A Merino line selected for a reduced fibre diameter (Fine wool line) was compared with a random selection control flock. Foundation sires in the Fine wool line were initially obtained from the parent stud and industry, followed by within flock selection. Selection was for a reduced fibre diameter, while maintaining live weight. Data of ~2700 records of hogget live weight (LW), clean fleece weight (CFW), staple length (SL), staple strength (SS), and fibre diameter (FD) recorded from 1998 to 2009 were used to derive genetic parameters for all traits in a five-trait animal model. Genetic parameters were consistent with literature values. Estimated breeding values (EBVs) in each year provided genetic trends in the two flocks. In the Fine wool line, EBVs for FD were reduced by 1.01% per annum (-0.19±0.02 μm) relative to the phenotypic mean for FD. There was also some evidence of favourable genetic change in LW and CFW in the Control flock. Fine wool line progeny maintained their LW, but showed declines in CFW, SL, and SS. The improvement in FD in the Fine wool line should be balanced against the deterioration of CFW and SS.

INTRODUCTION

Fibre diameter (FD) is commonly reported to be the most important determinant of the price of Merino wool (Cottle 2010). However, there are unfavourable genetic correlations of FD with other traits of economic importance (such as LW, CFW and SS) (Safari *et al*., 2005; 2007c; Huisman and Brown 2009). In view of the importance of FD, the South African sheep industry undertook the establishment of a genetic fine wool stud in the late 1980's (Schoeman *et al*. 2010). Cloete *et al*. (2007) reported that initial selection emphasis on LW and CFW in this stud resulted in an initially nonsignificant genetic trend for FD. After amending the selection strategy in 1995, the genetic response in LW was reduced, CFW remained stable, while FD declined by 0.67% per annum on the genetic level.

Emphasis on traits other than FD obviously compromised the genetic gain that could be obtained in FD in this stud. The interest in the response of FD in the absence of emphasis on other traits resulted in the establishment of a fine wool line at the Tygerhoek research farm where the focus was primarily on reducing FD. This study reports the genetic change in this line, in comparison with an unselected Control flock already present on the farm (Cloete et al. 1998).

MATERIALS AND METHODS

The experimental animals that were used were introduced from the Halesowen stud, as described by Olivier *et al*. (1999) and Schoeman *et al*. (2010). This stud was initially established by sourcing the finest maiden replacement ewes from industry flocks with a below average clip FD in the national clip, on the provision that they were above average for LW in their respective contemporary groups. The ewes were purchased from their original owners and mated to 4 Australian fine-wool rams (obtained from the Glenleigh stud in NSW and the Siera Park stud in Victoria), and were subsequently bred to rams from within the flock. During 1997, surplus ewes from this stud were transferred from the Halesowen research farm near Cradock to Tygerhoek research farm near Riviersonderend, to establish a fine-wool gene pool for further selection for a reduced fibre diameter (the Fine wool line) from 1998 up to and including 2009. This line was

maintained along with an unselected control flock described by Cloete *et al*. (1998) at Tygerhoek. The latter authors also described the experimental site. Rams in the Fine wool line were initially sourced from the parent stud (n=6) and from industry (n=4) to sire the bulk of progeny from 1998 to 2000. Such rams were treated as part of the base population in the analysis. Subsequently sires were selected from within the flock, while three migrant rams were also introduced. Two of these rams originated from the Grange stud in WA and sired progeny in 2002, while another ram from the parent stud sired progeny from 2007 to 2009. The low number of subsequent migrant rams did not validate special treatment as separate genetic groups. Selection was for a reduced FD, while it was attempted to maintain LW by ensuring that the mean estimated breeding value (EBV) for LW of rams selected for breeding exceeded the mean of all replacements.

Midrib wool samples were obtained from all hogget progeny after a growth period of approximately 10 months and analysed for clean scoured yield percentage (CY), SL, SS and FD. Greasy fleece weights were recorded at shearing about two months later. After being shorn, hogget LW was recorded. Greasy fleece weight and CY were used to calculate CFW. All the recorded traits were linked to pedigree information.

Each trait was initially subjected to single-trait genetic analyses to obtain prior values for a subsequent multi-trait analysis. The single random effect of animal was fitted, using ASREML (Gilmour *et al*. 2006). The data were then subjected to a five-trait animal model analysis to derive genetic (co)variance components to estimate the heritability (h^2) of all traits, as well as genetic and phenotypic correlations (r_g and r_p respectively). Animal solutions obtained in this way were used to construct genetic trends for the respective traits in the Control flock and the Fine wool line.

RESULTS AND DISCUSSION

The coefficients of variation (CV's) for the respective traits ranged between 10.4% for FD and 36.4% for SS (Table 1). Literature values suggested CV's of 14-28 % for LW, 18-42% for CFW and 9-11% for FD (Olivier and Cloete 2007; Safari *et al*. 2007a; Huisman and Brown 2009). The present CV's are well within this range of values.

Table 1. **Descriptive statistics for the traits included in the five-trait analysis, including means, standard deviations (s.d.) and coefficients of variation (CV)**

Trait	Number	Mean±s.d.	CV (%)
Live weight (kg)	2510	53.9±11.2	20.8
Clean fleece weight (kg)	2556	3.24±0.79	24.4
Staple length (mm)	2399	85.3±13.2	15.5
Staple strength (N/ktex)	2102	34.3±12.5	36.4
Fibre diameter (μm)	2622	18.3±1.9	10.4

Estimates of h^2 amounted to 0.50 for LW, 0.41 for CFW, 0.39 for SL, 0.19 for SS and 0.76 for FD (Table 2). Corresponding literature values ranged from 0.33-0.52 for LW, from 0.28-0.42 for CFW and from 0.55-0.74 for FD (Swan et al. 1995; Rose and Pepper 1999; Cloete *et al*. 2002; Safari *et al*. 2005; 2007b; Olivier and Cloete 2007). Estimates of h^2 for SL ranged from 0.26-0.46 (Swan *et al*. 1995; Safari *et al*. 2005; Olivier and Cloete 2007), while the h^2 of SS was estimated at 0.13-0.34 (Swan et al. 1995; Wuliji *et al*. 2001; Safari et al. 2005). The present results were all within these ranges, but the h^2 of FD (0.76) was slightly higher than the upper boundary. A similarly high h^2 estimate of 0.74 was reported for FD by Rose and Pepper (1999), suggesting that such high h^2 estimates for FD are indeed feasible. Genetic correlations were also consistent with those in the literature (Swan *et al*. 1995; Safari *et al*. 2005; 2007c; Olivier and Cloete 2007; Huisman and Brown 2009). The unfavourable r_g between FD and SS of 0.44 concurs with

literature values of 0.52 (Wuliji *et al*. 2001) and 0.37 (Safari *et al*. 2005).

Table 2. Additive and residual variance components (respectively σ^2_A and σ^2_E) and (co)variance ratios (±s.e.) for hogget live weight (LW), clean fleece weight (CFW), staple length (SL), staple strength (SS) and fibre diameter (FD)

Trait	LW	CFW	SL	SS	FD
Variance components					
σ^2_A	26.86	0.165	32.67	27.39	1.357
σ^2_E	26.86	0.236	50.48	113.28	0.422
(Co)variance ratios (h^2 in bold on the diagonal, r_g above the diagonal and r_p below the diagonal					
LW	**0.50±0.04**	0.33±0.07	0.14±0.08	0.30±0.10	-0.10±0.06
CFW	0.36±0.02	**0.41±0.04**	0.54±0.07	0.09±0.12	0.06±0.06
SL	0.17±0.02	0.49±0.02	**0.39±0.04**	0.34±0.11	0.05±0.07
SS	0.13±0.02	0.16±0.02	0.15±0.02	**0.19±0.04**	0.44±0.07
FD	0.06±0.02	0.16±0.02	0.14±0.02	0.28±0.02	**0.76±0.03**

Phenotypic means for the1998 progeny indicated that progeny from the Fine wool line was initially heavier (57.0±0.9 vs. 45.3±0.6 kg), with heavier fleeces (4.41±0.16 vs. 3.61±0.11 kg) and longer staples (94.7±1.9 vs. 85.9±1.3 mm) than the Control flock (all $P<0.01$). FD was nearly 1 μm lower in the Fine Wool line than in the Control flock (19.3±0.2 vs. 20.2±0.1 μm; $P<0.01$). These initial differences between flocks were expected, as the Fine wool line originated from the fine wool stud at Halesowen, while the Control flock was resident at Tygerhoek. Subsequent genetic change in the Control flock was below 0.3% of the overall phenotypic mean for SL, SS and FD. However, Control flock progeny became appreciably heavier with heavier fleeces with time (both $P<0.01$). A lack of genetic stability for LW has previously been reported in the Control flock, while a corresponding trend was reported for CFW (Cloete et al. 1998). EBVs for FD in the Fine wool line declined at 1.01% of the overall mean per annum, while a slight increase was noted for LW ($P<0.05$). Correlated responses in the Fine wool line were unfavourable in the other traits, leading to appreciable reductions in CFW and SS in particular ($P<0.01$; Table 3).

Table 3. Regressions of average EBVs on year (±s.e.), depicting genetic trends in the Control flock and Fine wool line at Tygerhoek

Regression Parameter	Traits*				
	LW	CFW	SL	SS	FD
Control flock					
Intercept	-6.21±0.32[a]	-0.22±0.05[a]	-2.00±0.04[a]	-0.46±0.42[a]	1.35±0.16[a]
Regression coefficient	0.41±0.05[a]	0.04±0.01[a]	0.13±0.04[a]	-0.02±0.06[a]	-0.04±0.02[a]
Fine wool line					
Intercept	3.59±0.42[b]	0.18±0.04[b]	1.94±0.06[b]	0.85±0.55[b]	-0.40±0.12[b]
Regression coefficient	0.16±0.07[b]	-0.03±0.01[b]	-0.19±0.09[b]	-0.30±0.09[b]	-0.19±0.02[b]

*Live weight (LW), clean fleece weight (CFW), staple length (SL), staple strength (SS), fibre diameter (FD); [a,b] – Denote differences between lines for regression parameters at $P<0.05$

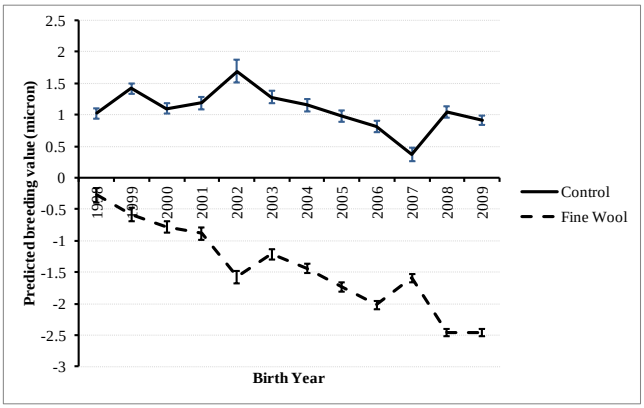


Figure 1. **Genetic trends for fibre diameter in the Control and Fine Wool lines at Tygerhoek. Vertical lines about means reflect standard errors.**

Averaged annual EBVs for FD in the control flock ranged from 0.5 to 1.5 µm (Figure 1). Average EBVs in the Fine wool line were reduced from -0.28 µm in 1998 to -2.46 µm in 2009.

CONCLUSIONS

Genetic trends indicated that FD was substantially reduced in the Fine wool line, while LW remained stable. However, the response in FD was associated with unfavourable correlated responses, particularly in CFW and SS. The correlated response in SS of the Fine wool line was consistent with an unfavourable within-flock genetic correlation of 0.44 between FD and SS. The premium paid for fine wool will determine whether reductions in FD will be economically viable.

REFERENCES

Cloete S.W.P., Greeff J.C., Lewer R.P. (2002) *Aust. J. Agric. Res.* **53**, 271.
Cloete S.W.P., Olivier J.J., Olivier W.J. (2007) *Proc. Assoc. Advmnt. Anim. Beed. Gen.* **17**: 320.
Cloete S.W.P., Olivier J.J., Snyman M.A., Du Toit, E. (1998) *Aust. J. Exp. Agric.* **38**: 427.
Cottle D.J. (2010) Wool Preparation, Testing and Marketing. In: Cottle D.J. (ed.) International Sheep and Wool Handbook, pp. 581-618. Nottingham University Press, Nottingham,UK.
Gilmour A.R., Gogel B.J., Cullis B.R., Thompson R. (2006) ASReml User Guide. Release 2.0. (VSN International Ltd, Hemel Hempstead, HP1 1ES: UK).
Huisman A.E., Brown D.J. (2009) *Anim. Prod. Sci.* **49**: 289.
Olivier J.J., Bezuidenhout A.G., Greyling A.C., Cloete S.W.P. (1999) *Proc. Assoc. Advmnt. Anim. Beed. Gen.* **13**: 62.
Olivier J.J., Cloete S.W.P. (2007) *Proc. Assoc. Advmnt. Anim. Beed. Gen.* **17**: 324.
Rose M., Pepper P. (1999) *Proc. Assoc. Advmnt. Anim. Beed. Gen.* **13**: 114.
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., 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., 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., van der Werf J.H.J. (2007c) *J. Anim. Breed. Gen.* **124**: 65.
Schoeman S.J., Cloete S.W.P., Olivier, J.J. (2010) *Livest. Sci.* **130**: 70-82.
Swan A.A., Lax J., Purvis I.W. (1995) *Proc. Aust. Assoc. Anim. Beed. Gen.* **11**: 516.
Wuliji T., Dodds K.G., Land J.T.J., Andrews R.N., Turner P.R. (2001) *Anim. Sci.* 72: 241.

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:381-384

GENETIC PARAMETERS FOR SLAUGHTER AND MEAT TRAITS IN OSTRICHES

A. Engelbrecht[1,2], S.W.P. Cloete[1,2], K.L. Bunter[3], J.B. van Wyk[4] and L.C. Hoffman[2]

[1]Directorate Animal Sciences, Western Cape Government Agriculture, South Africa
[2]Department of Animal Sciences, Stellenbosch University, Matieland 7602, South Africa
[3]Animal Genetics and Breeding Unit, University of New England, Armidale, NSW 2351, Australia
[4]Department of Animal, Wildlife and Grassland Sciences, University of the Free State, Bloemfontein 9300, South Africa

SUMMARY

Genetic parameters for ostrich slaughter and meat traits were estimated to determine whether the improvement of slaughter yield through genetic selection will be possible. Live weight before slaughter, *post mortem* weight, carcass weight, pelvic limb weight, muscle weights and fat depot weights were recorded. Abdominal and subcutaneous fat weights were highly variable, while coefficients of variation in the other traits ranged between 16 and 29%. All traits showed significant genetic variation, with estimates of heritability ranging from 0.21 to 0.34 for weight and carcass traits. Heritability estimates for individual muscle weights ranged from 0.14 to 0.43, while the genetic correlations among the individual muscle weights and with pre-slaughter live weight were all positive. The substantial variation, high and favourable genetic correlations between traits, and moderate to high heritability estimates indicate that genetic improvement in ostrich carcass traits is achievable.

INTRODUCTION

Genetic improvement of carcass traits in farmed livestock species has become an important tool at the disposal of producers specialising in meat production. Improving the yield of ostrich carcass components of economic importance also needs to be investigated in light of the importance of meat production for the ostrich industry. To establish a breeding program to improve carcass composition and yield in ostriches, it is essential to gain knowledge of genetic parameters involving carcass traits. Since this is currently lacking, this study aimed to estimate genetic parameters for quantitative ostrich slaughter, carcass and meat traits. The relationships of these traits with body weight were also investigated.

MATERIALS AND METHODS

Slaughter data were collected from the progeny of the commercial ostrich breeding flock maintained at the Oudtshoorn Research Farm, South Africa, slaughtered from 1997 to 2011. Only data from South African Black ostriches, slaughtered between 210 and 540 days of age, were used. Contemporary groups for on-farm weight were defined as year by season to represent chicks that were reared in the same environment and management regime. Weight data from contemporaries not slaughtered, some of which later on became parents, could therefore also be included in the analysis. Slaughter date was used to derive contemporary groups (slaughter groups) for slaughter traits, thereby representing ostriches slaughtered under the same slaughter conditions. Slaughter groups were large similarly aged birds that were reared together for the months immediately preceding slaughter. The final dataset analysed represented the progeny of 305 sires and 298 dams, pair-mated to each other in 382 unique combinations. Ostriches were weighed on-farm, before being transported for slaughter and dressing at a local abattoir. The ostriches were fasted for 24 h, electrically stunned, hoisted by the legs and bled before the feathers, skins, internal organs and abdominal fat were removed. Slaughter data routinely recorded at the abattoir included *post mortem* weight (after bleed out) and warm carcass weight. The weights of the subcutaneous

fat and abdominal fat depots were also determined for some of the slaughter groups. A number of right pelvic limbs from various slaughter groups were also removed from the abattoir for further dissection and investigation. The weights of the pelvic limbs (leg and thigh) were determined after being chilled at 0°C for 24 h. Ten of the major muscles that are sold as commercial cuts were also dissected and weighed (Table 1).

Statistical analysis. Age was fitted as a linear covariate for all traits. Fixed effects fitted for all traits initially included contemporary group for live weight, slaughter group for slaughter traits and gender (male or female). Two-way interactions between these effects were also included in the initial models. Effects found to be significant were fitted in the final models for each trait.

Variance components were estimated by REML procedures fitting an animal model (Gilmour *et al*. 2009). Two random models were fitted. Model 1 only included the direct genetic effects, while maternal permanent environmental effects were added in Model 2. Live weight was analyzed in multi-trait analyses with carcass traits and groups of muscle weight traits, respectively. The three lower leg muscles (*Muscularis gastrocnemius pars interna, M. fibularis longus* and *M. gastrocnemius pars externa*) were analysed together, as was the four post-acetabular muscles (*M. iliofemoralis, M. flexor cruris lateralis, M. iliofibularis* and *M. iliotibialis lateralis*). The rest of the muscles (*M. femorotibialis medius, M. iliofemoralis externus* and *M. iliotibialis cranialis*), which could not be conclusively grouped according to location, was analysed together. All analyses included the full pedigree file, consisting of 6 541 individuals, the progeny of 378 sires and 376 dams, mated to each other in 541 unique combinations. Individuals hatched from generation one through seven were included.

RESULTS AND DISCUSSION

The average slaughter age was 373 ± 76 days. Abdominal and subcutaneous fat weights were highly variable as reflected by high coefficients of variation of >50%, while coefficients of variation in the other traits ranged between 16 and 29% (Table 1).

Table 1 Means (± s.d.), coefficients of variation (CV) and ranges for ostrich slaughter traits

Trait	N	Mean (± s.d.)	CV (%)	Range
Slaughter traits:				
Live weight (kg)	1 897	90.3 ± 19.3	21	42 – 146
Post mortem weight (kg)	1 052	84.2 ± 13.1	16	43.5 - 134.7
Carcass traits:				
Carcass weight (kg)	1 268	41.8 ± 7.3	17	20.1 - 62.7
Pelvic limb weight (kg)	976	14.8 ± 3.0	20	6.5 - 23.4
Abdominal fat weight (kg)	424	3.708 ± 2.035	55	0.238 - 10.220
Subcutaneous fat weight (kg)	737	2.687 ± 1.593	59	0.293 - 10.166
Muscle traits:				
M. gastrocnemius pars interna (kg)	738	0.793 ± 0.191	24	0.284 - 1.446
M. fibularis longus (kg)	738	0.263 ± 0.067	25	0.102 - 0.487
M. gastrocnemius pars externa (kg)	738	0.527 ± 0.150	28	0.192 - 1.157
M. iliotibialis lateralis (kg)	740	0.988 ± 0.230	23	0.268 - 1.560
M. iliofibularis (kg)	871	1.394 ± 0.339	24	0.368 - 2.400
M. iliofemoralis (kg)	775	0.357 ± 0.102	29	0.105 - 0.637
M. flexor cruris lateralis (kg)	740	0.285 ± 0.083	29	0.103 - 0.531
M. iliotibialis cranialis (kg)	739	0.445 ± 0.109	24	0.171 - 0.852
M. iliofemoralis externus (kg)	740	0.168 ± 0.039	23	0.077 - 0.289
M. femorotibialis medius (kg)	739	0.659 ± 0.136	21	0.312 - 1.171

N: number of records; s.d.: standard deviation; CV: coefficient of variation

Most carcass and meat traits were dependent on age (with the exception of *post mortem* weight and *Muscularis iliofemoralis* weight), with yield generally increasing with an increased slaughter age. Gender seemed to influence fat weights, with males having less fat, both in the abdominal and subcutaneous depots. Broiler males were also shown to be leaner than females (Zerehdaran *et al*. 2004). Nonetheless, ostrich females produced the same pelvic limb weight as males, as reported by Hoffman *et al*. (2009). It seems, however, that the contribution of different muscles to the thigh and leg may vary between males and females. Some muscles (*M. iliofibularis* and *M. femorotibialis medius*) were heavier in females, while some were heavier in males (*M. iliofemoralis externus*). Carcass composition thus seemed to differ slightly between males and females. This would be consistent with studies on chickens, which showed that female chickens had a higher breast yield, but lower thigh and drumstick yields than males (Baeza *et al*. 2010).

(Co)variance components, ratios and correlations. The inclusion of the direct genetic component as a random effect in the operational model resulted in an improved log-likelihood for all traits, with the exception of abdominal fat weight. The additional inclusion of the maternal permanent environmental effect in the operational model resulted in an improved log-likelihood for live weight, pelvic limb weight, subcutaneous fat weight and some of the muscle weights.

The direct genetic component was thus fitted as default for all traits, resulting in heritability estimates (h^2) of 0.22 ± 0.05 for live weight, 0.44 ± 0.08 for *post mortem* weight, 0.29 ± 0.06 for carcass weight, 0.18 ± 0.09 for pelvic limb weight, 0.09 ± 0.09 for abdominal fat weight and 0.16 ± 0.12 for subcutaneous fat weight in single-trait analyses. Significant heritability estimates were also obtained for most of the muscle weights. Maternal permanent environmental variance ratios (pe^2) accounted for between 5% (live weight) and 16% (*M. gastrocnemius pars interna* weight) of the phenotypic variation for the respective traits. However, the maternal permanent environmental effect for all traits became insignificant in multi-trait analysis and was thus not included in the final analysis involving any combination of traits. Results from a five-trait model including live weight and various carcass traits are given in Table 2.

Table 2 (Co)variance components and ratios (± s.e.), along with residual and phenotypic variances and correlations between ostrich live weight and carcass traits from multi-trait analyses

Trait	Live weight	Carcass weight	Pelvic limb weight	Abdominal fat weight	Subcutaneous fat weight
Additive genetic correlations (h^2 in bold)					
Live weight	**0.34 ± 0.06**	0.94 ± 0.03	0.90 ± 0.04	0.56 ± 0.16	0.92 ± 0.07
Carcass weight		**0.27 ± 0.06**	0.99 ± 0.01	0.47 ± 0.18	0.73 ± 0.12
Pelvic limb weight			**0.32 ± 0.06**	0.41 ± 0.19	0.67 ± 0.13
Abdominal fat weight				**0.22 ± 0.08**	0.63 ± 0.18
Subcutaneous fat weight					**0.21 ± 0.06**
Residual correlations (σ^2_e in bold)					
Live weight	**81**	0.71 ± 0.02	0.75 ± 0.02	0.62 ± 0.05	0.52 ± 0.04
Carcass weight		**21.6**	0.93 ± 0.01	0.62 ± 0.05	0.38 ± 0.05
Pelvic limb weight			**2.68**	0.54 ± 0.06	0.34 ± 0.05
Abdominal fat weight				**1.5**	0.48 ± 0.05
Subcutaneous fat weight					**1.3**
Phenotypic correlations (σ^2_p in bold)					
Live weight	**123**	0.78 ± 0.01	0.80 ± 0.01	0.60 ± 0.03	0.62 ± 0.02
Carcass weight		**29.7**	0.94 ± 0.00	0.58 ± 0.03	0.46 ± 0.03
Pelvic limb weight			**3.9**	0.50 ± 0.03	0.42 ± 0.03
Abdominal fat weight				**1.9**	0.63 ± 0.03
Subcutaneous fat weight					**1.7**

All heritability estimates were moderate. Genetic correlations were very high between live weight and carcass weight, as well as between live weight and pelvic limb weight. Carcass weight and pelvic limb weight were also highly correlated, the derived genetic correlation not differing from unity. A high genetic correlation was also found between live weight and subcutaneous fat weight, while the correlation between live weight and abdominal fat weight was lower. The residual and phenotypic correlations between the various weight traits (live weight, carcass weight and pelvic limb weight) and abdominal fat were mostly higher than those with subcutaneous fat weight though. The genetic correlation between abdominal fat weight and subcutaneous fat weight amounted to 0.63, with the 95% confidence interval for the correlation (0.27 - 0.99) just excluding unity. Residual and phenotypic correlations were comparable to genetic correlations in sign, but in some cases somewhat smaller or larger in absolute magnitude.

Moderate to high heritability estimates were found in multi-trait analyses for most of the individual muscle weights, with a range from 0.14 to 0.43. The genetic correlations among these individual muscle weights and pre-slaughter live weight were all positive and ranged from 0.59 to 0.82. Accordingly, genetic correlations among the weights of the respective muscles were also positive and ranged from between 0.45 and 0.99.

All traits therefore showed significant genetic variation in multi-trait analyses, while no significant maternal permanent environmental effect was evident for ostrich carcass and meat traits in these analyses. The estimates from multi-trait analyses were generally slightly higher compared to single-trait estimates. Heritability estimates were comparable to estimates for other species. Lotfi *et al.* (2011), for instance, reported heritability estimates of 0.59 for carcass weight and 0.28 for abdominal fat weight of Japanese quail.

Pre-slaughter live weight was highly correlated with carcass weight and pelvic limb weight. Unfortunately the genetic correlation between live weight and subcutaneous fat was also very high; indicating that selection for increased live weight and slaughter yield will increase the subcutaneous fat weight as well. If this fat can be exploited as a valuable oil, as is done in other ratite species (Sales 2007; Bennett *et al*. 2008), this could be beneficial though. The possible uses of ostrich fat should therefore be investigated further.

CONCLUSIONS

The derived heritability estimates indicate that genetic improvement in ostrich carcass traits is achievable. The estimated genetic relationships are mostly favourable, with a few exceptions, namely those involving fat depots. However, even though ostrich fat is currently treated as a waste product, the possibility for exploiting the fat as valuable oil needs further attention.

REFERENCES

Baeza E., Chartrin P., Meteau K., Bordeau T., Juin H., Le Bihan-duval E., Lessire M. and Berri C. (2010) *Br. Poult. Sci.* **51:**344.

Bennett D.C., Code W.E., Godin D.V. and Cheng K.M. (2008) *Aus. J. Exp. Agr.* **48:**1345.

Gilmour A.R., Gogel B.J., Cullis B.R. and Thompson R. (2009) 'ASReml User Guide, Release 3.0' VSN International Ltd, Hemel Hempstead, HP1 1ES, UK. www.vsni.co.uk

Hoffman L.C., Cloete S.W.P., Van Schalkwyk S.J. and Botha S.StC. (2009) *S. Afr. J. Anim. Sci.* **39:**206.

Lotfi E., Zerehdaran S. and Ahani Azari M. (2011) *Poult. Sci.* **90:**2202.

Rosa F., de Almeida M.I.M., Rigoti F.H. and de Oliveira E.G. (2011) *Rev. Bras. de Zoo.* **40:**1686.

Sales J. (2007) *Avian & Poult. Biol. Rev.* **18:**1.

Zerehdaran S., Vereijken A.L.J., Van Arendonk J.A.M. and Van der Waaij E.H. (2004) *Poult. Sci.* **83:**521.

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:385-388

COMPARING GENOMIC RELATIONSHIP MATRICES WITH RELATIONSHIP ESTIMATED FROM PEDIGREE DATA

M.M. Farah[1], M.R.S. Fortes[2], L.R. Porto-Neto[3], C.T. Meira[1], M. Kelly[2], L.O.C. Duitama[1], A.V. Pires[4], R. Fonseca[1], S.S. Moore[2]

[1]São Paulo State University, Jaboticabal, SP, Brazil. [2]The University of Queensland, Queensland Alliance for Agriculture and Food Innovation, Centre for Animal Science, Brisbane, QLD 4062. [3]CSIRO Food Futures Flagship, and Division of Animal, Food and Health Science, Queensland Bioscience Precinct, Brisbane, QLD 4067. [4]Federal University of the Jequitinhonha and Mucuri Valleys, Diamantina, MG, Brazil

SUMMARY

In this study, we tested 3 methods of building variations of genomic relationship matrix and compared these with the classic **A** matrix (pedigree based). Brahman bulls (n = 1,695) genotyped for or imputed to more than 700,000 single nucleotide polymorphisms were used. The allele frequencies used to obtain the 3 variations of **G** were: 0.5 for all SNPs (**G50**), the average minor allele frequency (**GMF**), and the observed allele frequency of each SNP (**GOF**). Our results indicate that, it is relevant to evaluate the allele frequency in the population and select the method of building matrices to increase the importance of rare alleles, which can help with estimating more precise relationships.

INTRODUCTION

The use of genomic information has been growing in animal breeding programs. Several researchers are using this type of information to improve the accuracy of estimated breeding values (Hayes *et al.* 2010; Gianola *et al.* 2010; Erbe *et al.* 2012). Technology advancement and the possibility of genotyping many individuals made possible to use information on the alleles identical by state (IBS), not only identical by descendent (IBD) that can be shared through common ancestors. It is feasible to use a genomic relationship matrix (**G**) for estimating breeding values (Meuwissen and Goddard 1996).

Often it is not possible to obtain genomic information on the whole population and generate a relationship matrix based entirely on genomics due to the cost of genotyping and lack of samples to genotype. It is also difficult to estimate the allele frequencies of the base population. In the absence of this information, methods were developed to calculate the genomic relationship matrix using either an observed allele frequency, or a minor allele frequency or even a fixed value for allele frequency. These methods use observations from the genotyped population, which may be observed by actual genotyping or inferred with imputation from low density panels to high density panels. Forni *et al.* (2011) used genomic information from a population of pigs (1,919 females and 70 males) to test the impact on breeding values of using different approaches to build the **G** matrix and compared it to the **A** matrix (pedigree based). Forni *et al.* (2011) concluded that the breeding values estimated using the traditional **A** or an **H** matrix, that have both genomic and pedigree information, were similar. Their evidence suggested no real benefit from including genomic information in pig breeding programs. However, population structures in commercial pig lines are very different from the breed structure encountered in the beef cattle industry. Therefore, it is important to evaluate the contribution of genomic information to genetic evaluation processes in beef cattle. In this study, we test 3 methods of building genomic relationship matrices and compare these with the classic **A** matrix.

MATERIALS AND METHODS

Animals and genotypes. Data from 1,695 Brahman bulls were used in the current study. These cattle represent a subset of the population bred by the Beef CRC that was described in detail previously (Burns *et al*. 2013; Corbet *et al.* 2013). This population has information on 729,068 single nucleotide polymorphisms (SNPs). These SNPs were genotyped (97 animals) or imputed (1,598 animals) from a lower density Illumina chip (BovineSNP50). Although only 97 animals within this study were genotyped on the high density marker panel, the full reference used to impute genotypes contained 917 animals from the Beef CRC population. Quality control criteria excluded SNP if minor allele frequency was lower than 0.05. Also, if pairwise correlations between SNP alleles was stronger than 0.95, only one SNP of the pair remained in the analysis. After quality control, 569,620 SNPs were used to estimate **G**, as follows:

$$G = \frac{(M-P)(M-P)'}{2\sum_{j=1}^{m} p_j(1-p_j)} \ ,$$

where **M** is an allele-sharing matrix with *m* columns (*m* = 569,620 SNPs) and *n* rows (*n* = 1,695 individuals), and **P** is a matrix containing the frequency of the second allele (p_j), expressed as $2p_j$. $\mathbf{M}_{ij}$ was 0 if the genotype of individual *i* for SNP *j* was homozygous for the first allele, was 1 if heterozygous, or 2 if the genotype was the other homozygous state.

The frequencies used to obtain 3 variations of **G** were similar to the methods described by Forni *et al.* (2011) where **P** matrix was obtained with: 1) the observed allele frequency of each SNP for the population (**GOF**), 2) average minor allele frequency (**GMF**), and 3) 0.5 for all SNPs (**G50**). For comparison between these variations of **G** matrices and the **A** matrix two methods were used: descriptive statistics and the correlation between the estimated kinship of individuals. For this population, 7 generations pedigree was used to obtain the relationship between the genotyped animals, underpinning the A matrix (total number of animals 3030). The comparison between **A** and **G** variations was made using only the relationship estimated between genotyped individuals.

RESULTS AND DISCUSSION

Descriptive statistics for the **A** relationship matrix and the **G** relationship matrices, estimated for genotyped animals are provided in Table 1. The diagonal and off-diagonal elements were most similar for the matrices **A** and **GOF**, but the variances were very different. This lead to the differences between the matrices that can also be observed in Figure 1.

Table 1. Statistics of relationship coefficients estimated using pedigree and genomic data*

	Diagonal elements				**Off-diagonal elements**			
	Mean	Min.	Max.	Var.	Mean	Min.	Max.	Var.
A	1.00	1.00	1.12	$3.7\text{x}10^{-5}$	0.01	0.00	0.62	$1.4\text{x}10^{-3}$
GOF	1.03	0.90	1.26	$3.4\text{x}10^{-3}$	0.00	-0.11	0.66	$1.9\text{x}10^{-3}$
GMF	2.84	2.57	3.08	$3.6\text{x}10^{-3}$	1.91	1.55	2.58	$5.7\text{x}10^{-3}$
G50	1.36	1.20	1.52	$1.4\text{x}10^{-3}$	0.68	0.45	1.16	$2.6\text{x}10^{-3}$

*__A__ (relationship matrix pedigree-based); **GOF** (genomic relationship matrix with observed frequency); **GMF** (genomic relationship matrix with averaged minor allele frequency); **G50** (genomic relationship matrix with frequency 0.5 for all alleles).

Differences between the estimates for kinship based in either **A** or **G** calculations were observed (Figure 1). For some pairs of animals that **A** estimated as having no relationship (a value of zero), **G** matrices estimated values higher than zero suggesting that some of these animals share

alleles that are IBS, but may not be IBD, or they may have a common ancestor that was missing from the pedigree records.

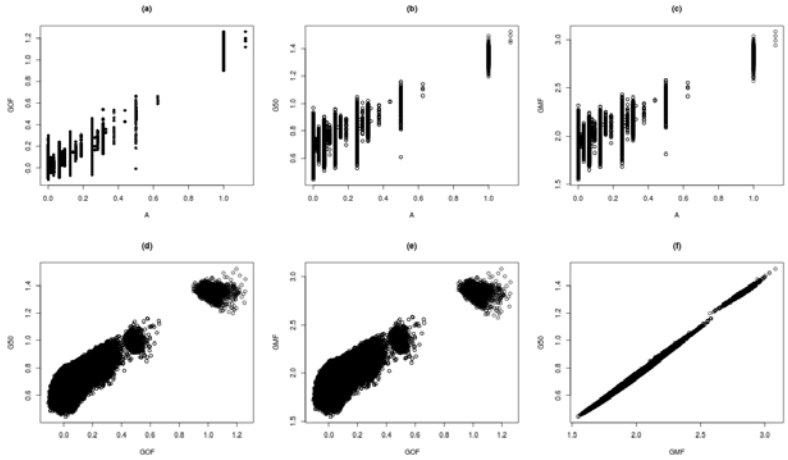


Figure 1. Pairwise comparisons between kinship values estimated by matrices: a) A vs. GOF, b) A vs. GMF, c) A vs. G50, d) GOF vs. GMF, e) GOF vs. GMF, and f) GMF vs.G50.

Lower variances for the matrix **A**, compared to **G**, can be explained because the method of calculating the relationship is by probability of two individuals sharing only alleles IBD. Higher variance among **G** elements, compared to **A**, can be expected because genomic relationships considered both alleles IBS and IBD, in agreement with Forni *et al.* (2011). Differences between matrix element variances are also reflected in the estimated correlations between **A** and **G** (Table 2). When the 3 variations of **G** were compared, we observed greater differences between **GOF** and the other 2 **G** matrices. The correlation between **G50** and **GMF** was high (Table 2), reflecting similar relationships estimated by these 2 methods (Figure 2).

Table 2. Correlations between individual kinship estimates from each relationship matrix*

	GOF	**GMF**	**G50**
A	0.85	0.50	0.54
GOF		0.58	0.63
GMF			0.99

***A** (pedigree based relationship matrix); **GOF** (genomic relationship matrix with observed allele frequencies); **GMF** (genomic relationship matrix with averaged minor allele frequencies); **G50** (genomic relationship matrix with a fixed allele frequency of 0.5 for all SNP).

The distribution of allele frequencies were shown in Figure 2. Extreme allele frequencies (higher than 0.8 and lower than 0.2) were observed in this population of Brahman bulls. This distribution of allele frequencies is in contrast with the distribution observed in pigs by Forni *et al.* (2011) that the distribution was much more uniform. The distribution of allele frequencies reflect the fact that the BovineSNP50 chip was developed primarily for *Bos taurus.* In *Bos taurus*, the allele frequencies are much more uniform in comparison to the distribution found in our Brahman population. The extreme frequencies in our allele distribution might help to explain the higher correlation between **GMF** and **G50**, and the differences between these 2 and **GOF**. Assumingly, if a fixed allele frequency such as 0.5 is used or minor allele frequencies are used instead of the observed frequencies, less importance is given to rare alleles and individual allele variation. As a

result, **GMF** and **G50** might estimate higher values of kinship and some individuals might be perceived to be more related than suggested by **GOF** or **A** results. Differences in estimated kinship will influence estimated breeding values, having an impact on cattle selection programs.

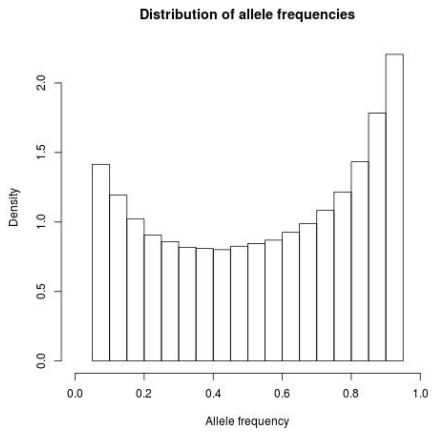


Figure 2. Distribution of observed frequencies of the second allele.

CONCLUSION

In this study, relationships between individuals estimated using genomic data were correlated with estimates based on pedigree information. Since **G** matrices are correlated but not identical to the **A** matrix, genomic data can add information and contribute to accurate relationship estimations. Appropriate use of genomic information can be achieved with different methods of calculating a **G** matrix. Our results indicate that it is relevant to evaluate the allele frequency in the target population and select the methodology accordingly. Presence of extreme allele frequencies might indicate the importance of rare alleles and the use of **GOF**. Future work should test the influence of the different **G** matrices in the estimative of breeding values.

ACKNOWLEDGEMENTS

The authors acknowledge that this research uses resources build by the Cooperative Research Centre for Beef Genetic Technologies (Beef CRC) and the financial support for genotyping Brahman animals was provided by Meat and Livestock Australia (project code B.NBP.0723). We thank the support of CAPES (Process: 13843/12-5). The Lab of scientific computation applied to animal science (LuCCA-Z) and QAAFI are acknowledge for providing the structure available.

REFERENCES

Burns B.M, Corbet N.J., Cobert D.H., Crisp J.M., Venus B.K, Johnston D.J, Li Y, McGowan M.R and Holroyd R.G. (2013) *Anim Prod Sci* **53:**87.

Corbet N.J, Burns B. M, Johnston D. J, Wolcott M. L, Corbet D. H, Venus B. K, Li Y, McGowan M. R and Holroyd, R. G (2013) *Anim Prod Sci* **53**:101.

Erbe M., Hayes B. J, Matukumalli L. K, Goswami S, Bowman P. J, Reich C. M, Mason B. A and Goddard M. E (2012) *J Dairy Sci* **95**: 4114.

Forni S., Aguilar I., and Misztal I. (2011) *Genet Sel Evol* **43**:1.

Gianola D., Simianer H., and Qanbari S. (2010) *Genet Res (Camb)* **92**: 141.

Hayes B. J., Pryce J., Chamberlain A.J., Bowman P.J. and Goddard M. E. (2010) *PLoS Genet* **6**: e1001139.

Meuwissen T., and Goddard M. (1996) *Genet Sel Evol* **28**: 161.

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:389-392

GENETIC MARKERS ASSOCIATED WITH MALE REPRODUCTIVE TRAITS ACROSS 2 BEEF CATTLE BREEDS: BRAHMAN AND TROPICAL COMPOSITE

M.R.S. Fortes[1], A. Reverter[2], L.R. Porto Neto[2], M. Kelly[1], S.S. Moore[1] and S.A. Lehnert[2]

[1]The University of Queensland, Queensland Alliance for Agriculture and Food innovation, Centre for Animal Science, Brisbane, Qld 4062
[2]CSIRO Food Futures Flagship, and Division of Animal, Food and Health Science, Queensland Bioscience Precinct, Brisbane, Qld 4067

SUMMARY

We report chromosomal regions identified as significant for bovine male fertility according to genome-wide association studies carried out in two independent populations of cattle. Reported chromosomal regions harboured single nucleotide polymorphisms (SNP) that were associated ($P < 0.01$) with inhibin, insulin growth factor 1 (IGF1), scrotal circumference (SC) or percentage of normal sperm (PNS) in both Brahman (n = 1,130) and Tropical Composite (n = 1,085) cattle. Bulls were genotyped with Illumina SNP chips (50K and 700K) and association analyses were performed using animal models. Chromosomes 2, 3, 5, 7 and 10 had SNP that were associated with inhibin in both breeds. SNP associated with IGF1 were located on chromosomes 5, 6, 10 and 14 in both breeds. SNP associated with SC mapped to chromosomes 9, 13 and X in both breeds. Only chromosome X had SNP associated to PNS in both breeds. Comparing the associations of SNP to traits measured in both Brahman and Tropical Composite cattle breeds is an important validation strategy for selecting markers that could be used for genomic selection in a multi-breed program. Markers associated with inhibin, IGF1, SC and PNS may contribute to the selection of bulls with improved reproductive performance.

INTRODUCTION

Single nucleotide polymorphisms (SNP) are used in genomic selection. Association of SNP across beef cattle breeds provides validation for independent genome-wide association studies (GWAS) and may contribute to increase the accuracy of genomic selection. Validation of SNP is a step towards the discovery of causative mutations that could aid genomic selection and genetic gain (Weller and Ron 2011; Snelling *et al.* 2012). Causative mutations have an advantage in comparison to random SNP: they are not dependent on linkage disequilibrium (LD), and so they can be used for selection over generations, across breeds and in breeds that were not in the reference population.

Reproductive performance of bulls has an impact on the economic gain of a farm. Measuring correlated traits, such as scrotal circumference (SC) and percent normal sperm (PNS) allows for selection of bulls with improved reproductive performance (Holroyd *et al*. 2002; Moser *et al.* 1996). Hormonal levels of inhibin and insulin growth factor 1 (IGF1) correlate with reproductive traits and may also aid selection (Corbet *et al.* 2013). The aim of this study is to report validated SNP associated with inhibin, IGF1, SC and PNS, by comparing GWAS carried out in 2 independent populations of bulls: Brahman (BRAH) and Tropical Composites (TC). Only BRAH results were published previously (Fortes *et al.* 2012). Reports of validated SNP associations point to genomic regions that merit further research targeting the discovery of causative mutations for fertility in bulls.

MATERIALS AND METHODS

Animals, Traits and Genotypes. Blood samples for DNA extraction were obtained from 1,130 BRAH and 1,085 TC bulls. These bulls were bred by the Cooperative Research Centre for

Beef Genetic Technologies and details concerning project design and measurement of reproductive traits have been reported (Burns *et al*. 2013; Corbet *et al*. 2013). In short, blood levels of inhibin were measured at 4 months of age, circulating IGF1 at 6 months, scrotal circumference (SC) at 12 months and percent normal sperm (PNS) at 24 months. BovineSNP50 chips (Matukumalli *et al*. 2009) were used to genotype all bulls. Some samples were replicated for quality control and Bead Studio software (Illumina Inc., San Diego, CA 2006) was used to call genotypes. SNP with call rates < 80% or minor allele frequency < 0.01 were discarded. High-density (HD) genotyping of selected TC cattle was performed. Missing 50K genotypes for BRAH and HD genotypes for TC were imputed using BEAGLE (Browning and Browning 2010). Quality control and imputation resulted in 50,354 SNP genotypes for 1,115 BRAH and 729,068 for 1,019 TC.

Statistical Analyses. GWAS were performed for each breed and each trait separately. SNP effects were estimated using an animal model. Solutions were estimated with Qxpak5 (Perez-Enciso and Misztal 2011), using a likelihood ratio test to compare the model with versus the model without each SNP against a chi-squared distribution with 1 degree of freedom. This test was performed for one SNP at a time.

RESULTS AND DISCUSSION

Table 1. SNP associated with reproductive traits in Brahman and Tropical Composite bulls*

Chromosome	Mb (number of sig. SNP $P < 0.01$)	Brahman		Tropical Composite	
		max	min	max	min
Inhibin					
2	104(1)	8.1×10^{-3}		4.3×10^{-3}	
3	60(1)	2.1×10^{-3}		7.0×10^{-4}	
5	108(1)	2.8×10^{-3}		5.8×10^{-3}	
7	73(1)	5.1×10^{-3}		3.6×10^{-3}	
10	17(1)	5.0×10^{-3}		3.8×10^{-3}	
IGF1					
5	34(2), 42(1)	7.0×10^{-7}	6.8×10^{4}	2.7×10^{-3}	8.7×10^{-3}
6	105(1)	1.4×10^{-3}		7.0×10^{-5}	
10	47(1)	7.8×10^{-3}		9.7×10^{-3}	
14	21(1), 23(1), 24(2), 25(7), 26(1), 27(1), 28(4), 30(2), 32(1), 33(1)	1.0×10^{-16}	5.9×10^{3}	3.1×10^{-8}	9.6×10^{-3}
SC					
9	91(1)	4.3×10^{-3}		6.4×10^{-3}	
13	78(1)	6.9×10^{-3}		2.0×10^{-3}	
X	54(1), 60(1), 62(1), 63(2), 65(1), 66(6), 68(2), 69(3), 70(2), 71(1), 72(2), 73(3), 75(1), 76(1), 77(1), 80(1), 81(3), 82(2), 84(2), 85(3), 86(3), 87(2), 91(2), 92(4), 93(3), 94(1), 98(1), 100(1), 102(2), 105(2), 108(2)	4.9×10^{-10}	8.7×10^{-3}	1.1×10^{-29}	3.6×10^{-3}
PNS					
X	40(3), 41(1), 43(1),47(1), 50(1),52(1),53(1),55(1)	6.9×10^{-7}	3.3×10^{-13}	7.9×10^{-7}	3.8×10^{-10}

*Traits: Inhibin, IGF1, scrotal circumference (SC) and percentage of normal sperm (PNS) in Brahman and Tropical Composite bulls. Mega base pairs (Mb) position, number of significant SNP within the Mb, and minimum and maximum *P*-values are reported for each breed.

The GWAS performed in 2 independent populations revealed SNP that were associated ($P < 0.01$) across 2 beef cattle breeds (Table 1). More than one chromosome had SNP associated with BRAH and TC for Inhibin, IGF1 and SC. Only chromosome X had validated SNP associated with PNS, but these were located in multiple regions. Together, BRAH and TC results are evidence for polygenic regulation of these reproductive traits.

The X chromosome harboured validated SNP for SC and PNS, spread across millions of base pairs (Figure 1). These regions with multiple associated SNP might be an indication for multiple quantitative trait loci (QTL). Within our results, SNP associations that point to QTL close to 48 and 110 Mb of the X chromosome provide supporting evidence for results that were first reported in Holstein bulls (Blaschek *et al*. 2011). Further, candidate genes underpinning these QTL on the X chromosome were proposed in the previous Brahman study (Fortes *et al.* 2012). For example, the androgen receptor gene (*AR*) localized at 88 Mb of the X chromosome is a candidate gene for SC, due to the position of associated SNP and its physiological role (Quigley 1998).

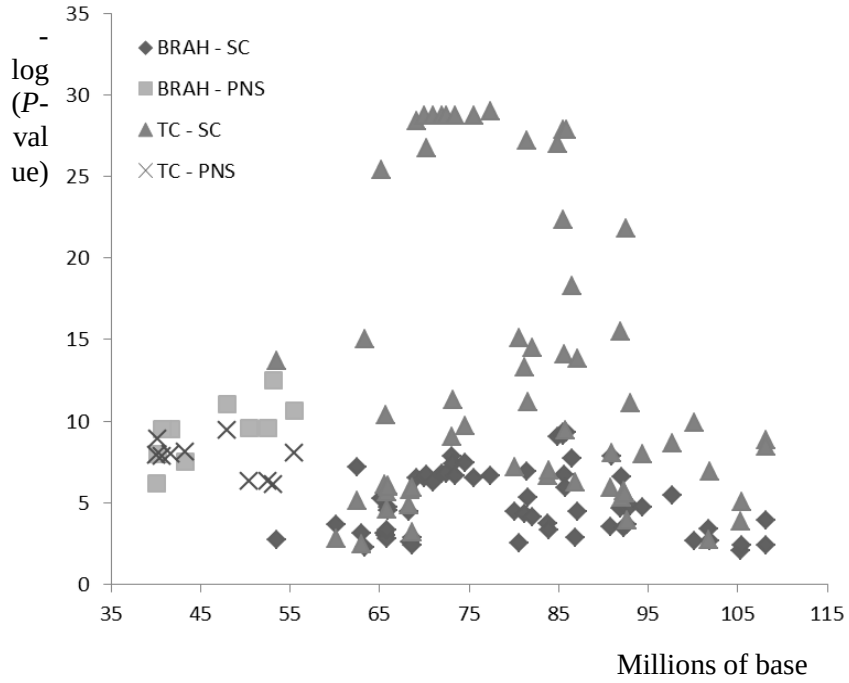


Figure 1. Validated polymorphisms in the X chromosome. Acronyms in figure: Brahman (BRAH), Tropical Composite (TC), Scrotal Circumference (SC), Percent Normal Sperm (PNS).

The region around 25 Mb of chromosome 14 had the highest number of validated SNP for IGF1 and it confirms a known QTL associated with IGF1 in female cattle as well as height, weight and puberty, across various breeds (Karim *et al*. 2011; Littlejohn *et al*. 2011; Hawken *et al*., 2012; Nishimura *et al*. 2012). A putative causative mutation on chromosome 14 near the pleiomorphic adenoma 1 (*PLAG1*) gene was proposed by a study on Holstein and Jersey cattle (Karim *et al*. 2011). However, a direct effect of this mutation on IGF1 levels remains to be investigated. The molecular mechanism linking *PLAG1* function to IGF1 levels is unclear. It is possible that *PLAG1* acts as a transcription factor regulating the expression of the *IGF1* gene.

CONCLUSION

The QTL presented here for inhibin, IGF1, SC and PNS were confirmed across independent cattle populations. These validated QTL point to genomic regions that merit further research, specifically targeting the discovery of causative mutations affecting reproductive traits in bulls. A putative causative mutation for chromosome 14 was proposed and merits functional investigation. Causative mutations underpinning the other validated QTL are unknown. Targeting candidate genes that emerge from the cross-validation of significantly associated SNP between BRAH and TC could lead to the discovery of causative mutations. Knowledge on causative mutations would improve the accuracy of genomic selection and facilitate its use across cattle breeds and over multiple generations.

ACKNOWLEDGEMENTS

Authors acknowledge that this research uses resources built under the CRC for Beef Genetic Technologies. We acknowledge Dr. R. G. Holroyd for leading the experiments that created the Beef CRC male fertility dataset and Ms B. K. Venus for assessing of sperm morphology. Support for TC genotyping was provided by Meat and Livestock Australia (project code B.NBP.0723).

REFERENCES

Blaschek M., Kaya A., Zwald N., Memili E. and Kirkpatrick B.W. (2011) *J. Dairy Sci.* **94**: 4695.

Burns B.M., Corbet N.J., Corbet D.H., Crisp J.M., Venus B.K., Johnston D.J., Li Y., McGowan M.R., and Holroyd R.G. (2013) *Anim. Prod. Sci.* **53**: 87.

Corbet N.J., Burns B.M., Johnston D.J., Wolcott M.L., Corbet D.H., Venus B.K., Li Y., McGowan M.R., and Holroyd R.G. (2013) *Anim. Prod. Sci.* **53**: 101.

Fortes M.R.S., Reverter A., Hawken R., Bolormaa S. and Lehnert S. (2012) *Biol. Reprod.* **87**: 58.

Hawken R.J., Zhang Y.D., Fortes M.R.S., Collis E., Barris W.C., Corbet N.J., Williams P.J., Fordyce G., Holroyd R.G., and Walkley J.R.W. (2012) *J. Anim. Sci.* **90**: 1398.

Holroyd R.G., Doogan W., De Faveri J., Fordyce G., McGowan M.R., Bertram J.D., Vankan D.M., Fitzpatrick L.A., Jayawardhana G., and Miller R.G (2002) *Anim. Reprod. Sci.* **71**: 67.

Karim L., Takeda H., Lin L., Druet T., Arias J.A.C., Baurain D., Cambisano N., Davis S.R., Farnir F., Grisart B., Harris B.L., Keehan M.D., Littlejohn M.D., Spelman R.J., Georges M., and Coppieters W. (2011) *Nature Genetics* **43**: 405.

Littlejohn M., Grala T., Sanders K., Walker C., Waghorn G., Macdonald K., Coppieters W., Georges M., Spelman R., Hillerton E., Davis S. and Snell R. (2011) *Animal Genetics* **43**: 591.

Matukumalli L.K., Lawley C.T., Schnabel R.D., Taylor J.F., Allan M.F., Heaton M.P., O'Connell J., Moore S.S., Smith T.P., Sonstegard T.S., and Van Tassell C.P (2009) *PLoS One* **4**: e5350.

Moser D.W., Bertrand J.K., Benyshek L.L., McCann M.A., and Kiser T.E. (1996) *J. Anim. Sci.* **74**: 2052.

Nishimura S., Watanabe T., Mizoshita K., Tatsuda K., Fujita T., Watanabe N., Sugimoto Y., and Takasuga A. (2012) *BMC Genetics* 13.

Perez-Enciso M. and Misztal I. (2011) *BMC Bioinformatics* 12.

Quigley C.A. (1998) In "Testosterone: action, deficiency, substitution.", pp.33-106, editor Nieschlag E. and Behre H.M. Heidelberg, Springer-Verlag.

Snelling W.M., Cushman R.A., Fortes M.R.S., Reverter A., Bennett G.L., Keele J.W., Kuehn L.A., McDaneld T.G., Thallman R.M., and Thomas M.G. (2012) *J. Anim. Sci.* **90**, 1152.

Weller J.I. and Ron M. (2011) *J. Dairy Sci.* **94**: 1082.

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:393-396

GENETIC PARAMETERS FOR STAPLE STRENGTH AND COEFFICIENT OF VARIATION OF FIBRE DIAMETER IN MERINO WOOL OF DIFFERENT STAPLE LENGTH

J C Greeff and A C Schlink
Department of Agriculture and Food Western Australia, 3 Baron Hay Court Perth, WA, 6051

SUMMARY

Staple strength (SS) and coefficient of variation of fibre diameter (FDCV) were measured on full length staples from 12 months of wool growth, and wool staples cut to 80 or 60 % of their full length from the tip. The results show that SS and FDCV for the different staple length treatments are heritable but lower for staples 60% or shorter than their full length. SS and FDCV measured on wool staples greater than 60 % of their normal length are genetically the same trait.

BACKGROUND

Staple strength (SS) is the second most important wool trait affecting the price of wool after fibre diameter, and tender wools are consistently penalised, especially for the finer types. Stott (2004), as quoted by Smith *et al.* (2004), indicated that SS was the single largest contributing factor to variation in price received for wool less than 19.5 micron. SS is a heritable trait exhibiting large phenotypic variation (Greeff *et al.* 1995) and it will respond to selection (Greeff *et al.* 1997). SS is generally expensive to measure but it is genetically highly correlated to coefficient of variation of fibre diameter (FDCV). The latter trait is obtained automatically when measuring fibre diameter with the Optical Fibre Diameter Analyser (OFDA) or Laserscan®. FDCV is therefore an effective indicator trait to indirectly select for SS.

However, ram breeders have recently expressed concern that SS could be affected by staple length (SL). Anecdotal information indicates that shorter staples from premature shearing have higher SS than full grown staples, which could impact on the effectiveness of SS measurements for selection purposes. This paper aims to determine whether reducing SS has a significant effect on the genetic parameters of SS and FDCV measurements and whether SS should be adjusted for SS to increase the accuracy of breeding values for SS.

MATERIAL AND METHODS

Midside wool samples were collected from 2642 Merino hoggets, born from 1992 to 2005 in a fully pedigreed Merino resource flock based at Katanning, Western Australia. The progeny were produced by 120 sires that were mated to 1081 dams. The establishment of this flock was described by Greeff *et al.* (1997) and consisted of high, average, and low staple strength lines. The animals were born in July/August in a Mediterranean environment and managed as one group from marking to weaning. All males were castrated during the first 4 years. In 1996 every second male within a sire progeny group was castrated at marking. From 1997 all males were left intact. After weaning at approximately 100 days of age, the animals were separated on sex and managed separately up to hogget shearing.

Sampling of wool. All lambs received an even-up shearing after weaning in November and were shorn again as hoggets in November the following year. Wool samples were collected with Oster® clippers immediately prior to shearing, on the midside of all animals. Ten pencil sized wool staples were pulled from each midside sample and measured for SS (SSMS) and SS.

From 1997, 1060 individual fleeces of the 1996, 1997 and 1998 born progeny were stored in separate plastic bags. These fleeces were re-sampled in 2012 at the estimated midside site. All the

fleeces were in good condition and allowed to recover from compaction in the bale before sampling. Thirty pencil sized staples were pulled from each of the stored fleece samples.

Measurement of staple sections. SS of the 2642 midside samples were measured with using ATLAS. SS of the 1060 stored fleeces, were measured with an Agritester. SS can only be reliably measured on staples that are longer than 40mm (MicronMan, personal communication). Therefore, from each stored fleece, SS was measured on 10 full length wool staples (SS100), on 10 staples that were cut with a guillotine at 80% of their length (SS80), and on a further 10 wool staples that were cut at 60% of their length (SS60). All staples were measured and cut from the tip end of the staple to achieve the required length reductions in order to leave the tip intact.

FDCV of all the wool samples were measured with an OFDA2000 but the original midside samples were measured on an OFDA100. The traits were defined as CVMS (FDCV of the midside samples), CV100, CV80 and CV60 (FDCV of staples from stored fleeces not cut, cut at 80% or at 60% of their length, respectively), and CVbutt80 and CVbutt60 (FDCV of leftover butts from staples of stored fleeces cut at 80% or 60% of their length).

Data analysis. The data were analysed with ASREML (Gilmour *et al.* 2009). An animal model was fitted with year of birth, sex, age of the dam and birth status as fixed factors and all 2 way interactions. Univariate analyses were carried out to identify significant fixed effects, followed by bivariate analysis to obtain variances and covariances for genetic parameter estimation.

RESULTS AND DISCUSSION

The average fibre diameter of this flock was 19.8 ± 1.5 micron with an average SS of 107 ± 12 mm for wool samples collected on the midside. The average SS of the full length staples sampled from the stored fleeces was 99 ± 10 mm, demonstrating that wool samples collected on the shearing board were shorter on average than midside wool samples removed with Oster® clippers prior to shearing. The average SS of the staples cut at 80% of their SS was 78 ± 9 mm, and 61 ± 6 mm for staples cut at 60% of their length. The butts of the 80% and 60% cut staples were 21 mm and 38 mm long, respectively, which relates to 20% and 40% of the full length staples. Thus, the data set consisted of FDCV measured on midside staples (CVMS), on staples that were collected on the shearing board (CV100), and on staples that were 80% (CV80), 60% (CV60), 40% (CVbutt60) and 20% of full length staples (CVbutt80).

A significant sex x birth year interaction ($P<0.01$) was found for all the SS and FDCV traits, while age of the dam, birth status and day of birth were not statistically significant ($P>0.10$). The sex x birth year interaction could be explained by the fact that the males and females were managed separately from weaning to hogget shearing. Table 1 shows the basic statistics in wool from different length staples treatments.

SS of the different length treatments differed significantly ($P<0.01$). Table 1 show that SSMS exhibited double the amount of phenotypic variation compared to SS100, SS80 or SS60. However, SSMS and SS100, SS80 and SS60 measurements were carried out with different machines and with different wool sampling protocols. Pre-testing evaluation of the Agritester and ATLAS found that both instruments produced the same outcomes in SS when evaluated using a common set of samples. Table 1 also shows that as SS reduced, SS increased from 27.4 N/Ktex for SS100 to 31.3 N/Ktex for the SS60 length staples.

CVMS showed higher levels of phenotypic variation compared to CV100, CV80, CV60, CVbutt60 and CVbutt80. CVMS was measured on an OFDA100 whereas the CV100, CV80, CV60, CVbutt60 and CVbutt80 measurements were measured on single, greasy staples using OFDA2000. FDCV tended to decrease as the staples became shorter.

Table 1. Number of samples, mean, standard deviation (SD), minimum, maximum and the phenotypic variation (Vp) of staple strength and coefficient of variation of fibre diameter traits for wool with different length staple outcomes

Trait	n	Mean	SD	Min	Max	Vp
SSMS (N/Ktex)	2642	30.9	10.76	3.7	69.6	84.53
SS100 (N/Ktex)	1060	27.4	7.85	4.0	74.5	44.10
SS80 (N/Ktex)	1060	28.9	7.44	4.9	54.0	40.11
SS60 (N/Ktex)	1060	31.3	7.52	9.9	99.4	44.23
CVMS (%)	2642	21.8	3.08	14.3	35.8	8.08
CV100 (%)	1060	22.3	2.93	15.6	35.0	6.98
CV80 (%)	1060	19.8	2.40	14.8	28.6	4.75
CV60 (%)	1060	19.9	2.63	13.5	29.0	5.38
CVbutt60 (%)	1060	17.3	2.43	12.4	27.5	5.05
CVbutt80 (%)	1060	16.9	2.30	12.0	26.7	5.16

Heritability estimates. The heritability estimates for the SS measurements are shown in Table 2. Except for the low estimate of 0.31 for the SS60 measurement, the heritability of the other SS traits agrees with previous (Greeff *et al*. 1995; Greeff and Paganoni 2004). This shows that the heritability of SS is low in very short staples. However, adjusting SS measurements for SS resulted in no significant changes in heritability estimates for SS. The only improvement was a small increase from 0.31 ± 0.07 to 0.33 ± 0.07 for SS60.

Table 2. Heritability (on diagonal) of, and the phenotypic (above diagonal) and genetic (below diagonal) correlations between the different staple strength measurements and their standard errors in brackets

	SSMS	SS100	SS80	SS60
SSMS	**0.44 (0.05)**	0.73 (0.01)	0.69 (0.02)	0.51 (0.02)
SS100	0.99 (0.02)	**0.50 (0.08)**	0.69 (0.02)	0.48 (0.03)
SS80	0.99 (0.02)	1.01 (0.02)	**0.49 (0.08)**	0.52 (0.02)
SS60	0.85 (0.08)	0.98 (0.05)	0.97 (0.05)	**0.31 (0.07)**

The heritability estimates of the CVFD traits are shown in Table 3. All the h^2 estimates were above 0.5 with CV100 having the highest heritability estimate of 0.73 ± 0.07. The lowest heritability was found for CVbutt80 (0.51 ± 0.08) which represents only 20% of the staple. However, all these estimates agree strongly with published results in the literature (Greeff *et al*. 1995) and on estimates of short wool sections (Greeff and Paganoni 2004). Adjusting FDCV traits for SS had no effect on the heritability of FDCV traits.

Phenotypic correlations. Table 2 shows the phenotypic correlations between the different SS measurements, while that of the CVFD traits are shown in Table 3. Moderately high phenotypic correlations were found between the SS measurements with the lowest correlation between SSMS

Table 3. Heritability (on diagonal) and the phenotypic (above diagonal) and genetic (below diagonal) correlations between the different coefficient of variation of fibre diameter measurements (standard errors (SE) in brackets)

Trait	CVMS	CV100	CV80	CV60	CVbutt60	CVbutt80
CVMS	**0.63 (0.04)**	0.66 (0.02)	0.67 (0.02)	0.69 (0.02)	0.62 (0.02)	0.60 (0.02)
CV100	0.97 (0.02)	**0.73 (0.07)**	0.59 (0.02)	0.63 (0.02)	0.61 (0.02)	0.57 (0.02)
CV80	0.97 (0.02)	0.97 (0.02)	**0.53 (0.05)**	0.72 (0.02)	0.60 (0.02)	0.65 (0.02)
CV60	1.02 (0.02)	0.97 (0.02)	0.97 (0.02)	**0.61 (0.07)**	0.68 (0.02)	0.60 (0.02)
CVbutt60	1.00 (0.02)	0.97 (0.03)	0.96 (0.03)	0.95 (0.03)	**0.62 (0.07)**	0.70 (0.02)
CVbutt80	1.02 (0.02)	0.98 (0.03)	1.01 (0.02)	1.02 (0.02)	1.02 (0.01)	**0.51 (0.08)**

and SS60. The correlations between SS measurements of the different length staples decreased from 0.73 for SSMS and SS100, to 0.69 for SSMS and SS80, and to 0.51 for SSMS and SS60. The moderately high phenotypic correlation between SSMS, SS100, SS80 and SS60 of greater than 0.5, suggests that the measurements are repeatable.

The phenotypic correlations between the different FDCV traits are shown in Table 3. Again moderately high correlations were found with the lowest correlation of 0.57 being between CV100 and CVbutt80. No clear pattern was found as the staples became shorter indicating that FDCV of a short section of wool gives a reliable indication of FDCV of whole staple.

Genetic correlations. The genetic correlations among the SS measurements (Table 2) and among the FDCV measurements (Table 3) of the different length staples were generally very high. Some estimates fall outside the parameter space but were not significantly different from unity. This indicates that the same genes are controlling SS in staples that are up to 60 % shorter than full length staples.

CONCLUSIONS

This study shows that SS measured on short staples are heritable and genetically representative of SS measured on 12 month wool. SS measured on staples that are approximately 60 % or longer than their normal length, is genetically the same trait. Similar results were obtained for FDCV on staples of different length. The results confirm that SS and FDCV are reliable measurements to estimate a breeding value for SS in Merino sheep. However more work is needed to determine whether this finding will also apply in production systems with 8 months or shorter shearing times.

REFERENCES

Gilmour A.R., Gogel B.J., Cullis B.R. and Thompson R. 2009. ASREML User Guide Release 3.0 VSN International Ltd., Hemel Hempstead. HP1 1ES, UK. www.vsni.co.uk.

Greeff J.C., Lewer R.P., Ponzoni R.W. and Purvis I. (1995) *Proc. Aust. Assoc. Anim. Breed. Genet*. **11**:595.

Greeff J.C., Ritchie A.M. and Lewis R.M. (1997) *Proc. Assoc. Advmt. Anim. Breed. Genet.* **12**:714.

Greeff J.C. and Paganoni B. (2004) *Aust J. Exp. Agric.* **45:** 347.

Smith J., Purvis I.W., and Lee G.J. (2004) Proc. of the 2006 Australian Sheep Industry CRC Conference Orange, pp 170. Editor P.B. Cronje and D.K. Maxwell.

Stott K. (2004) Interactive pricemake On-line. http://www/pricemaker.info/index.html.

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:397-400

DOGS CAN DIFFERENTIATE BETWEEN ODOURS FROM SHEEP THAT ARE RESISTANT OR SUSCEPTIBLE TO BREECH STRIKE

J C Greeff[1], A Biggs[2], W Grewar[2], P Crumblin[2], L J E Karlsson[1], A C Schlink[1] and J Smith[3]

[1]Department of Agriculture and Food Western Australia, Perth, WA
[2]Hanrob Dog Academy, 1800 Princess Hwy, Heathcote, NSW
[3]CSIRO, Armidale, NSW

SUMMARY

Merino ewes that were genetically resistant or susceptible to breech strike were identified in the Australian Wool Innovations breech strike flocks at Mt Barker (WA) and in Armidale (NSW). Wool crutchings were regularly collected on the Mt Barker sheep and used to train two dogs over a 12 month period to identify animals from the resistant line as the target group and to ignore wool from the susceptible group. After successful training, the dogs were evaluated in a test to determine whether the dogs can successfully differentiate between crutched wool samples from high and low breech strike resistant Merino ewes in the Armidale flock that were raised in a different environment. The results showed that the dogs could identify the resistant animals with an accuracy of 82%, and can ignore the susceptible animals with an accuracy of 92%.

INTRODUCTION

Breech strike is a serious disease of wool sheep. Seddon *et al.* (1931) showed that wrinkles play an important part in making sheep more susceptible to breech strike. Greeff *et al.* (2009; 2013) and Smith *et al.* (2009) confirmed the role wrinkles play and showed that dags, breech cover and urine stain also contribute to make sheep more susceptible to breech strike. However, these traits explain only about 15-20% of the total variation in breech strike and dags was the most important and explained up to 10% of the total variation in breech strike (Greeff, unpublished) . This indicates that other traits may be involved in making sheep more or less resistant to breech strike.

Ashworth and Wall (1994) showed that putrefactive sulphur rich compounds originating from bacterial decomposition in fleece rot and Dermatophilosis affected sheep, attract *Lucilia cuprina* and *Liculia sericata*. However, recent research has found that specific semiochemicals secreted by the animals attract and/or repel horn flies in cattle (Oyarzun *et al* 2009). Similar results have been found for biting midges in humans in Scotland (Logan *et al.* 2009). No research has been carried out on unstruck sheep that are genetically resistant or susceptible to breech strike to find out whether genetically resistant sheep secrete specific semiochemicals that repels or attract blowflies. This paper reports on preliminary findings from training dogs to detect odour differences between unstruck sheep that were genetically resistant or susceptible to breech strike.

MATERIAL AND METHODS

Animals

Sheep (Mt Barker flock). Records of 748 unmulesed Merino hoggets, that were the progeny of 21 individual sires which were born in 2008 on the Mt Barker research station in Western Australia were available for this study. The Mt Barker Research station is situated in a rainfall region with an annual rainfall of approximately 700mm. These animals were naturally challenged by flies

from birth to post hogget shearing. They were only shorn after weaning or shorn as hoggets after the fly season expired. No preventative fly treatments such as crutching and jetting were applied to any of the sheep to ensure that the animals were appropriately challenged. Animals were allowed to be struck naturally by flies, and any struck sheep was identified, treated with a short acting chemical and returned to the flock. The total number of breech strikes from birth to hogget shearing, were recorded on all animals. After hogget shearing, the ewes were crutched prior to lambing in winter but no other preventative treatments were applied. All flystrikes were recorded on the ewes during their lifetime in the flock.

The 2 most resistant and 2 most susceptible sire progeny groups were identified after hogget shearing, and within each sire progeny group the three most resistant and three most susceptible ewes were identified. None of the six resistant ewes were ever struck up to the 2012 shearing while every ewe in the susceptible group was struck in 2008, 2009 and 2010. A very poor fly season was experienced in 2011, but 4 of these susceptible ewes were again struck in 2012. No chemicals or preventative treatments apart from crutching, were applied. During 2010, these 12 ewes were regularly crutched outside the fly season, and their wool samples sealed in plastic bags and forwarded to Hanrob International Dog Academy Pty in Sydney. No wool grown from the previous shearing was used for training, and no fly struck wool or wool treated with insecticides collected during the sampling period, were used for training.

Sheep (Armidale flock). In 2012, the 10 most resistant and 10 most susceptible sheep from the breech strike flock in Armidale (Smith *et al.* 2009) were identified amongst the 2005 and 2006 born ewes. The selected susceptible ewes were struck between 2 and 8 times, while none of the resistant ewes were struck over their lifetime prior to being selected up to sampling. However, 2 ewes deemed to be resistant prior to sampling have been subsequently struck. None of the ewes that were sampled for training purposes were struck during the sampling period. The ewes were crutched and the crutched wool samples forwarded to Hanrob International Dog Academy for testing.

Dogs. Three dogs were initially sourced by Hanrob Dog Academy and trained by a qualified trainer over 6 months to identify the crutched wool samples of the resistant ewes as target group. One dog was later excluded from the program as he did not make sufficient progress.

Training methodology. Eighteen tins were used as search items per exercise. Five tins were allocated to the resistant group and another 5 tins to the susceptible group. The remaining tins were allocated to the various items that contributed background odour such as gloves of the operator who sampled the sheep, scissors, unused empty plastic bags used to send and stored wool samples, and wool from unknown sheep. A wool sample was inserted in a clean 4 litre tin and covered with its lid which had approximately 10 pencil sized holes in, through which the odour from the item in the tin could escape. The wool from each group was kept separate for training. A separate pair of tongs was used to insert wool from the resistant or susceptible groups in the tins allocated to the different groups to prevent any cross contamination.

The dogs were trained by a senior qualified trainer using a positive reward system to identify the crutched wool from the resistant ewes as target group. The dogs were trained to sit when they

smelt the target (resistant) group. The samples were randomly replaced by new wool samples from both the target and susceptible groups and the tins were regularly shuffled to prevent any pattern formation. Fresh crutched wool samples were collected every 5-6 weeks on the 2 groups of ewes from the Mt Barker flock, sealed and forwarded to the training centre. Between zero or 5 target (resistant) or non-target (susceptible) items were included amongst 18 tins. The remainder of tins contained a variety of other items as mentioned above.

The dogs were tested over 4 occasions (27th July, 2nd, 9th and 15th August 2012) for their ability to correctly identify the target (resistant) group or to ignore the non-target (susceptible) items from both flocks. On 27th July 2 samples, on 2nd August again 2 samples, and on 9th August 1 sample from the Armidale flock were included amongst the Mt Barker samples, respectively. Both resistant (target) and susceptible (non-target) samples were used. An exercise test consisted of 18 tins which contained between zero or 5 target or non-target items at any time. The dogs were walked past the tins twice in opposite directions and scored for their ability to accurately identify the target and to ignore the non-target items. Between 2 and 5 exercises were carried out per test.

The final evaluation test on 15th August involved only crutched wool from the resistant and susceptible ewes from the Armidale flock to which the dogs have not been exposed previously. Five tests were carried out and the samples were varied and mixed with the other items to determine whether the dogs can differentiate between resistant and susceptible wool samples from sheep in this flock.

Statistical analysis. The number of successes and failures by correctly identifying the targets and ignoring the non-targets were determined. A Chi-square test was carried out to determine whether the proportions obtained were significantly higher than 75%. This value is the critical limit for novice dogs to be considered sufficiently accurate for the sniffer dog industry.

RESULTS AND DISCUSSION

Prior to the final test on the Armidale samples only, the dogs were tested on 27th July, 2nd and 9th August and six exercises were carried out over these days. Both dogs were 100% accurate in identifying the resistant and 100% accurate in ignoring the susceptible wool samples from the Mt Barker ewes on which they have been trained. Both dogs were also 100% accurate in identifying the resistant samples from Armidale when tested on 27th July and 2nd August. However, on 9th August both dogs failed to identify the same resistant sample from Armidale.

The final test was carried out on 15th August and the outcome from the test using the wool samples from resistant and susceptible sheep for the Armidale flock is shown in Table 1. Exactly the same number of successes and failures were recorded for both dogs. They were 82% accurate in identifying the resistant samples correctly, and were 92% accurate in ignoring susceptible samples. This is significantly ($P<0.05$) higher than the expected 75% accuracy for novice dogs.

Table 1. Number of successes and failures of dogs to correctly identify crutched wool samples sourced from genetically resistant and susceptible ewes to breech strike.

	Exercise number						
	1	**2**	**3**	**4**	**5**	**Total**	**Accuracy**
Targets (Resistant)	2	1	3	0	5	11	
Success (identify targets)	2	0	3	0	4	9	0.82
Failures (ignoring targets)	0	1	0	0	1	2	(P<0.05)
Non targets (Susceptible)	5	5	5	5	5	25	
Success (ignore non targets)	5	3	5	5	5	23	0.92
Failures (identify non-targets)	0	2	0	0	0	2	(P<0.05)
Other items	11	12	10	13	8	54	
Total number of search items	18	18	18	18	18	90	

CONCLUSIONS

The results showed that the two dogs were able to differentiate highly successfully between the crutched wool samples from the resistant and susceptible from the Mt Barker flocks on which they have been trained. The final test also showed that they were able to differentiate (P<0.05) between the resistant and susceptible sheep from the Armidale flock. This indicates that resistant and susceptible ewes to breech strike may have a common odour associated with their level of resistance or susceptibility. Logan *et al.* (2009) and Oyarzun *et al.* (2009) have identified that volatile chemicals play a part in attracting or repelling midges and flies, respectively. Thus different semio-chemicals may be operating as repellent and as attractants in resistant and susceptible sheep. This warrants further studies in identifying specific odour compounds that are common to resistant and susceptible sheep.

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

Ashworth J.R. and Wall R. (1994) *Med. Vet. Entomol.* **8**:303.

Greeff J.C. and Karlsson L.J.E. (2009) *Proc. Assoc. Advmt. Anim. Breed. Genet* 18:272.

Greeff J.C., Karlsson L.J.E. and Schlink A.C. (2013) *Anim. Prod. Sci.* http://dx.doi.org/10.1071/AN12233.

Logan J.G., Seal N.J., Cook J.I., Stanczyk N.M., Birkett M.A., Clark S.J., Gezan S.A., Wadhams L.J., Pickett J.A. and Mordue A.J. (2009) *J. Med. Entomol.* **46**:208.

Oyarzun M.P., Palma R., Alberti E., Hormazabal E., Pardo F., Birkett M.A. and Quiroz A. (2009) *J. Med. Entomol.* **46**:1320.

Seddon H.R., Belschner H.G. and Mulhearn C.R. (1931) NSW Dept. Agric. Sci. Bull. No.37.

Smith J., Brewer H. and Dyall T. (2009) *Proc. Assoc. Advmt. Anim. Breed. Genet.* **18**:334.

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:401-404

THE ROLE OF AI IN GENETIC PROGRESS - NEW OPPORTUNITIES FROM NEW TECHNOLOGIES AND NEW APPROACHES

R.E. Green, P.R. Amer and P.F. Fennessy

AbacusBio Limited, PO Box 5585, Dunedin 9058, New Zealand

SUMMARY

Currently, the use of artificial reproductive technologies (ART) in the New Zealand sheep and beef industries is limited. While the past 30 years have seen rapid development of new reproductive technologies and improvements in existing technologies, there are issues in practical implementation. A simple tool like Artificial Insemination (AI) would greatly facilitate and enhance genetic improvement programmes for sheep and beef cattle. Effective application of ART provides an opportunity to increase the rate of dissemination of superior animals and the rate of genetic gain, but there is little evidence to support their practical and economic value in the NZ sheep and beef industries. Complementary technologies such as monitoring the reproductive cycle, estrous synchronization, and semen sexing can improve the efficiency of AI. This paper reviews on-going and recent technological advances that have the potential to significantly improve the genetic merit of sheep flocks and beef herds when implemented as part of ART program.

INTRODUCTION

In the last few decades, a significant improvement in the efficiency of sheep and beef production has been achieved due to implementation of several new or improved technologies and production practices. Major advances in genetics and genomic applications can be expected to further accelerate genetic gain. For example, the rate of genetic gain in the New Zealand sheep flock has increased with the introduction of the Central Progeny Test and SILACE (Amer 2009).

However, there is considerable potential to enhance the rate of dissemination of superior genetic material among breeders, multipliers and commercial flocks and herds, given the availability of a practical and economically-feasible reproductive technology. Artificial insemination (AI) is the technique of choice for widespread dissemination of desirable genetics in farmed livestock.

Currently, the NZ dairy cattle industry is a major user of AI, applying this technology in around 70% of dairy cows (NZ Dairy Statistics, 2011-2012). However very few beef and sheep breeders make use of the technology and application is limited to some breeders producing bulls and rams for sale. The only reproductive technology in widespread use is ultrasound pregnancy scanning. While breeders recognise the potential to enhance the rate of genetic improvement through the use of ART, it seems that both breeders and commercial producers need to be familiar with the opportunities presented by these methods. Artificial insemination, estrous synchronization, embryo transfer, *in vitro* fertilization, and semen sexing are all procedures that have influenced, or can be expected to have a major influence in the beef and sheep industries. The major reason for the slow adoption of technologies such as AI is their relatively high cost in relation to both the risks and the short-term benefits.

In order to assess the opportunities to utilise AI more widely in the NZ sheep industry, it is necessary to define the conditions under which the technology can be expected to operate. The most likely applications will involve AI within the normal breeding season, although there may be some options out of season where problems associated with seasonality can be expected. Aspects to consider include estrous synchronization and the use of frozen, chilled or sexed semen. Therefore this paper highlights some practical aspects of AI in sheep and beef cattle, with an emphasis on the benefits in terms of genetic progress.

ESTROUS SYNCHRONIZATION

Beef cattle. The synchronization of estrous facilitates the application of AI and it may also facilitate better feeding and calving management as all cows will be at the same stage of pregnancy. The extensive nature of cattle production systems makes the use of AI a more challenging option than in the intensively-managed dairy industry (Hall 2011). Ideally, estrous synchronization should be cost-effective, simple and practical to implement, with minimal animal handling and without the need to detect the females on heat (Busch *et al.* 2008). Success with such practices could result in a more highly synchronized and fertile estrous with excellent pregnancy rates from fixed-time AI (FTAI). Currently, several types of GnRH/PGF2α and progesterone/progestin-based estrous synchronization protocols have been developed for use in cattle, allowing farmers to perform timed AI (TAI) without detection of estrous 48 hours after progesterone withdrawal. However there is a need for further development to effectively utilise sex-sorted sperm in association with TAI protocols (Sales *et al.* 2011).

Sheep. Estrous synchronization is fundamental to the application of most ARTs. The most common protocol for sheep estrous synchronization is based on intravaginal devices with progesterone or progestagens for 12-14 days. They can also be combined with equine chorionic gonadotrophin (eCG) to increase ovulatory efficiency and ovulation rate (Letelier *et al.* 2011); this can improve results from FTAI (without estrous detection). Fixed timed AI is usually performed at 48 to 60 hours after progesterone withdrawal, depending on the type of semen (fresh or frozen) and technique used (cervical or laparoscopic). It seems that the time of ovulation is critical for the use of cryopreserved semen in TAI programs. Several studies have sought to improve the time of ovulation especially with the use of sexed semen (de Graaf *et al.* 2007a, 2007b; Beilby *et al.* 2009). Increasing the precision of the time of ovulation using GnRH 36 hours after progesterone/progestagen withdrawal can play an important role in obtaining satisfactory fertility (Hollinshead *et al.* 2002).

ARTIFICIAL INSEMINATION WITH LIQUID STORED SEMEN

The storage of semen can be achieved through methods that reduce the metabolic rate of spermatozoa, thereby prolonging their fertile life. Liquid storage of semen is carried out using temperatures low enough to depress sperm metabolism (5 or 15°C) (Anel *et al.* 2006). Fresh-cooled (15°C) or chilled (5°C) semen is a good alternative to frozen semen, when it is used within a short period after collection. In New Zealand, chilled semen remains the method of choice in dairy cattle, mainly because of restricted seasonal breeding (early September to December) and the development of technology which allows the use of low numbers of sperm, with excellent longevity and viability of sperm during storage and post-insemination (Vishwanath *et al.* 1996).

On the other hand, for sheep, much higher numbers of spermatozoa are required for effective transcervical AI using chilled semen. However lower doses of semen can be used for laparoscopic insemination. An important limitation for the use of chilled semen, especially in sheep, is the logistics of transporting semen given the maximum shelf life of 24 hours with acceptable fertility (Salamon and Maxwell 2000).

ARTIFICIAL INSEMINATION WITH FROZEN SEMEN

The goal of a good sperm freezing protocol is the production of a bank of sperm for AI. However, various biochemical and anatomical compartments in the sperm cells may be altered during freezing and thawing (Amirat *et al.* 2004). Consequently, the fertility is normally lower than that achieved for fresh semen. The most important advantage of frozen-thawed semen is the long-term storage capability. This allows for extensive testing of the processed semen, and is a reliable method for genetic insurance of valuable bulls (Vishwanath and Shannon 2000) and rams.

Due to practicability and a general consistency of results, users in NZ generally have a preference for this type of semen administered by laparoscopic intrauterine insemination. Laparoscopic artificial insemination (LapAI) with frozen semen has been used in the Beef + Lamb NZ Central Progeny Test since 2002. This provides genetic connectedness across years and across the three CPT flocks. In a summary of data for the first six years, McLean *et al.* (2008) reported that the highest conception rate was 83% in 2006.

ARTIFICIAL INSEMINATION WITH SEXED SEMEN

Sex-pre-selection is a potentially attractive technology that could increase uptake of AI in New Zealand. Generally protocols are well-established for sex-sorting of livestock spermatozoa using the Beltsville sperm sexing technology and offspring have been produced in several species by the combination of flow cytometric sperm sorting with a range of ART methods. As with most ART, the practical use of sexed spermatozoa depends on the cost, the fertility results, efficiency and ease of use as assessed against the benefits in terms of genetic gain or the production of a higher proportion of the more valuable sex of offspring.

Furthermore, another limitation from a commercial perspective is the minimum effective dose of sexed semen and the cost of production of the semen. In general, fertility results for beef cattle seem to be lower than with unsorted semen (Seidel, 2011). However the sheep data are relatively good; perhaps surprisingly, low-dose AI of sex-sorted ram sperm has produced similar, if not superior, fertility to non-sorted controls (de Graaf *et al* 2007a, 2007b). A key factor in good results is the time of insemination, which should be close to the time of ovulation. Therefore hormonal protocols to control ovulation, combined with the use of low sperm numbers, provides a highly encouraging outlook for the commercial application of sex-sorted, frozen-thawed ram (Beilby *et al.* 2009) and bull sperm.

Commercial use of sex-sorted sperm depends on its price. It has to be low enough to allow a reasonable profit for farmers. In New Zealand, sexed semen is now being used in dairy cattle, but the cost and possibly slightly lower conception rates are apparently regarded as a deterrent. There is very limited use or beef cattle and there are no indications of any use at all in sheep flocks. However, this technology could be used by breeders and multipliers, allowing more selective production of males and female for replacement and sale.

POTENTIAL BENEFITS OF AI IN GENETIC IMPROVEMENT

The application of ART, such as AI, has had a major impact on the structure of breeding programs, the rate of genetic gain and the dissemination of genetic gain in livestock production (Van Arendonk *et al.* 2011). The industry impact of genetic improvement in the breeder sector is absolutely dependent on the effective dissemination of genetic material from the breeders to the target population (commercial farms). In this situation, AI is very important for effective dissemination.

The major benefit for the NZ beef and sheep industries from the increased adoption of AI is the increase of genetic gain rates (ΔG). To realize this goal it is essential to:

- improve flock (genetic) connectedness across breeder flocks and also with multiplier flocks, in order to generate better estimated breeding values and so better identify superior individuals; and to
- develop better methods to utilize such superior individuals.

Both aspects will be facilitated by increased use of AI (and potentially other ART) in breeding flocks/herds, and in their associated multiplier flocks/herds. In essence, such practices will increase the intensity of selection, which in turn can result in an increase in the average genetic merit of offspring as outlined by Nicholas (1996).

CONCLUSION

Faster and more widespread dissemination of genetic improvement from proven sires in sheep and beef cattle through AI is limited due to a number of factors. These include animal handling and the labour requirement. In cattle, the need for estrous detection or reliable synchronization is a critical issue. In sheep, the insemination technique and consistent conception rates for frozen semen are important aspects to be considered if more widespread of use of AI is to be achieved.

There has been considerable investment in New Zealand in the development and application of genetic technologies. The use of AI and other reproductive technologies has the potential to increase the return on this investment through a greater rate of genetic response, which is also dependent on breeding scheme structure. In this respect, data on the current rate of use of AI and the situations where it is used would be useful to provide a solid base to development of the value proposition for the use of AI (and potentially other artificial breeding technologies). Greater use of AI could be achieved through commitment from organizations in the breeding sector, especially around how to reduce barriers for adoption, together with a clear value proposition of the benefits from genetic improvement and its ultimate impact on farm profitability.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge financial support for this study from Ovita Ltd.

REFERENCES

Amer, P.R. (2009) *Proc. of the NZ Soc. of Anim. Prod*. **69**:155.

Amirata L., Tainturiera D., Jeanneaua L., Thorina C., Ge´rardb O., Courtensc J.L. and Anton M. (2004) *Theriogenology* **61:** 895.

Anel L., Alvarez M., Martinez-Pastor F., Garcia-Macias V., Anel E. and de Paz P. (2006*) Reprod Dom Anim* **41**: 30.

Beilby K.H., Grupen C.G., Thomson P.C., Maxwell W.M. and Evans G. (2009) *Theriogenology* **71**:829.

Busch D.C., Schafer D.J., Wilson D.J., Mallory D.A., Leitman N.R., Haden J.K., Ellersieck M.R., Smith M.F. and Patterson D.J. (2008) *J. Anim. Sci*. **86**: 1519.

de Graaf S.P., Evans G., Maxwell W.M., Downing J.A. and O"Brien J.K. (2007a) *Reprod Domest Anim* **42**: 648.

de Graaf, S.P., Evans, G., Maxwell, W.M.C., Cran, D.G. and O"Brien, J.K. (2007b) *Theriogenology* **67:** 391.

Gordon I. (1999). Reproduction in Sheep and Goats. CAB International, UK.

Hall J. B. (2011) *Proc, Appl Reprod Strat in Beef Cattle* - Northwest September 30 – October 1.

Hollinshead F.K., O'Brien J.K., Maxwell W.M.C. and Evans G. (2002) *Reprod Fertil and Develop* **14**: 503.

Letelier C.A., Contreras-Solisa I., García-Fernándezc *et al* (2011) *Anim Reprod Sci* **126:** 61

McLean N.J., Dodds K.G., Knowler K.J., Jopson N.B., Behrent M., Logan C.M., Muir P.D., Shackell G.H. and McEwan J.C. (2008) *Proc. of the NZ Soc. of Anim. Prod*. **68**: 77.

Nicholas F.W. (1996) *Anim Reprod Sci* **42**: 205.

Rath D. and Johnson L.A. (2008) *Reprod. Domest. Anim.* **43:** 338.

Salamon S., Maxwell W.M.C. (2000) *Anim Reprod Sci* **62**: 77.

Sales J.N.S., Neves K.A.L., Souza A.H., Crepaldi G.A., Sala R.V.,Fosado M., Filho E.P.C. and Baruselli P.S. (2011) *Theriogenology* **76**: 427.

Seidel G.E. Jr. (2011) *Proc. App. Reprod. Strat. In Beef Cattle*, 349.

Van Arendonk, J.A.M. (2011) *Livestock Sci* **136**: 29.

Vishwanath R., Pitt C.J. and Shannon P. (1996). *Proc. of the NZ Soc. of An. Prod*. **56**:31.

Vishwanath R., Shannon P. (2000) *Anim Reprod Sci* **62**: 23.

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:405-407

IN-SILICO APPROACH IDENTIFIED POLYMORPHISM ASSOCIATED WITH WOOL TRAITS IN SHEEP

E. Jonas[1,2], P.C. Thomson[1] and H.W. Raadsma[1]

[1] ReproGen, Faculty of Veterinary Science, The University of Sydney, Camden, NSW, 2570
[2] Swedish University of Agricultural Sciences – Department of Animal Breeding and Genetics, 750 07 Uppsala, Sweden

SUMMARY

A number of factors contribute to the economic revenue of sheep farmers, but breed improvement via selection is one factor affecting the efficiency of sheep production particularly in relation to wool. The complexity of traits, with partly opposing relationships makes it difficult to select for improved economic values of wool. A further investigation of the genetic background of wool quality, quantity as well as pigmentation traits might assist to unravel the basis of this relationship. Our approach was to identify and analyse possible candidate genes in major linkage regions. Using a combination of positional mapping and literature finding of the gene function, we identified lysosomal trafficking regulator (LYST) as strong candidate gene. Polymorphisms were identified using in-silico screening of scaffolds on the virtual sheep genome assembly v2.0. One of four polymorphisms analysed in the ovine LYST gene was significantly associated with clean fleece weight and coefficient of variation of fleece diameter. However, this polymorphism explained only a small proportion of the phenotypic variation, contradicting unpublished findings of major effects of QTL on chromosome 25 for the same traits. Further analysis is needed to analyse the function of LYST and to find additional genes either having a direct effect on wool quality and quantity or regulating the function of the ovine LYST gene.

INTRODUCTION

Largely as a consequence of the antagonistic relationship between two of the major determinants of profit (fleece weight and fibre diameter), the rate of genetic improvement of sheep bred primarily for wool production has been relatively slow (Purvis and Franklin 2005). A better knowledge on the genetic background of wool quantity and quality traits might assist increasing the gain from the farming of wool-sheep by understanding some of the basis for this opposing relationship. A number of linkage studies have identified major loci for wool quality and quantity traits in sheep. Many QTL have been reported for wool quality parameters such as clean or greasy fleece weight and yield and for wool quality including fibre diameter, staple length, coefficient of variation and standard variation of fibre diameter across many chromosomes (Parsons et al. 1994; Allain et al. 1998; Beh et al. 2001; Ponz et al. 2001; Allain et al. 2006; Bidinost et al. 2006; Bidinost et al. 2008). Preliminary studies using a resource population have also verified highly significant QTL for wool quality and quantity on chromosome 25 and a meta-assembly of QTL in sheep has suggested major genes on chromosomes 3 and 25 (Raadsma et al. unpublished).

The study presented here aimed to identify polymorphisms in the main candidate region on chromosome 25 using published sequence information and to analyse the association with wool quality and quantity traits in an ovine sheep resource population.

MATERIALS AND METHODS

Animals. For the analysis presented here, a total 170 wether backcross progeny from a resource population of crosses between Awassi and Merino sheep (Raadsma et al. 2009) were used. A total of five wool quality and three quantity traits were recorded from wool samples collected at 75

weeks of age. Measurement of most fibre characteristics were performed by Riverina Wool Testers in Wagga Wagga, Australia (http://www.wooltesters.com.au/).

Genotyping. In a previous conducted linkage study using backcross Awassi x Merino x Merino (AMM) sheep, QTL were identified for wool traits across different chromosomes using the program QTL-MLE (Raadsma et al. unpublished). Highly significant QTL were identified on chromosome 25, where many QTL were located around 15 cM with a 1-LOD drop off confidence interval of 0-32 cM. The region around microsatellite marker DIK2451, where most of the QTL were identified (Raadsma et al. unpublished) was used to identify the most likely location of an underlying candidate gene. Genes on the virtual sheep genome assembly v2.0 (http://www.livestockgenomics.csiro.au/perl/gbrowse.cgi/vsheep2/) were further taken into consideration if their described function suggested a connection to wool development. Scaffolds from the ovine whole-genome sequence on the virtual sheep genome assembly v2.0 were screened for potential SNPs in the candidate gene. A total of four SNPs identified in the region were genotyped using the iPLEX system (Sequenom).

Association analysis. Association was tested using analysis of variance in R (version 2.15.1) where the SNPs were fitted as fixed effects. A number of traits were investigated including greasy fleece weight, clean fleece weight (CFW), fleece yield, fibre diameter, standard deviation of fibre diameter, coefficient of variation (CV), percentage fibres greater than 30 µm and fleece rot. No additional effects were included in the models as all animals were from the same resource flock and kept under the same conditions. Animals were from the same sire.

RESULTS AND DISCUSSION

Previously a number of QTL were identified for wool quality and quantity in the AMM resource population (Raadsma et al., unpublished). In particular, the many QTL on chromosome 25 were located within the same marker intervals. QTL on this chromosome were previously published in a Sarda x Lacaune backcross population (Allain et al. 2006), Merino sheep (Bidinost et al. 2006; Bidinost et al. 2008), and animals from the synthetic INRA401 breed (Ponz et al. 2001). Genome-wide-association studies using data from the AMM animals validated these results and the regions were further fine-mapped (data not shown). The in-silico analysis of the candidate region on chromosome 25 suggested lysosomal trafficking regulator (LYST) as a potential candidate gene. A total of four SNP were reported on the published sequence including a non-synonymous (G/A) polymorphisms in exon 29, two non-synonymous SNP (T/C and A/G) in exon 36 and one synonymous SNP (G/A) in exon 20. All four SNP segregated in the genotyped animals from the Awassi Merino population.

LYST has been previously been related to melanosome formation, a process positioned between melanocyte development and pigment production in the development of pigmentation cells. Association analysis showed that the non-synonymous SNP in exon 29 of the ovine LYST gene was significantly associated ($P < 0.01$) with CFW and CV. However, none of the other polymorphisms were significantly associated with the investigated wool traits. Additionally the explained variation of the phenotype was rather small using these models. The unpublished study of QTL for wool traits using animals from the same population suggested a major gene especially for wool quality traits, explaining up to 70% of the phenotypic variation for CV for example (Raadsma et al. unpublished). Despite our findings suggesting some effect of the polymorphisms within the ovine LYST gene, we assume that the significant polymorphism in our study is unlikely the major gene underlying the strong QTL identified earlier (Raadsma et al. unpublished). The in-silico approach taken for this study was surprisingly successful as all four polymorphisms identified in the published ovine sequence segregated in our population. We conclude that it is a useful approach screening such data as useful source for preliminary studies. However, further

screening of other genes or genes regulating the function of LYST might be useful to continue future efforts to identify loci underlying the major QTL for wool quality and quantity traits.

ACKNOWLEDGEMENT

We acknowledge the support received from all staff involved in different stages of the project.

REFERENCES

Allain D., Lantier I., Elsen J.M., Francois D., Brunel J.C., Weisbecker J.L., Schibler L., Vaiman D., Cribiu E., Gautier A, Berthon P. and Lantier F. (1998) *Proc. 6th World Congr. Genet. Appl. Livest. Prod., Armidale, Australia*. **24**: 51.

Allain D., Schibler L., Mura L., Barillet F., Sechi T., Rupp R., Casu S., Cribiu E. and Carta A. (2006*) Proc. 8th World Congr. Genet. Appl. Livest. Prod., Belo Horizonte, Minas Gerais, Brazil*

Beh K.J., Callaghan M.J., Leish Z., Hulme D.J., Lenane I. and Maddox J.F. (2001) *Wool Techn. Sheep Breed.* **49**: 88.

Bidinost F., Roldan D.L, Dodero A.B., Cano E.M., Taddeo H.R., Mueller J.P. and Poli M.A. (2008*) Small Rum. Res.* **74**: 113.

Bidinost F., Roldan D.L., Dodero A.M., Cano E.M., Taddeo H.R., Mueller J.P. and Poli M.A. (2006) *Proc. 8th World Congr. Genet. Appl. Livest. Prod., Belo Horizonte, Minas Gerais, Brazil*

Parsons Y.M., Cooper D.W. and Piper L.R. (1994) *Anim. Genet.* **25**: 105.

Ponz R., Moreno C., Allain D., Elsen J.M., Lantier F., Lantier I., Brunel J.C. and Perez-Enciso M. (2001) *Mamm. Genome* **12**: 569.

Purvis I.W. and Franklin I.R. (2005) *Genet. Sel. Evol.* **37**: S97-S107.

Raadsma H.W., Jonas E., Fleet M.R., Fullard K., Gongora J., Cavanagh C.R., Tammen I., and Thomson P.C. (2013) *Anim. Genet.*, accepted.

Raadsma H.W., Jonas E. and Thomson P.C. In preparation.

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) *Genet. Select. Evol.* **41**: 34.

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:408-411

FIBRE DIAMETER CORRECTED WOOL CLEAN COLOUR - THE IMPACT ON GENETIC PARAMETERS

A.Jones[1] and S.Hatcher[1,2]

[1]NSW Department of Primary Industries, Orange Agricultural Institute, Locked Bag 6006 Forest Road Orange NSW 2800, Australia
[2]CRC for Sheep Industry Innovation, Homestead Building, University of New England, Armidale NSW 2350, Australia
sue.hatcher@dpi.nsw.gov.au

SUMMARY

The brightness (Y) and yellowness (Y-Z) of wool is highly correlated with mean fibre diameter (MFD). The Cooperative Research Centre for Sheep Industry Innovation (Sheep CRC) and the Australian Wool testing Authority Limited (AWTA Limited) have recently developed an algorithm to correct Y and Y-Z for MFD which significantly reduces the fibre diameter covariance with colour. This paper demonstrates that correcting Y, Z and Y-Z for fibre diameter had little to no impact on the phenotypic or genetic relationships between the various wool colour traits. Therefore removing the diameter covariance in Y, Z or Y-Z using fibre diameter correction algorithm will not dramatically change the phenotypic and genotypic correlations between colour and other key wool production and quality traits. This analysis indicates it is not necessary to correct wool colour for fibre diameter when colour is included in Merino breeding programs.

INTRODUCTION

A key objective of the Sheep CRC's Wool Program is to provide Australian sheep breeders with the genetic information required to improve the clean wool whiteness and brightness of the Australian wool clip. The yellowness (i.e. Y-Z) of clean wool is highly correlated with the mean fibre diameter (MFD) of wool fibres (Fleet *et al*. 2009; Hebart and Brien 2009; Smith and Purvis 2009). This has the potential to confound genetic parameter estimates for colour traits; significantly impairing the ability of the Australian wool industry to improve wool colour using genetic selection (Wang *et al*. 2011). The Sheep CRC and AWTA Limited developed a mathematical diameter-scatter correction algorithm for clean colour measurement using selected midside samples from the Information Nucleus Flock (INF). The algorithm was subsequently validated on all INF samples measured during 2008-2009 which proved it was possible to remove the fibre diameter artefact from the phenotypic relationship between clean colour and fibre diameter (Wang *et al*. 2011). The objective of this paper was to investigate the impact of the diameter-scatter correction on the genetic parameter estimates for brightness (Y) and Y-Z and their phenotypic and genetic correlations with wool production and other wool quality traits. The likely consequences for Merino breeding objectives aimed at producing whiter wool are discussed.

MATERIALS AND METHODS

Data from the Sheep CRC's INF (van der Werf *et al*. 2010) Merino progeny run at each of the 8 IN sites measured as yearlings (approx. 11months, n = 4,019) and adults (approx. 23 months, n = 2,012) were used for this analysis. The sheep were born between 2007 and 2009 and a comprehensive suite of wool production and quality traits (Hatcher *et al.* 2010) were assessed or measured at each shearing (2008-2010 for yearlings and 2009-2010 for adults). Midside samples (75–85 g) taken from the right side of each animal prior to shearing and measured at AWTA Limited using standard IWTO test methods as described in Hatcher *et al.* (2010). Briefly 10 staples from each midside sample were sub sampled for measurement of staple length (SL) and staple strength (SS). The remainder of each sample was weighed, washed, oven dried (105°C),

carded and conditioned for 24 hours prior to weighing to determine the washing yield (using 16% regain). A mini-corer was then used to sample 2mm fibre snippets which were measured for mean fibre diameter (FD), FD standard deviation (FDSD), FD coefficient of variation (FDCV) and mean fibre curvature (CURVE) using Sirolan™ Laserscan. The carded sample was further sub-sampled and measured for clean colour (Y–Z, yellowness) and the X, Y and Z tristimulus vales (X, red; Y, green brightness and; Z, blue). Clean colour (Y-Z) and the Y and Z tristimulus values were then adjusted for fibre diameter using the correction algorithm described by Wang *et al*. 011); yielding the additional traits CYY, CYZ and CYY-Z and CAY, CAZ and CAY-Z for the yearling (Y) and adult (A) stages respectively where 'C' denotes corrected for fibre diameter. The greasy fleece weight (GFW) of each unskirted fleece (belly wool included) was recorded at shearing with clean fleece weight (CFW) calculated as the product of GFW and the washing yield. GFW and CFW were corrected to 365-day growth equivalents (Hatcher *et al*. 2010).

ASReml 3.0 (Gilmour *et al*. 2009) was used to estimate fixed effects, variance components and genetic parameters using general linear mixed models and the residual maximum likelihood method as described by Hatcher *et al.* (2010). Phenotypic and genetic covariances were estimated from a series of bivariate analyses using fixed effects and their interactions, as appropriate from the univariate analyses. Genetic and phenotypic correlations, and their standard errors, were estimated from the appropriate variances and covariances using ASReml.

RESULTS AND DISCUSSION

There was no significant difference in heritability between any of the yearling and adult corrected and uncorrected colour traits (Table 1). Similarly the phenotypic variance, coefficient of variation, residual, additive variance and sire.flock variance were similar for the corrected and uncorrected yearling and adult colour traits. A maternal variance term was fitted in all models but was not significant for either the corrected or uncorrected colour traits.

Table 1: Mean, variance components, coefficient of variation and heritability for yearling and adult uncorrected and fibre diameter corrected colour traits

Trait	Mean (*tristimulus values*)	Phenotypic variance	Coefficient of variation (%)	Residual variance	Additive variance	Sire.flock variance	Heritability
Yearling							
YY	67.63	2.70	2.22	2.07	0.59	0.05	0.22 ± 0.04
YZ	62.68	4.19	3.11	2.31	1.72	0.16	0.41 ± 0.05
YY-Z	4.95	0.74	10.48	0.22	0.47	0.05	0.63 ± 0.05
CYY	73.90	2.71	2.22	2.08	0.59	0.04	0.22 ± 0.04
CYZ	65.62	4.13	3.08	2.34	1.66	0.13	0.40 ± 0.05
CYY-Z	8.28	0.69	10.18	0.24	0.41	0.05	0.59 ± 0.05
Adult							
AY	74.14	2.46	2.11	1.62	0.74	0.10	0.30 ± 0.07
AZ	65.81	3.39	2.79	2.11	1.13	0.15	0.33 ± 0.07
AY-Z	8.32	0.40	7.64	0.25	0.14	0.01	0.35 ± 0.07
CAY	74.15	2.46	2.11	1.62	0.74	0.10	0.30 ± 0.07
CAZ	66.01	3.37	2.78	2.12	1.10	0.15	0.33 ± 0.07
CAY-Z	8.14	0.35	7.31	0.25	0.09	0.01	0.27 ± 0.07

The means for the yearling corrected colour traits were higher than the uncorrected traits (+6.3, +2.9 and +3.3 for YY, YZ and YY-Z respectively), but smaller differences were evident for the adult colour traits (+0.01, +0.20 and -0.18 for AY, AZ and AY-Z respectively). These differences

could be due to a 1.2µm difference between the yearling and adult FD (YFD = 16.9 µm AFD = 18.1µm) in this study as higher diameter wool tends to be more yellow (Wang *et al*. 2011). The heritability estimate for YY-Z was significantly higher than previous reports (Fleet *et al*. 2009; Hebart and Brien 2009; Smith and Purvis 2009) which may be due to the differing average FD and low FD spread of the flocks in those studies compared to the 14 µm FD range in the INF. The heritability of AY-Z in the INF agrees with the adult estimate of Smith and Purvis (2009).

Correcting Y, Z and Y-Z for fibre diameter had little to no impact on the phenotypic (r_p) or genetic (r_g) relationships between the various colour traits (Table 2). Each of the correlations, both r_p and r_g, between the uncorrected and corrected colour traits were all high (i.e. >0.6), except for the r_p between CY and Y-Z and between CY-Z and Y which were low (both -0.26). These low r_p are the result of the relatively lower heritability of Y compared to Z (Table 1), which indicates that reflectance of green light from wool fibres (i.e. Y) is more affected by the environment than the animals genes compared to the reflectance of blue (i.e. Z) leading to the lower r_p. The same trends occurred between the corrected and uncorrected adult measurements of these colour traits. The generally high r_p and r_g between the corrected and uncorrected colour traits indicates that removing the diameter co-variance does not change the phenotypic or genetic relationships between the colour traits whether measured either as yearlings or adults.

Table 2- Phenotypic (r_p) and genetic (r_g) correlations between uncorrected and fibre diameter corrected colour traits

FD corrected colour traits	Uncorrected colour traits					
	YY		YZ		YY-Z	
	r_p	r_g	r_p	r_g	r_p	r_g
CYY	0.99 ± 0.00	0.99 ± 0.00	0.92 ± 0.00	0.92 ± 0.02	-0.26 ± 0.02	-0.65 ± 0.07
CYZ	0.91 ± 0.00	0.91 ± 0.02	0.99 ± 0.00	0.99 ± 0.00	-0.63 ± 0.01	-0.89 ± 0.02
CYY-Z	-0.26 ± 0.02	-0.64 ± 0.07	-0.62 ± 0.01	-0.86 ± 0.03	0.98 ± 0.00	0.98 ± 0.00

Correcting Y, Z or Y-Z for fibre diameter had very little impact on the r_p or r_g between the various colour and key wool production and quality traits (Table 3). The r_p between the uncorrected colour traits and key wool production and quality traits were all negligible in magnitude (i.e. <0.2) except for between YFD and corrected Y-Z (0.28) was lowly correlated. The r_p between the corrected colour traits and key wool production and quality traits were all negligible. Similarly the majority of the r_g were also negligible except for those between Y and CY with YSS (0.34 and 0.33 respectively), Z and YFD (0.29), Y-Z and CY-Z with YCFW (0.41 and 0.37), Y-Z and YFD (0.39) and Y-Z and YSL (0.22). The r_p and r_g between YFD and YY-Z are similar to those of Smith and Purvis (2009) but lower than those of Hebart and Brien (2009).

There was no major difference in either the r_p or r_g between YY and CYY and any of the key wool production and quality traits (Table 3). Similarly there was no difference between the correlations with YZ or CYZ, except for the r_p and r_g with YFD, which were both lower with CYZ (-0.17 vs -0.08 and -0.29 vs -0.18 respectively). The same trend was evident for the r_p and r_g between YY-Z and YFD (0.28 vs -0.08 and 0.39 vs -0.18 respectively). The same trends were evident for the adult measurements of these traits. A possible reason for the change observed in the r_p and r_g with YFD could be due to using fibre diameter as the basis for correcting the colour traits. The FD covariances are therefore likely to be most impacted by the correction thus producing the greatest difference observed in the estimated correlations. Small differences in the r_p and r_g with YSL and YSS were also identified; however given the relative size of the standard error for each estimate these differences are unlikely to be important. The observed differences in the YSL and YSS covariance with colour when correcting for FD maybe related to the antagonistic phenotypic

and genetic relationships between FD, SL and SS (i.e. finer fleeces are associated with shorter weaker staples) (Safari *et al*. 2005).

Table 3. Phenotypic and genetic correlations between uncorrected and FD corrected colour traits and key wool production and quality traits measured as yearlings

	YY		YZ		YY – Z	
	Uncorrected	Corrected	Uncorrected	Corrected	Uncorrected	Corrected
Phenotypic correlations						
YCFW	0.10 ± 0.02	0.10 ± 0.02	0.01 ± 0.02	0.03 ± 0.02	0.16 ± 0.02	0.12 ± 0.02
YFD	-0.06 ± 0.02	-0.05 ± 0.02	-0.17 ± 0.02	-0.08 ± 0.02	0.28 ± 0.02	0.08 ± 0.02
YFDCV	-0.07 ± 0.02	-0.07 ± 0.02	-0.05 ± 0.02	-0.06 ± 0.02	-0.02 ± 0.02	0.01 ± 0.02
YSL	-0.01 ± 0.02	0.00 ± 0.02	-0.04 ± 0.02	-0.01 ± 0.02	0.08 ± 0.02	0.02 ± 0.02
YSS	0.12 ± 0.02	0.13 ± 0.02	0.08 ± 0.02	0.11 ± 0.02	0.05 ± 0.02	0.00 ± 0.02
Genetic correlations						
YCFW	0.09 ± 0.11	0.09 ± 0.11	-0.18 ± 0.09	-0.13 ± 0.09	0.41 ± 0.07	0.37 ± 0.07
YFD	-0.12 ± 0.09	-0.10 ± 0.09	-0.29 ± 0.07	-0.18 ± 0.07	0.39 ± 0.06	0.18 ± 0.06
YFDCV	-0.17 ± 0.10	-0.18 ± 0.10	-0.11 ± 0.08	-0.12 ± 0.09	-0.03 ± 0.07	0.00 ± 0.08
YSL	-0.15 ± 0.10	-0.15 ± 0.10	-0.19 ± 0.08	-0.16 ± 0.08	0.22 ± 0.07	0.16 ± 0.07
YSS	0.34 ± 0.10	0.33 ± 0.11	0.16 ± 0.09	0.20 ± 0.09	0.11 ± 0.08	0.04 ± 0.09

These differences between the corrected and uncorrected colour traits in either the phenotypic or genetic correlations are likely be of low practical relevance to Merino breeding programs as all the correlations are classified as low to negligible. Therefore removing the diameter covariance in Y, Z or Y-Z by using the correction algorithm will not dramatically change the phenotypic and genotypic correlations between colour and other key wool production and quality traits. This analysis indicates that it is not necessary to correct wool colour for FD when colour is included in Merino breeding programs.

Light reflectance in lower wavelength zone of the spectrum (i.e. < 470nm or the 'blue' light range represented by the Z tristimulus value) is more affected by light scattering behaviour due to the morphological structure of the scales on the surface of the fibre (Wang *et al*. 2011). This may explain the differences observed in the variance components as well as the heritability, phenotypic and genetic correlation estimates between the colour traits and warrants further investigation.

ACKNOWLEDGEMENTS

We acknowledge the dedicated team of scientists and technical staff at each of the 8 INF sites who are responsible for management of each of the flocks and implementation of the comprehensive sampling, assessment and measurement protocols.

REFERENCES

Fleet M.R., Millington K.R., Smith D.H. and Grimson R.J. (2009) *Assoc. Advmt. Anim. Breed. Genet.* **18**: 556

Gilmour A.R., Gogel B.J., Cullis B.R. and Thompson R. (2009) 'ASReml User Guide Release 3.0'. (VSN International Ltd: Hemel Hempstead, HP1 1ES, UK)

Hatcher S., Hynd P.I., Thornberry K.J. and Gabb S. (2010) *Anim. Prod. Sci.* **50**: 1089

Hebart M.L. and Brien F.D. (2009) *Assoc. Advmt. Anim. Breed. Genet.* **18**: 500

Safari E., Fogarty N.M. and Gilmour A.R. (2005) *Livest. Prod. Sci.* **92**: 271.

Smith J.L. and Purvis I.W. (2000) *Assoc. Advmt. Anim. Breed. Genet.* **18**: 390

van der Werf J.H.J., Kinghorn B.P., and Banks R.G. (2010) *Anim. Prod. Sci.* **50**: 998.

Wang H., Mahar T., Liu X., Swan P. and Wang X. (2011) *J. Text. Inst.* **102**: 1031.

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:412-415

PARTITIONING THE GENETIC VARIANCE INTO GENOMIC AND PEDIGREE COMPONENTS FOR PARASITE RESISTANCE IN SHEEP

M. Al Kalaldeh J.P. Gibson, J.H.J. van der Werf and C. Gondro

University of New England, School of Environmental and Rural Science,
Armidale 2351, Australia

SUMMARY

In this study, we estimated the additive genetic variance explained by genomic markers for parasite resistance in a large mixed population of sheep and compared this estimate to the additive genetic variance explained by pedigree. Furthermore, we partitioned the total genetic variance by fitting both of genomic relationship matrix (GRM) and numerator relationship matrix (NRM) simultaneously into a genomic component explained by genomic relationships and a polygenic component explained by pedigree relationships. In this analysis, all the genetic variation explained by pedigree could be captured by the 50K SNP chip markers. When both of GRM and NRM were fitted simultaneously, 73.7% of total genetic variance was explained by genomic effects while the remaining variance (26.3%) was explained by pedigree effects. The proportion of genetic variance explained by genomic effects was further partitioned into 26 chromosomes. A significant relationship was found between chromosome-specific variance and the length of the chromosome ($R^2 = 0.26$). This indicates that disease resistance is a largely polygenic trait with a large number of genes involved in the mechanisms of resistance but there are some chromosomal regions that explain a larger proportion of the variation.

INTRODUCTION

Parasite resistance for nematode infection is a complex trait of great importance in sheep and other livestock species. Breeding for sheep resistance is a viable method to reduce the effect of these nematodes on production and to reduce the cost of anthelminitic treatments (Dominik 2005). The identification of genes or genomic regions associated with sheep resistance would greatly accelerate genetic improvement in breeding programs. To date, genome wide association studies (GWAS) for parasite resistance have identified genetic variants that together explain only a small proportion of genetic variance of the trait (Kemper *et al.* 2012). Recently, Yang *et al.* (2010) showed that a considerable proportion of genetic variance can be explained by considering all single-nucleotide polymorphism SNPs simultaneously in a mixed linear model analysis. This mixed model has the potential to accumulate the effects of associated SNPs that might be too small to pass the significance threshold of single-SNP GWAS analysis.

To investigate in more details the role of SNP markers in parasite resistance, we used data from a large mixed breed population of sheep naturally challenged with *Haemonchus contortus*, and genotyped with the Illumina OvineSNP50 BeadChip. We estimated the additive genetic variance explained by genome-wide SNP data and compared this estimate to the additive genetic variance explained by pedigree. Furthermore, we partitioned the total genetic variance explained into genomic and polygenic components by fitting both of genomic data and pedigree simultaneously, and quantified the amount of genomic variance that can be explained by each chromosome.

MATERIALS AND METHODS

Animals and Phenotypes. Parasite resistance trait, as measured by WEC, was investigated in a multi-breed sheep population from the Sheep Cooperative Research Centre information nucleus flock (INF). A total of 7153 animals with both genotype and phenotype data were included in this analysis. Sires were either from Merino, terminal or maternal breeds and the size of resulting half-

sib families ranged from 20 to 91 with a median of 33 progeny. The breed content of the sheep population is shown in Table 1. Various breeds were represented in the population but with a significant proportion of Merino sheep, and only this breed had a substantial proportion of purebred animals. The remaining breeds were mainly represented by their crosses with Merino.

Table 1. Proportions of different breeds in the population

Breed	BL	COR	DH	SD	CO	PD	TX	AF	PER	PS	ME
Proportion	11.9	0.74	0.02	0.48	10.7	1.7	2.48	3.16	0.04	0.88	67.9

Border Leicester: BL, Corriedale: COR, Dorset Horn: DH, SD: Southdown, Coopworth: CO, Poll Dorset: PD, Texel: TX, Australian Finnsheep: AF, Perendale: PER, Prime Samm: PS, Merino:ME

Genotypes. Sheep were genotyped using the Illumina OvineSNP50 BeadChip. The following quality control measures were applied to the SNP data: SNPs were removed if they had a minor allele frequency (MAF) < 0.1%, a genotyping call rate < 90%, were not in Hardy-Weinberg equilibrium, and had no mapping information. The final data comprised genotypes for 47306 SNPs on 7153 animals.

Data analysis. The data were analyzed using the following mixed linear models:

Model1: $y^* = a + e$

Model2: $y^* = g + e$

Model3: $y^* = g + a + e$

Model4: $y^* = \sum_1^{26} g_i + a + e$

where y^* is a vector of adjusted phenotypic records, a is a vector of random additive genetic effects and assumed to be normally distributed with $N(0, A\sigma_a^2)$, A is the numerator relationship matrix (NRM) calculated from the pedigree data and σ_a^2 is the additive genetic variance, g is a vector of additive genetic effects accounted by all SNPs and assumed to be normally distributed with $N(0, G\sigma_g^2)$, G is the genomic relationship matrix (GRM) and σ_g^2 is the variance explained by all SNPs, g_i is a vector of additive genetic effects accounted by SNPs on the i^{th} chromosome and assumed to be normally distributed with $N(0, G_i\sigma_{gi}^2)$, G_i is the GRM built based on SNPs of the i^{th} chromosome, σ_{gi}^2 is the variance explained by SNPs on the i^{th} chromosome, and e is a vector of random residuals. The variance components were estimated using GCTA software (Yang *et al.* 2011).

Phenotypic records were adjusted for systematic environmental effects using the following model: $y = 1\mu + Xb + ZQa + e$, where y is a vector of cube root transformed WEC records, μ is the mean, X and Z are design matrices of fixed and random effects respectively, Q is a matrix containing breed proportions for each animal calculated from the pedigree records, b is a vector of fixed effects, a is a vector of random breed effects assumed to be normally distributed $\sim N(0, \sigma_q^2 I)$, where σ_q^2 is the variance of breed effects. The following fixed effects were included in the model: age at WEC recording, sex, rearing type, and contemporary groups formed using INF flock, group of management and year of birth.

RESULTS AND DISCUSION

The proportion of additive genetic variance explained by SNP markers or pedigree relative to the total variance corresponds to heritability of WEC. The estimated variance components from models 1to 4 are shown in Table 2. In this analysis, all the additive genetic variance explained by pedigree could be captured by the Ovine 50K SNP chip markers. This clearly indicates that the Ovine 50K SNP chip markers can trace all polygenic relationships due to sharing of causative

variants in this large mixed breed population of sheep.

Table 2. Genetic and genomic variances for WEC estimated in models 1 to 4

Model 1		Model 2		Model 3			Model 4
$\hat{\sigma}_a^2$	h_a^2	$\hat{\sigma}_g^2$	h_g^2	$\hat{\sigma}_a^2$	$\hat{\sigma}_g^2$	h_{a+g}^2	$\sum_{i=1}^{26} \sigma_{gi}^2$
0.67	0.147	0.67	0.147	0.20	0.56	0.1645	0.55

In model3, Both of GRM and NRM were fitted simultaneously in order to separate effects of pedigree (polygenic) relationships from genomic (SNPs) relationships. The total genetic variance estimated when both effects were fitted simultaneously was higher than the situation where each of them was fitted alone. Moreover, the residual (unexplained) variance of the total phenotypic variance was reduced in model 3 compared to the residual variance in model 1 and model 2. This indicates that there is not complete overlap between polygenic and genomic effects. In this model, a large proportion of total genetic variance was explained by genomic relationships (73.7%) while the remaining variance was explained by pedigree effects (26.3%).

In model4, the genomic variance explained by genomic relationships was further partitioned into 26 chromosomes. A GRM was built for each individual chromosome then all GRMs were fitted simultaneously to estimate the amount of genomic variance that can be attributed to each chromosome. The sum of estimates due to individual chromosomes was slightly lower than the genomic variance explained by genomic relationships in model 3. This suggests a very weak covariance between genomic relationships on different chromosomes.

A significant relationship between chromosomal length and the genomic variance explained by each chromosome (Figure 1 over page) is consistent with the hypothesis that many alleles with small effects contribute much of the genetic variation of the trait. It is notable, however, that five chromosomes exhibited higher contributions to genetic variance than expected given their size. This demonstrates that some chromosomal regions have effects larger than expected on a purely infinitesimal model.

In conclusion, our results suggest that the Ovine SNP50 array can capture a large proportion of genetic variance for WEC trait in a large multi-breed population of sheep. The same proportion of genetic variance can be attributed to individual chromosomes with a significant relationship between chromosomal length and the genomic variance explained by each chromosome.

REFERENCES

Dominik S. (2005) *Genet. Sel. Evol* **37**: S83.

Kemper K. E., Emery D., Bishop S. C., Oddy H., Hayes B. J., Dominik S., Henshall J. M. and Goddard M. E. (2011) *Genet Res Camb.* **93**: 203.

Yang J., Benyamin B., McEvoy B. P., Gordon S., Henders A. K., Nyholt D. R., Madden P. A., Heath A. C., Martin N. G., Montgomery G. W., Goddard M. E. and Visscher P. M. (2010) *Nature Genetics*. **42**: 565.

Yang J., Lee H., Goddard, M. E. and Visscher P. M. (2011) Am *J Hum Genet*. **88**: 76.

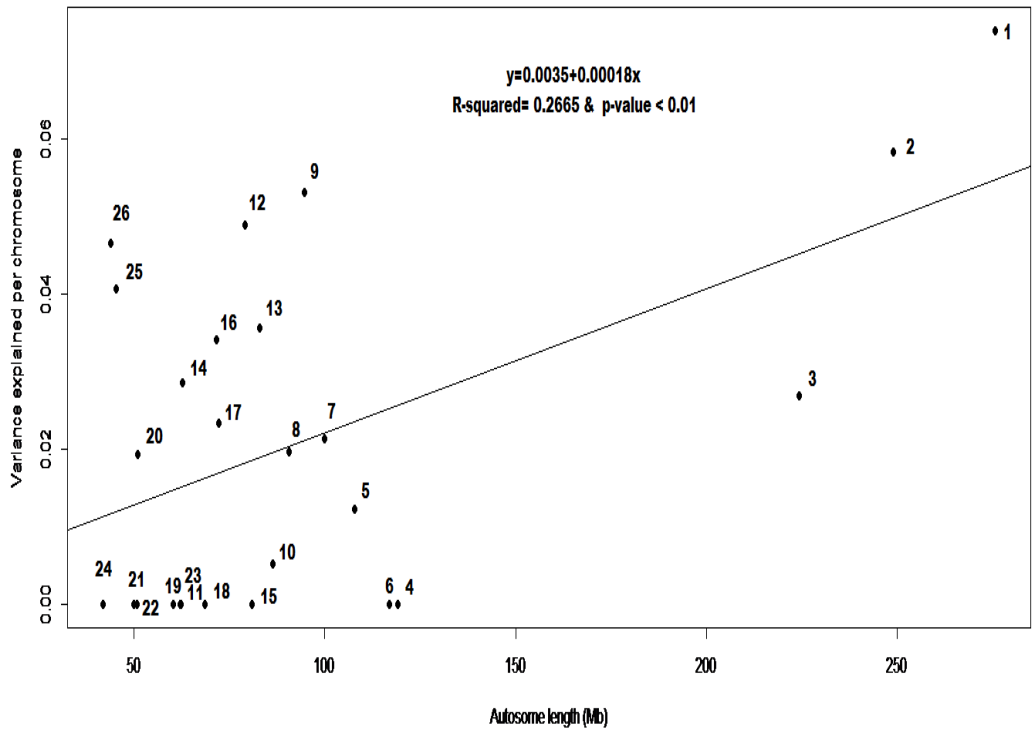


Figure 1. Amount of genomic variance explained per chromosome. The equation ($y = 0.0035 + 0.00018x$) corresponds to linear regression where y is the genomic variance explained by each chromosome and x is the chromosomal length in mega bases (Mb).

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:416-419

MANAGING COST OF PHENOTYPING

Cécile Massault[1], Brian Kinghorn[1] and Julius van der Werf[1,2]

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

SUMMARY

We investigate a way to reduce phenotyping cost with an approach that uses a differential evolution algorithm to optimize which sets of animals to phenotype. The "fitness function" to optimize was the average accuracy of selection candidates (prediction error variance covariance function) when phenotyping 15 animals of four small illustrative pedigrees (different family structure, balanced and unbalanced family sizes). We compared these results to three other strategies (random phenotyping, 15 selection candidates and sires and 12 selection candidates) for different heritability. The tactical approach was either the best strategy or shared highest position on the podium with an equally performing strategy. Phenotyping patterns are impacted by heritability, family structure and family size. This tactical approach to phenotyping is a first step towards a tool that can more generally help decide which animals should be phenotyped or genotyped to optimize breeding programs.

INTRODUCTION

Some traits which impact on profit are not routinely recorded for genetic evaluation or evaluated through a correlated trait, such as traits that are difficult to record (e.g. carcass traits) and/or expensive to measure (e.g. methane emission of cattle). A strategy that reduces costs of phenotyping would allow the integration of valuable novel traits and reduces the cost of implementing genomic selection for some species (i.e. sheep, beef cattle).

We investigate in this paper an approach to manage total cost of phenotyping. The objective of our approach is to find the best set of animals to phenotype for a given number of phenotypes that maximises the EBV accuracy for selection candidates. The total number of phenotypes to be measured will be dictated by the financial resources of the breeder or farmer. We used a differential algorithm to find the set of animals to phenotype that incur the highest average accuracy. The fitness of the solution was evaluated by an objective function that calculated the average accuracy of selection candidates based the prediction error variance covariance (PEVC) given by the traditional best linear unbiased prediction method (BLUP, Henderson 1984). Using four simple pedigrees, we investigate the phenotyping patterns of optimal average accuracy of selection candidates that results from our tactical approach (depending on family structure and family size) when phenotyping 15 animals. We also compared our approach to three others phenotyping strategies (phenotyping random animals, only selection candidates or sires and selection candidates).

MATERIAL AND METHODS

Phenotyping approach. To find which animals should be phenotyped given a maximum number, we used a differential evolution algorithm (DE, Price and Storn 1997), that judges on the fitness of solutions by maximizing accuracy obtained through the fitness function described below. The DE is set up with a population of 10 solutions that evolved for a maximum of 100,000 generations. Two other criteria can stop the evolution prematurely when they have been all met: the DE run for at least 1000 generations without any improvement and solution has to exceed 99.5% of current predicted asymptotic maximum solution (see Kinghorn 2008 for full description).

Fitness function. The fitness function establishes the average accuracy of selection candidates given a set of animals to phenotype for the simple case of a single trait with no fixed effects. The prediction error variance-covariance corresponds to the inverse of the left hand side of the mixed model equations used in traditional BLUP genetic evaluation:

$$PEVC_i = diagonal(Z'Z + \lambda A^{-1})^{-1} \times MSE$$

where Z is the matrix that contains information on which animals is phenotyped, A is the relationship matrix and $\lambda = \frac{1-h^2}{h^2}$. MSE represents the mean squared error, assumed to be 1. Individual accuracy was then calculated as:

$$r_{IH\,i} = \sqrt{1 - \lambda PEVC_i}$$

The value maximized by the fitness function was the average accuracy of selection candidates.

Pedigrees. We designed 4 representative pedigrees with specific characteristics to evaluate our approach for common population structures found in animal breeding. The pedigrees are described in **Table 1**. They all comprise 3 generations and the last generation (offspring) represents selection candidates. We investigated the impact of different family structure (half-sib and full-sib), of balanced and unbalanced family sizes. Family sizes in the unbalanced pedigrees are 2 (small family), 10 (medium family), and 18 (large family) offspring. We looked at the particular cases of phenotyping 15 animals which corresponds to half of the selection candidates and for three heritabilities of 0.1, 0.5 and 0.8 (TACT strategy). We also examined the accuracy when phenotyping 15 animals randomly across generations (RAND strategy), phenotyping 15 selection candidates randomly (OFFS strategy) and phenotyping the sire of each family and 12 random offspring (SIRE strategy) for each pedigree and each heritability (0.1, 0.5 and 0.8).

Table 1. Illustrative pedigrees family structures

	Structure	**# GP**	**# sires**	**# dams**	**# offspring**	**Family size**		
PED1	HS	66	3	30	30	10	10	10
PED2	FS	12	3	3	30	10	10	10
PED3	HS	66	3	30	30	2	10	18
PED4	FS	12	3	3	30	2	10	18

RESULTS AND DISCUSSION

Table 2 reports the percentage of maximum accuracy achieved by the four different phenotyping strategies. As expected, RAND was the least efficient strategy. Out of the 12 scenarios, TACT always performs best. In 2 cases, TACT was equivalent to OFFS (PED2 heritability of 0.5 and 0.8) and in 2 cases equivalent to SIRE (PED1, heritability 0.1 and 0.5). This indicates that our tactical approach found the best strategy, which also happened to be one of other strategy. Figure 1 shows the distribution of phenotypes for the families when phenotyping 15 animals in PED1 to PED4 for heritability of 0.1, 0.5 and 0.8 using the TACT approach. We are considering three factors that impact on the phenotyping pattern given a fixed number of individuals to phenotype and a fixed number of selection candidates: heritability, family structure and balanced/unbalanced family size.

Table 2. Percentage of maximum accuracy (when all animals are phenotyped) captured when phenotyping 15 animals for PED1 to PED4 for heritability of 0.1, 0.5 and 0.8 with TACT, RAND, OFFS and SIRE phenotyping strategies

	Heritability	TACT	RAND	OFFS	SIRE
PED1	0.1	**63**	33	51	**63**
	0.5	**69**	38	60	**69**
	0.8	**72**	40	62	69
PED2	0.1	**73**	63	71	69
	0.5	**84**	75	**84**	80
	0.8	**84**	76	**84**	80
PED3	0.1	**66**	32	61	64
	0.5	**71**	40	69	69
	0.8	**73**	42	70	71
PED4	0.1	**77**	67	75	73
	0.5	**85**	77	84	82
	0.8	**85**	77	82	81

Optimal phenotyping patterns vary with heritability except for the case of PED2. We can also observe difference in optimal phenotyping pattern when the structure is different. For example, for the same heritability of 0.1, TACT recommends to phenotype 1 sire and 4 offspring per family with a half-sib structure (PED1), while it recommends to phenotype 5 offspring per family with a full-sib structure (PED2). Finally, we can note that phenotyping patterns are different between balanced and unbalanced family sizes (e.g. PED2 and PED4). Phenotyping patterns are clearly impacted by family structure, heritability and family size.

We also use the tactical approach with a pedigree of 50 half-sib families of various sizes (3-16) for 3 heritabilities and phenotyped patterns observed were similar to the ones described in Figure 1 for PED3. The phenotyping approach is a useful tool to find the best set of animals to phenotype for a given number of phenotypes. The advantage of such an approach is that accuracy is only slightly lower, while cost of phenotyping can be significantly reduced. This could permit the inclusion in the breeding goal of new traits that are expensive to measure, including genotyping. In this study, we maximized accuracy of the youngest cohort of selection candidates, as it affects short term response. Genetic gain is impacted by accuracy, as well as generation interval. Increasing accuracy of younger animals tends to reduce the generation interval and therefore increases genetic gain. Optimal response in the longer term is obtained when genetic diversity is also considered. Phenotyped animals are more informative and are more likely to be selected for the next generation and phenotyping many related animals is more likely to increase inbreeding Constraints on inbreeding as used in optimum contributions selection (Sonesson and Meuwissen 2000) can also be added to the tactical phenotyping approach described here. A further step is to extend the tactical approach to phenotyping to a multi-trait case, and subsequently to include genotyping for the case of gnomic selection. Further work will also address the need to reduce the computational time e.g. by using approximate accuracies (Meyer 1989) for larger pedigrees.

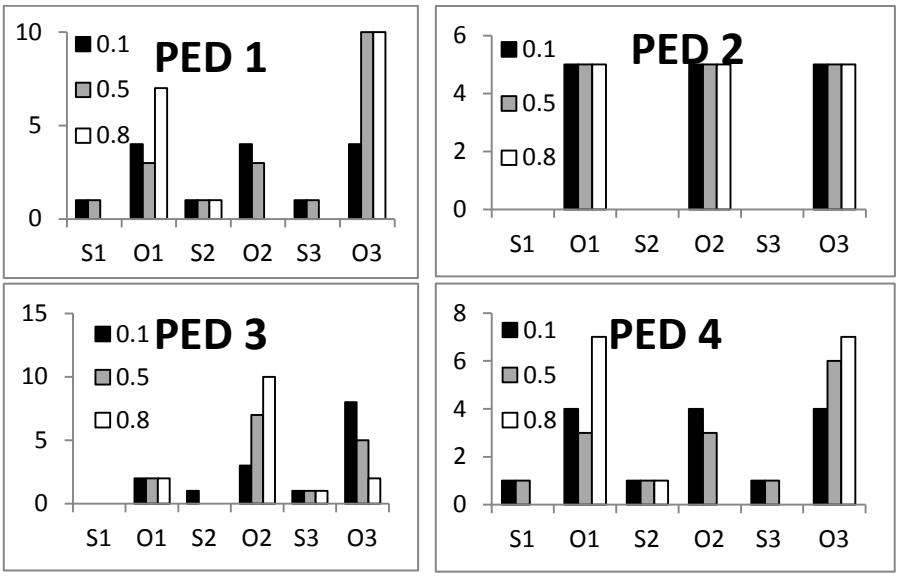


Figure 1. Optimal phenotyping patterns for PED1, PED2, PED3 and PED4 when phenotyping 15 animals for heritability 0.1 (black), 0.5 (grey) and 0.8 (white). S*i* refers to sire of family *i*, O*i* refer to offspring of family *i*.

REFERENCES

Kinghorn B.P. (2008) "Application of Evolutionary Algorithms to solve complex problems in Quantitative Genetics and Bioinformatics" ed. Centre for Genetic Improvement of Livestock University of Guelph.

Meyer, K. (1989) *Livestock Animal Production.* **21**:87

Sonesson, A.K. and Meuwissen, T.H.E. (2000*) Genetics Selection Evolution* **32**: 231

Storn, R. and Price, K. (1997) *Journal of global optimization.* **11**:341

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:420-423

A GENOME-WIDE ASSOCIATION STUDY FOR HEIGHT AT WITHERS IN RACING QUARTER HORSE

C.T. Meira[1], M.R.S. Fortes[2], M.M. Farah[1], L.R. Porto-Neto[3], R.A. Curi[4], S.S. Moore[2] and M.D.S. Mota[4]

[1]Faculdade de Ciências Agrárias e Veterinárias, Universidade Estadual Paulista, Jaboticabal, SP 14884-900, Brazil
[2]The University of Queensland, Queensland Alliance for Agriculture and Food innovation, Centre for Animal Science, Brisbane, Qld 4062, Australia
[3]CSIRO Food Futures Flagship and Animal, Health and Food Sciences, 306 Carmody Road, St Lucia, QLD 4067, Australia
[4]Faculdade de Medicina Veterinária e Zootecnia, Universidade Estadual Paulista, Botucatu, SP 18618-000, Brazil

SUMMARY

Height is a classic polygenic trait and an important conformation characteristic for horses. A genome-wide association study for height at withers was conducted in 120 racing Quarter Horses, which were genotyped using the Illumina EquineSNP50 Bead chip. Association analysis was performed with 40,787 SNPs (after quality control) using Qxpak5 software. The analysis revealed 8 chromosomal regions that harboured 23 SNPs associated ($P < 0.0001$) with height at withers. These regions were on chromosomes 3, 4, 5, 6, 7, 9, 17 and 30. A positional and functional candidate gene emerging from this study is *WISP1*, a gene previously related to bone development. Further studies are required to confirm these SNP associations and candidate genes in additional populations and other horse breeds.

INTRODUCTION

One of the highlights from the analysis of the horse genome project is its complete sequencing from a Thoroughbred animal (EquCab2.0) and, from this, the identification of 1,162,753 single-nucleotide polymorphisms (SNPs) in different breeds (Wade *et al*. 2009). As a result, genome-wide association studies (GWAS) based on SNPs and high density chips have been used to detect changes caused by genetic selection and to identify quantitative trait loci (QTLs). For example, GWAS have identified SNPs on *Equus caballus* autosomes (ECA) 18, within and proximal to the myostatin gene that are associated with racing performance in Thoroughbred horses (Hill *et al*. 2010). According to Signer-Hasler *et al*. (2012), less information is available on the genetics of polygenic quantitative traits in horses than in other species. Examples of polygenic traits are conformation traits including height at withers. Conformation traits are important criterion for selection and evaluation of a horse, particularly if a horse is a candidate for an athletic career (such as racing), long years of sound service or breeding (Thomas 2005). The objective of our study was to perform GWAS to identify chromosomal regions and positional candidate genes associated with height at withers in the racing Quarter Horse.

MATERIALS AND METHODS

Animals, Traits and Genotypes. Blood samples for DNA extraction were obtained from 120 racing Quarter Horses, born between 1985 and 2007 and registered at the Brazilian Association of Quarter Horse Breeders. Animals of this racing line, including 18 males and 102 females born to 48 stallions and 107 mares, were from five properties located in São Paulo state, Brazil. Height at withers was measure from the tallest point of the thoracic vertebrae to ground (Thomas 2005). This measure was performed by the same person with a tape measure and measuring stick, always

on the right side of the animal, with the horse standing with front and rear legs perpendicular to the ground. The DNA samples were genotyped with the Equine SNP50 BeadChip (Illumina, Inc., USA) using the HiScan system at the Faculty of Agricultural and Veterinary Sciences, Unesp, Jaboticabal, São Paulo, Brazil. This array contains 54,602 SNPs derived from the EquCab2.0 SNP Collection database, with a mean density of one SNP per 43.2 kb. Repeat samples were included for quality control and Genome Studio 2011.1 software (Illumina Inc., San Diego, CA) was used to call genotypes. SNP with call rates < 90%; cluster separation < 0.3; call frequency < 0.9; p-value < 1 x 10-3 for Hardy-Weinberg equilibrium and minor allele frequency < 0.05 were discarded. Quality control resulted in 40,787 autosomal SNPs for 120 animals.

Statistical Analysis. An association analysis was performed with the 40,787 SNPs for height at withers using Qxpak5 (Perez-Enciso and Misztal 2011) and fitting one SNP at a time. Qxpak5 relies on the well-known theory of mixed models, performing a likelihood ratio test with every SNP in turn, testing the model with the SNP versus the model without the SNP, against a chi-square distribution with 1 degree of freedom. Correction for multiple tests considered two metrics: false discovery rate (FDR) and Qvalue calculated with the package for R (Version 2.10). The percentage of the genetic variance accounted by the *i*-th SNP was estimated according to the following formula:

$$\%V_i = 100\left(\frac{2p_i q_i \hat{a}_i^2}{\sigma_g^2}\right)$$

where p_i and q_i are the allele frequencies for the *i*-th SNP estimated across the entire population, a_i is the estimated additive effect of the *i*-th SNP on the trait in question, and σ^2_g is the REML estimate of the (poly-)genetic variance for the trait.

RESULTS AND DISCUSSION

Descriptive statistics of the data are reported in Table 1. In accordance to previous reports (Signer-Hasler *et al*. 2012) moderate estimated heritability for height at withers was observed. Low phenotypic, genetic and residual variances were also observed.

Table 1. Summary statistics and heritability estimates for height at withers (meter) of the racing Quarter Horse

Parameter	Height at withers
N (animals)	120
Mean	1.55
Sample variance	0.002
Minimum	1.46
Maximum	1.72
σ^2_q	0.0008
h^2	0.58

We report the results of GWAS for height at withers using a mixed-model with fixed effect of sex and age at the time of trait measurement as a linear covariate. Twenty three significant SNPs were detected (P<0.0001) on ECA 3, 4, 5, 6, 7, 9, 17 and 30 (Figure 1 and Table 2). This *P*-value corresponds to FDR of 0.17 and q-values between 0.08 and 0.15, which are indicative of a possible true association, given the small sample size and the fact that each of these SNP accounted for an important proportion of the genetic variance (Table 2). On ECA 3, 4, 5, 7, 9 and 30 were found 2, 30, 19, 2, 2 and 1 genes located within the associated region, respectively. Among these results

some of the most interesting findings were on ECA 9. In ECA 6 and 17 no annotated genes were found within the associated region (1 Mb window around associated SNP).

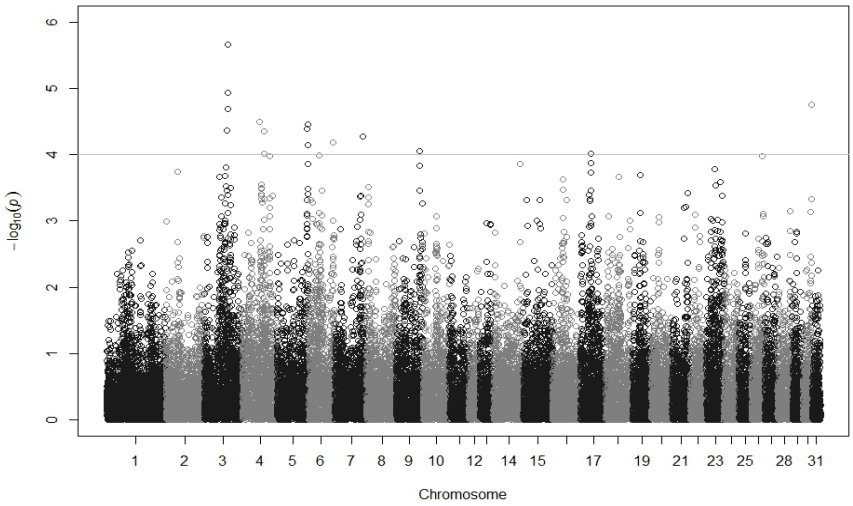


Figure 1. Manhattan plot of *P*-value for height at withers. The log inverse *P-values* estimated for each polymorphism are plotted in the y-axis. Chromosome number is plotted in the x-axis. Horizontal line indicates the threshold $P<0.0001$.

Signer-Hasler *et al*. (2012) identified eight SNPs within two QTL regions for height at withers on ECA 3 and ECA 9 in Franches-Montagnes horses. The two QTL regions are mapped near the *LCORL*/*NCAPG* (ECA 3) and *ZFAT* (ECA 9) genes. Tetens *et al*. (2013) have also identified significant association signal on the distal end of ECA 3 for height at withers in German Warmblood horses explaining ~18% of the phenotypic variance. In our study associations on ECA 3 and 9 were also reported but in different regions (implicating other genes: *LPHN3*, *TECRL* on ECA 3 and *WISP1*, *NDRG1* on ECA 9). Of these genes, *WISP1* is the one with a reported functional connection to growth and development playing an important regulatory role during bone development and fracture repair (French *et al*. 2004). This gene encodes a member of the *WNT1* inducible signaling pathway (WISP) protein subfamily, which belongs to the connective tissue growth factor (CTGF) family. *WNT1* is a member of a family of cysteine-rich, glycosylated signaling proteins that mediate diverse developmental processes (NCBI, 2013).

The percentage of the genetic variance explained by each SNP ranged 20.01% – 36.75% (Table 2), which was higher than normally encountered in GWAS (Tetens *et al*. 2013). These high values may be due to the small number of animals used in the study. Also, it is important to notice that these high percentages are in relation to a low genetic variance and that they were estimated in the same population used for the discovery of the SNP association. Despite these limitations, our results are similar to those of Makvandi-Nejad *et al*. (2012) that identified four loci on chromosomes 3, 6, 9 and 11, which together explained 83% of size variance in 48 horses from 16 breeds. According to Signer-Hasler *et al*. (2012), the genetic architecture of the digressed estimated breeding values (dEBV) for height at withers is characterized by a few genes with major effects and a large number of genes with small effects. Therefore, results reported here and elsewhere seem consistent.

Table 2. Effect, *P*-values and proportion of the variance explained for SNPs associated (*P*<0.0001) with height at withers of the racing Quarter Horse

SNP Name	ECA	Position	MAF	Effect	*P*-value	%Variance
BIEC2-789895	3	69,347,313	0.33	-0.021	4.29003E-05	24.19
BIEC2-789896	3	69,347,393	0.33	-0.021	4.29003E-05	24.19
BIEC2-789900	3	69,351,329	0.33	0.021	4.29003E-05	24.19
BIEC2-792718	3	73,581,276	0.38	-0.024	2.19999E-06	34.04
BIEC2-792768	3	73,797,941	0.34	0.022	2.03002E-05	27.05
BIEC2-792776	3	73,811,436	0.34	-0.022	2.03002E-05	27.05
BIEC2-792783	3	73,813,893	0.34	0.022	2.03002E-05	27.05
BIEC2-792814	3	73,881,123	0.38	-0.023	1.16001E-05	31.23
BIEC2-792839	3	74,052,752	0.34	0.022	2.03002E-05	27.05
BIEC2-866887	4	54,805,325	0.38	-0.02	3.18002E-05	23.44
BIEC2-869000	4	66,635,563	0.36	0.022	4.49997E-05	27.82
BIEC2-869084	4	67,096,727	0.35	-0.021	9.66006E-05	24.94
BIEC2-929949	5	93,714,559	0.48	-0.02	4.06004E-05	24.94
BIEC2-931221	5	95,656,828	0.18	-0.027	7.05992E-05	26.31
BIEC2-931466	5	95,894,134	0.40	0.021	3.51002E-05	26.55
BIEC2-931509	5	95,951,807	0.18	-0.027	7.05992E-05	26.31
BIEC2-931513	5	95,955,507	0.18	-0.027	7.05992E-05	26.31
BIEC2-931518	5	95,959,168	0.18	-0.027	7.05992E-05	26.31
BIEC2-1186793	6	76,390,403	0.45	0.018	6.58006E-05	20.01
BIEC2-1011792	7	87,083,836	0.18	0.028	5.33998E-05	28.48
BIEC2-1105149	9	73,886,834	0.42	0.02	8.71004E-05	24.37
BIEC2-375392	17	33,111,862	0.49	-0.018	9.5801E-05	20.24
BIEC2-828161	30	26,347,707	0.25	-0.028	1.75001E-05	36.75

CONCLUSION

Genomic regions on ECA 3, 4, 5, 6, 7, 9, 17 and 30 were associated with height at withers in the racing Quarter Horse. A total of 56 genes mapped to these regions and thus emerge as positional candidates for height at withers. However, most of these genes have no obvious function related to height. A positional and functional candidate gene from this study is *WISP1*, a gene related to bone development. Further studies are required to confirm these SNP associations and candidate genes in other populations and breeds.

ACKNOWLEDGEMENTS

The authors thank the financial support of FAPESP and CAPES. (Proc. n° 13842-12-9).

REFERENCES

French D.M., Kaul R.J., D'Souza A.L., Crowley C.W., Bao M., Frantz G.D., Filvaroff E.H. and Desnoyers L. (2004) *Am. J. Pathol.* **165**: 855.

Hill E.W., Mcgivney B.A., Gu J., Whiston R. and MacHugh D.E. (2010) *BMC Genomics.* **11**:552.

Makvandi-Nejad S., Hoffman G.E., Allen J.J., Chu E., Chandler A.M., Loredo A.I., Bellone R.R., Mezey J.G., Brooks S.A. and Sutter N.B. (2012) *PloS One.* **7**:e39929.

Perez-Enciso M. and Misztal I. (2011) *BMC Bioinformatics.* **12**:202.

Signer-Hasler H., Flury C., Haase B., Burger D., Simianer H., Leeb T. and Rieder S. (2012) *PloS One.* **7**:e37282.

Tetens J., Widmann P., Kühn C. and Thaller G. (2013) *Anim. Genet.* doi: 10.1111/age.12031

Thomas H.S. (2005) 'The Horse Conformation Handbook'. Storey Publishing, North Adams.

Wade C.M. and Equine Genome Sequencing Consortium (2009) *Sci.* **326**: 865.

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:424-427

POST-ESTIMATION PENALIZATION: MORE 'PEP' FOR ESTIMATES OF GENETIC COVARIANCE MATRICES

Karin Meyer

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

SUMMARY

Maximum likelihood estimation of genetic covariances subject to a penalty to reduce sampling variation has been shown to yield improved estimates, especially for analyses comprising many traits. However, this can increase computational requirements substantially. Similarly, penalties have been found to be beneficial in a maximum likelihood based approach for pooling results from analyses of subsets of traits. This paper examines the scope for using the latter method to apply penalties to results from multivariate analyses in a computationally undemanding post-estimation step. A simulation study is presented demonstrating that even slight changes to estimates in this way can result in 'regularized' values markedly closer to population values than standard, unpenalized estimates.

INTRODUCTION

Restricted maximum likelihood (REML) estimation of genetic covariance matrices subject to a penalty to borrow strength from their phenotypic counterparts has been shown to 'improve' estimates, i.e. to result in estimates which are, on average, closer to the population values than standard (unpenalized) estimates (Meyer and Kirkpatrick 2010; Meyer 2011b). Whilst highly appealing, penalized estimation can increase computational requirements by orders of magnitude. This may be prohibitive for multivariate analyses comprising numerous traits where penalization is likely to be most beneficial. Recently, Meyer (2013) demonstrated that penalization can also yield 'better' estimates when employing a maximum likelihood approach to combine estimates from analyses of overlapping subsets of traits to construct overall covariance matrices. This suggests that the same procedure might be used to modify estimates from a single, unpenalized multivariate analysis in a simple, computationally undemanding post-estimation penalization (PEP) step. This paper presents a simulation study examining the scope for PEP.

PENALIZING ESTIMATES

Penalized REML estimates are obtained by maximising the log likelihood ($\log \mathcal{L}$) in a multivariate analysis subject to a penalty ($\mathcal{P}$), $\log \mathcal{L} - \frac{1}{2}\psi\mathcal{P}$, with $\mathcal{P}$ a suitable function of the covariance components to be estimated and $\psi \geq 0$ the so-called tuning factor determining the stringency of penalization. For PEP, unpenalized estimates ($\psi = 0$) of covariance matrices are first obtained preforming a standard, multivariate analysis. In a second step, these are 'converted' to 'data' by forming matrices of mean squares and crossproducts corresponding to a selected simple, balanced pedigree structure from the estimates. Together with the assumed pseudo pedigree, these matrices then provide a likelihood function which again is maximised subject to a penalty. Further details are given in Meyer (2013).

MATERIAL AND METHODS

Data for 10 traits were simulated for 250 independent families of size 8, as per Bondari *et al.* (1978)'s design, sampling genetic and residual effects from appropriate multivariate Normal distributions for two sets of population parameters. For case A, all heritabilities were assumed equal to 0.4, for case B values ranged from 0.6 to 0.2, $0.2 + 0.1 \bmod(i, 5)$ for trait i. All genetic correlations were set to 0.5 and all residual values to 0.2. Phenotypic variances for the $i-$th trait were $\bmod(i, 3) + 1$.

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

This yielded canonical eigenvalues (λ_i) of 0.57 and 9×0.29 for case A and from 0.69 to 0.14 for B. A total of 250 replicates per case were carried out.

Analyses. Estimates of genetic ($\mathbf{\Sigma}_G$) and residual ($\mathbf{\Sigma}_E$) covariance matrices were obtained from multivariate REML analyses (MUV), with and without penalties. Unpenalized estimates were then modified by PEP, considering a paternal half-sib design (PHS) comprising $s = 2$ sires and $n = 2$ progeny per sire, a hierarchical full-sib design (HFS) with $s = 2$, $d = 2$ dams per sire and $n = 2$, and 2 families with Bondari's design (BON, $n = 8$) as pseudo pedigree structures. Penalties considered were

$$\mathcal{P}_\lambda = \textstyle\sum_i (\hat{\lambda}_i - \bar{\lambda})^2 \tag{1}$$

$$\mathcal{P}_\lambda^{\ell 2} = \textstyle\sum_i (\log(\hat{\lambda}_i) - \bar{\lambda}_1)^2 + (\log(1 - \hat{\lambda}_i) - \bar{\lambda}_2)^2 \tag{2}$$

$$\mathcal{P}_\Sigma = \log|\hat{\mathbf{\Sigma}}_G| + \mathrm{tr}(\hat{\mathbf{\Sigma}}_G^{-1}\hat{\mathbf{\Sigma}}_P^0) + \log|\hat{\mathbf{\Sigma}}_E| + \mathrm{tr}(\hat{\mathbf{\Sigma}}_E^{-1}\hat{\mathbf{\Sigma}}_P^0) \tag{3}$$

$$\mathcal{P}_R = \log|\hat{\mathbf{R}}_G| + \mathrm{tr}(\hat{\mathbf{R}}_G^{-1}\hat{\mathbf{R}}_P^0) + \log|\hat{\mathbf{R}}_E| + \mathrm{tr}(\hat{\mathbf{R}}_E^{-1}\hat{\mathbf{R}}_P^0) \tag{4}$$

with $\bar{\lambda}$, $\bar{\lambda}_1$ and $\bar{\lambda}_2$ the means of estimates $\hat{\lambda}_i$, $\log(\hat{\lambda}_i)$ and $\log(1 - \hat{\lambda}_i)$, respectively, $\hat{\mathbf{\Sigma}}_P^0$ the unpenalized estimate of the phenotypic covariance matrix, $\hat{\mathbf{R}}_G$ and $\hat{\mathbf{R}}_E$ the estimates of the genetic and residual correlation matrix, and $\hat{\mathbf{R}}_P^0$ their unpenalized, phenotypic counterpart. In addition, simple 'bending' (BEN) was applied, regressing $\hat{\lambda}_i$ towards $\bar{\lambda}$, as proposed by Hayes and Hill (1981).

Degree of penalization. Tuning factors for each replicate were determined as values of ψ for which a) the sum of losses in $\hat{\mathbf{\Sigma}}_G$ and $\hat{\mathbf{\Sigma}}_E$ was smallest ("Optimum"), and b) the largest value for which the deviation (absolute value) of $\log \mathcal{L}$ from the (unpenalized) maximum did not exceed $\chi^2_{1,5\%} = 1.92$ ("$\Delta\mathcal{L}$"). In addition, fixed values selected to provide "very mild" and "mild" penalties were used, c) $\psi = 0.1$ for MUV and $\psi = 0.001$ for PEP, and d) $\psi = 1.0$ (MUV) and $\psi = 0.01$ (PEP). For BEN, regression coefficients were set to 0.98 for "very mild" and 0.90 for "mild" shrinkage.

Summary statistics. The deviation of estimated covariance matrices ($\hat{\mathbf{\Sigma}}$) for q traits from the respective population values ($\mathbf{\Sigma}$) was evaluated as the so-called entropy loss (L_1) and, with $\bar{L}_1(\cdot)$ denoting the mean over replicates and $\hat{\mathbf{\Sigma}}^\psi$ the estimate for a tuning factor of ψ, the percent reduction in average loss (PRIAL),

$$L_1(\mathbf{\Sigma}, \hat{\mathbf{\Sigma}}) = \mathrm{tr}(\mathbf{\Sigma}^{-1}\hat{\mathbf{\Sigma}}) - \log|\mathbf{\Sigma}^{-1}\hat{\mathbf{\Sigma}}| - q \qquad \text{and} \tag{5}$$

$$\mathrm{PRIAL} = 100\big[1 - \bar{L}_1\big(\mathbf{\Sigma}, \hat{\mathbf{\Sigma}}^\psi\big)/\bar{L}_1\big(\mathbf{\Sigma}, \hat{\mathbf{\Sigma}}^0\big)\big]. \tag{6}$$

RESULTS

The distribution of losses in estimates of $\mathbf{\Sigma}_G$ for case B is summarized in Figure 1. Shown on the left of each panel are losses for unpenalized estimates from standard, multivariate analyses. Penalization using the optimum tuning factor (top panel) reduced both the mean and variation in losses dramatically for all penalties and both MUV and PEP. Moreover, simple 'bending' performed similar to a penalty encouraging shrinkage of the canonical eigenvalues towards their mean. In line with previous experience with MUV for cases with a substantial spread of population canonical eigenvalues (Meyer 2011b), a penalty shrinking correlation matrices towards their phenotypic counterpart ($\mathcal{P}_R$) was most effective, with MUV yielding a PRIAL of 74% and PEP of 61%.

In practice, the optimal tuning factor is unknown and, for MUV, estimating ψ using cross-validation techniques not only imposes a considerable computational burden but also has been found to reduce PRIALs achieved, typically by at least 10-15%. Hence, selecting a value of ψ which limits the change in $\log \mathcal{L}$ from the maximum (at $\psi = 0$) has been suggested as a simple, pragmatic alternative, and has been shown to yield losses $L_1(\cdot)$ closely related to optimal values (Meyer 2011a,b). As demonstrated

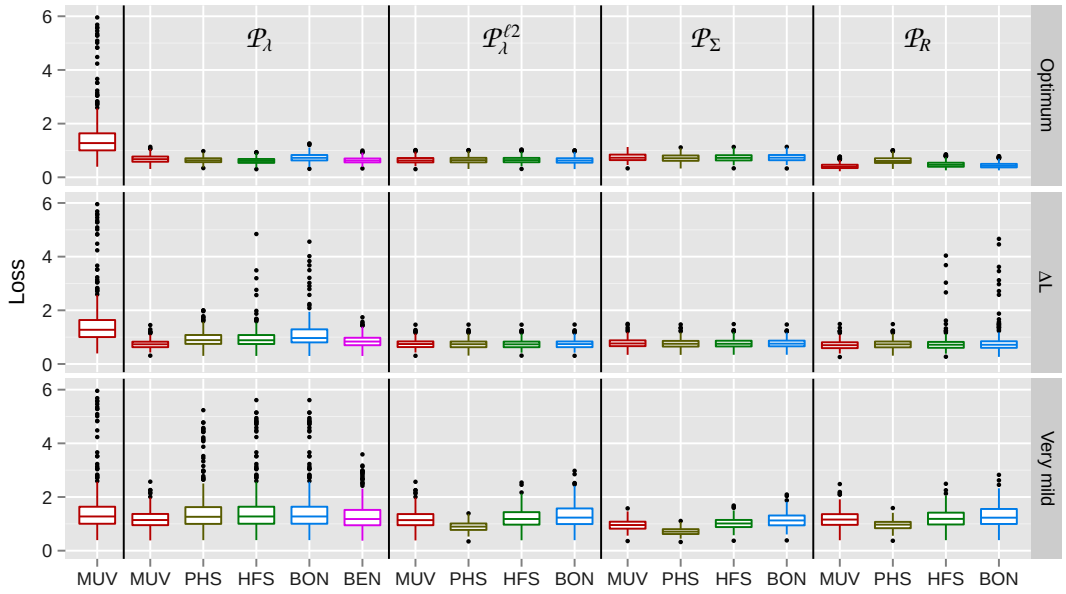


Figure 1. Distribution of entropy loss in estimates of the genetic covariance matrix for case B

in Figure 1 (middle panel), this strategy also performed well for PEP, especially for the simplest pseudo pedigree structure. For $\mathcal{P}_R$, PRIALs obtained were 55 and 53% for MUV and PEP, respectively. Limiting $\Delta\mathcal{L}$ to a value for which the change in even a single parameter was not statistically significant (at an error probability of 5%) yielded much milder penalization than for the optimum values of ψ, which resulted in average changes in $\log \mathcal{L}$ from -7.1 to -16.8. However, even such a mild penalty consistently provided substantial reductions in sampling variation and losses in estimates of the genetic covariance matrix. In contrast, whilst beneficial throughout, effects of penalization for a small, fixed value of ψ varied markedly with the type of penalty and pseudo-pedigree structure chosen.

Table 1 summarizes PRIALs and the corresponding mean change in $\log \mathcal{L}$ for selected examples. With 9 of the population canonical eigenvalues equal, stringent penalties on the λ_i, $\mathcal{P}_\lambda$ or $\mathcal{P}_\lambda^{\ell 2}$, performed best for case A, achieving optimum PRIALs (not shown) as high as 79% accompanied by changes in $\log \mathcal{L}$ around -17, with little difference between MUV and PEP. Conversely, choosing ψ on the basis of $\Delta\mathcal{L}$ was further from the optimum than for case B, but still achieved worthwhile PRIALs of more than 40% for MUV. Corresponding values for PEP were somewhat lower, but not too disconcertingly, especially as constellations of population values as for case A are uncommon in practice. Again, depending on the penalty, a fixed value of ψ resulted in substantial improvement in estimates of $\mathbf{\Sigma}_G$ for both cases, but with more fluctuations than the likelihood based choice.

With penalties designed to shrink both $\mathbf{\Sigma}_G$ and $\mathbf{\Sigma}_E$, a similar pattern of improvements was observed for estimates of $\mathbf{\Sigma}_E$ though PRIALs obtained were considerably lower, ranging from 14 to 28% for case A and 10 to 20 % for case B when selecting the tuning factor on the basis of $\Delta\mathcal{L}$. Corresponding values for estimates of $\mathbf{\Sigma}_P$ were small throughout, ranging from 0 to 3%.

DISCUSSION

Estimates of covariance components from multivariate analyses comprised of more than a few traits are subject to substantial sampling variation. Regularization can reduce this dramatically and thus yield estimates closer to the population values and, ultimately, result in better predictions of genetic merit and increased response to selection, in particular if weights for selection indices need

Table 1. Percentage reduction in average loss (PRIAL) and corresponding mean change in log likelihood ($\log \mathcal{L}$) for estimates of the genetic covariance matrix imposing different penalties.

Case	Value	Tune	$\mathcal{P}_\lambda$			$\mathcal{P}_\lambda^{\ell 2}$		$\mathcal{P}_\Sigma$		$\mathcal{P}_R$	
			MUV	PHS	BEN	MUV	PHS	MUV	PHS	MUV	PHS
A	PRIAL	$\Delta\mathcal{L}$	46	32	38	45	43	37	38	40	41
		mild	18	8	21	20	52	37	51	19	45
	$\log \mathcal{L}$	$\Delta\mathcal{L}$	-1.88	-1.89	-1.91	-1.88	-1.88	-1.83	-1.86	-1.89	-1.86
		mild	-0.19	-0.15	-0.50	-0.24	-3.02	-1.71	-11.58	-0.27	-2.34
B	PRIAL	$\Delta\mathcal{L}$	53	41	46	53	53	51	52	55	53
		mild	47	19	40	48	60	55	45	44	58
	$\log \mathcal{L}$	$\Delta\mathcal{L}$	-1.87	-1.88	-1.91	-1.86	-1.88	-1.84	-1.86	-1.87	-1.88
		mild	-0.81	-0.20	-1.02	-0.92	-7.12	-3.96	-13.88	-0.57	-3.16

to be derived from these estimates. REML estimation subject to a penalty provides such improved estimates but, while desirable, can be computationally demanding and accurate estimation of the optimum tuning factor remains problematic. Hence we propose a two-step procedure as alternative, in which standard, unpenalized estimates are modified post-estimation applying a mild penalty.

A suitable choice of the tuning factor may be based on limiting the change in $\log \mathcal{L}$ from the maximum to a relatively small value. For a limit corresponding to the significance threshold in a likelihood ratio test for one parameter, results showed reductions in loss in the range of 30 to 50%, and, except for a penalty on canonical eigenvalues on the original scale ($\mathcal{P}_\lambda$), differences to values for a penalized multivariate analyses were small. For an animal model with only two sources of variation, choosing a paternal half-sib design as pseudo-pedigree structure generally performed best.

REML estimates of covariance components are biased due to constraints on the parameter space. Improvements in estimates due to penalization generally come at the price of additional bias. While a mild penalty may not fully exploit the scope for reducing losses, the impact of penalization is not linear and such strategy can thus achieve a substantial proportion of the potential benefits at little effort. In addition, mild penalization will keep the extra bias created small and often result in estimates of individual components barely changed from unpenalized values.

CONCLUSIONS

Post-estimation penalization of multivariate estimates of covariance matrices using a likelihood approach teamed with a mild penalty can yield substantially improved estimates. It is recommended for routine analyses involving more than a few traits.

ACKNOWLEDGEMENTS

This work was supported by Meat and Livestock Australia under grant B.BFG.0050.

REFERENCES

Bondari K., Willham R.L. and Freeman A.E. (1978) *J. Anim. Sci.* **47**:358.
Hayes J.F. and Hill W.G. (1981) *Biometrics* **37**:483.
Meyer K. (2011a) *Proc. Ass. Advan. Anim. Breed. Genet.* **19**:83.
Meyer K. (2011b) *Genet. Sel. Evol.* **43**:39.
Meyer K. (2013) *J. Anim. Breed. Genet.* **130**:270.
Meyer K. and Kirkpatrick M. (2010) *Genetics* **185**:1097.

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:428-431

PENALIZED ESTIMATION OF COVARIANCE MATRICES WITH FLEXIBLE AMOUNTS OF SHRINKAGE

Karin Meyer

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

SUMMARY

Penalized maximum likelihood estimation has been advocated for its capability to yield substantially improved estimates of covariance matrices, but so far only cases with equal numbers of records have been considered. We show that a generalization of the inverse Wishart distribution can be utilised to derive penalties which allow for differential penalization for different blocks of the matrices to be estimated. However, this requires multiple tuning factors to be determined and thus can increase computational requirements markedly. Simulation results are presented which indicate that the additional gains obtainable for estimates of genetic covariance components – over and above those from a simple, non-differential scheme – are moderate, even if numbers of records for different traits differ by orders of magnitude.

INTRODUCTION

Estimation of covariance components by restricted maximum likelihood (REML) subject to a penalty borrowing strength from the phenotypic covariance matrix, has been shown to yield estimates closer to the population values than their 'standard' unpenalized counterparts (Meyer 2011). So far, studies to evaluate the properties of penalized estimates only considered equal numbers of measurements for all traits. In practice, however, we may have subgroups of traits with greatly differing numbers of records. A particular type of penalty – motivated by Bayesian estimation – is given by minus the logarithmic value of the density of an inverse Wishart (IW) distribution added to the REML log likelihood. Using the phenotypic covariance as scale matrix, this shrinks individual, e.g. genetic, matrices towards the former (Meyer *et al.* 2011). A drawback of this 'prior' is the rigidity imposed by a single parameter for the degrees of freedom. Hence an extension to a generalized inverse Wishart (GIW) distribution (Brown 2006) has been proposed as a more flexible alternative. Munilla and Cantet (2012) give details together with an application to account for differential uncertainty in genetic parameters in a Bayesian analysis.

This paper describes a penalty based on the GIW distribution and presents a simulation study examining the effect of this penalty on sampling properties of penalized REML estimates of covariance matrices for unequal numbers of records between traits.

THE GIW PENALTY

Consider q traits with covariance matrix $\mathbf{\Sigma}$, ordered so that 1 to q_1 are the subset of traits measured on a group of individuals without records for traits $q_1 + 1$ to q, while a second group has all q traits measured. Assume $\mathbf{\Sigma}$ has an IW distribution with scale matrix $\mathbf{\Omega}$. This gives 'whole matrix' penalty

$$\mathcal{P} = C \log|\mathbf{\Sigma}| + \mathrm{tr}\left(\mathbf{\Sigma}^{-1}\mathbf{\Omega}\right) \tag{1}$$

with $C \approx 1$ a constant depending on q and the degrees of freedom. Partition $\mathbf{\Sigma}$ and $\mathbf{\Omega}$ into

$$\mathbf{\Sigma} = \begin{pmatrix} \mathbf{\Sigma}_{11} & \mathbf{\Sigma}_{12} \\ \mathbf{\Sigma}_{21} & \mathbf{\Sigma}_{22} \end{pmatrix} \quad \text{and} \quad \mathbf{\Omega} = \begin{pmatrix} \mathbf{\Omega}_{11} & \mathbf{\Omega}_{12} \\ \mathbf{\Omega}_{21} & \mathbf{\Omega}_{22} \end{pmatrix}$$

according to the subsets of traits. $\mathbf{\Sigma}_{11}$ is independent of $\mathbf{\Sigma}_{22.1} = \mathbf{\Sigma}_{22} - \mathbf{\Sigma}_{21}\mathbf{\Sigma}_{11}^{-1}\mathbf{\Sigma}_{12}$ and $\mathbf{\Sigma}_{11}^{-1}\mathbf{\Sigma}_{12}$ and has an IW distribution. The penalty for the first sub-matrix is then simply obtained ignoring

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

the remaining traits. Similarly, for $\boldsymbol{\Omega}_{22.1} = \boldsymbol{\Omega}_{22} - \boldsymbol{\Omega}_{21}\boldsymbol{\Omega}_{11}^{-1}\boldsymbol{\Omega}_{12}$ and C_{ij} similar to C the conditional distribution for the second block given the first is IW, which yields penalties

$$\mathcal{P}_{11} = C_{11}\log|\boldsymbol{\Sigma}_{11}| + \operatorname{tr}\left(\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Omega}_{11}\right) \qquad \text{and} \qquad \mathcal{P}_{22} = C_{22}\log|\boldsymbol{\Sigma}_{22.1}| + \operatorname{tr}\left(\boldsymbol{\Sigma}_{22.1}^{-1}\boldsymbol{\Omega}_{22.1}\right) \tag{2}$$

Expanding (1) in terms of the submatrices and subtracting $\mathcal{P}_{11}$ and $\mathcal{P}_{22}$ gives the penalty for the remaining covariance components (assuming $C_{11} = C_{22} = C$)

$$\mathcal{P}_{12} = \operatorname{tr}\left(\boldsymbol{\Sigma}_{22.1}^{-1}\left[\left(\boldsymbol{\Omega}_{21}\boldsymbol{\Omega}_{11}^{-1} - \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\right)\boldsymbol{\Omega}_{12} + \left(\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Omega}_{11} - \boldsymbol{\Omega}_{21}\right)\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{12}\right]\right) \tag{3}$$

It can be shown that $\mathcal{P}_{12}$ is proportional to minus the log density for $\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{12}$ assumed to have a matrix-variate Normal distribution. These arguments are readily generalized to more subsets of traits; Brown (2006) summarizes the GIW as a series of sequential, conditional distributions.

MATERIAL AND METHODS

Data for 14 traits were simulated by sampling genetic and residual effects from appropriate multivariate Normal distributions for a paternal half-sib design and different combinations of population heritabilities and correlations. For case A and B, all heritabilities were assumed equal, 0.4 and 0.2, respectively. For case C, values for traits 1 to 14 were 2×0.6, 0.55, 2×0.5, 0.45, 2×0.4, 0.35 2×0.3, 0.25 and 2×0.2. For scenario I, all correlations were assumed to be zero and all phenotypic variances were set to 1. For II, all genetic and residual correlations were equal, 0.5 and 0.2, respectively, and for III correlations between traits i and j were set to $0.5^{|i-j|}$ (genetic) and $0.2^{|i-j|}$ (residual), while phenotypic variances were set to $\operatorname{mod}(i,3)+1$. This yielded nine sets of population parameters, referred to as A-I to C-III henceforth. Records for all traits were obtained for $s_1 = 400$ sires with 10 progeny each. In addition, records for the first $q_1 = 3$, 5, 7, 9 and 11 traits only were sampled for $s_2 = 400$ or $s_2 = 2000$ sires with 20 progeny. A total of 500 replicates per case were carried out.

Analyses. For each replicate, REML estimates of genetic ($\boldsymbol{\Sigma}_G$) and residual ($\boldsymbol{\Sigma}_E$) covariance matrices were obtained subject to five types of penalty, involving up to three different tuning factors (ψ_i)

$$\mathcal{P}_a = \psi_1(\mathcal{P}_{22} + \mathcal{P}_{12} + \mathcal{P}_{11}) = \psi_1\mathcal{P} \qquad \mathcal{P}_c = \psi_1\mathcal{P}_{22} + \psi_2\mathcal{P}_{12}$$
$$\mathcal{P}_b = \psi_1(\mathcal{P}_{22} + \mathcal{P}_{12}) = \psi_1(\mathcal{P} - \mathcal{P}_{11}) \qquad \mathcal{P}_d = \mathcal{P}_e = \psi_1\mathcal{P}_{22} + \psi_2\mathcal{P}_{12} + \psi_3\mathcal{P}_{11}$$

and without penalization. Tuning factors were estimated by constructing matrices of means squares and cross-products corresponding to the data structure for the population parameters (which are unknown in practice), and maximizing the likelihood of estimates of $\boldsymbol{\Sigma}_G$ and $\boldsymbol{\Sigma}_E$ in these 'validation data'. This was done using a derivative-free search as implemented in routine `NEWUOA` (Powell 2008), maximizing with respect to $\log\psi_i$ to ensure that estimates were positive. Any estimates exceeding 1,000 were set to this value. For $\mathcal{P}_e$ maximization was performed in two steps by first estimating ψ_3, considering records for traits 1 to q_1 only, and then (jointly) estimating ψ_1 and ψ_2 for ψ_3 fixed at its estimate from step 1.

Summary statistics. The deviation of estimated covariance matrices ($\hat{\boldsymbol{\Sigma}}$) from their population values ($\boldsymbol{\Sigma}$) was evaluated as the entropy loss (L_1) and, with $\bar{L}_1(\cdot)$ denoting the mean over replicates and $\hat{\boldsymbol{\Sigma}}^{\psi}$ the estimate for a tuning factor of ψ, the percent reduction in average loss (PRIAL)

$$L_1(\boldsymbol{\Sigma}, \hat{\boldsymbol{\Sigma}}) = \operatorname{tr}(\boldsymbol{\Sigma}^{-1}\hat{\boldsymbol{\Sigma}}) - \log|\boldsymbol{\Sigma}^{-1}\hat{\boldsymbol{\Sigma}}| - q \qquad \text{PRIAL} = 100\left[1 - \bar{L}_1\left(\boldsymbol{\Sigma}, \hat{\boldsymbol{\Sigma}}^{\psi}\right)/\bar{L}_1\left(\boldsymbol{\Sigma}, \hat{\boldsymbol{\Sigma}}^{0}\right)\right]$$

In addition, the deviation in likelihood from the (unpenalized) maximum ($\Delta\log\mathcal{L}$) was calculated.

RESULTS

Figure 1 shows the distribution of losses in estimates of $\boldsymbol{\Sigma}_G$ for different values of q_1 for one of the cases examined (C-III for $s_2 = 2000$). Patterns for other constellations were similar. As to be expected, losses in unpenalized estimates decreased substantially as the number of traits (q_1) with

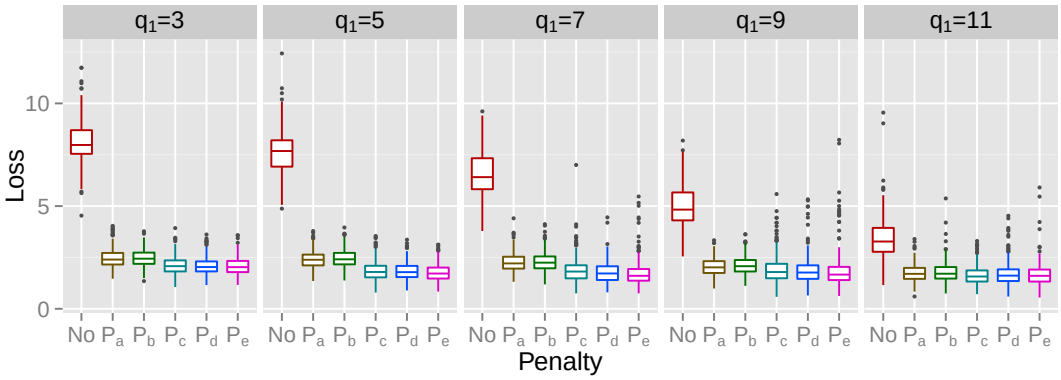


Figure 1. Distribution of loss in estimates of covariance matrices for case C-III (s_2 = 2000)

many records increased. Penalization reduced losses in $\hat{\mathbf{\Sigma}}_G$ and their spread throughout with relatively small differences between types of penalty, especially for larger values of q_1.

Means for tuning factors (across population values and q_1), PRIALs and $\Delta \log \mathcal{L}$ (across population values) for both sample sizes are summarized in Figures 2 and 3. With tuning factors obtained by exploiting knowledge of the population values, mean PRIALs were high, especially for small proportions of traits with many records. The number of sire families with records only for the first q_1 traits appeared unimportant until these represented at least half the traits. For $\hat{\mathbf{\Sigma}}_G$, differences in mean PRIAL between penalties $\mathcal{P}_a$ and $\mathcal{P}_e$ increased with q_1, amounting to 13 to 20% Corresponding values for $\hat{\mathbf{\Sigma}}_E$ ranged from 22 to 30% for s_2 = 400 and 13 to 25% for s_2 = 2000. Whilst only $\hat{\mathbf{\Sigma}}_G$ was penalized directly, previous studies found marked associated improvements in $\hat{\mathbf{\Sigma}}_E$, due to strong negative sampling correlations (Meyer 2011). For unequal numbers of records, the effect of penalties involving a single tuning factor ($\mathcal{P}_a$ and $\mathcal{P}_b$) on $\hat{\mathbf{\Sigma}}_E$ for low numbers of q_1 were substantially less than those with multiple factors. Again there was comparatively little difference between $\mathcal{P}_c$, $\mathcal{P}_d$ and $\mathcal{P}_e$, suggesting that the main benefits were obtained by penalizing submatrices $\mathbf{\Sigma}_{22}$ and $\mathbf{\Sigma}_{12}$ differentially. Higher PRIALs for $\mathcal{P}_c$, $\mathcal{P}_d$ and $\mathcal{P}_e$ were accompanied by larger changes in likelihood. This was due to much more stringent penalization of block $\mathbf{\Sigma}_{12}$. Similarly, estimating ψ_3 separately to ψ_1 and ψ_2 ($\mathcal{P}_e$) resulted in higher estimates of ψ_3 and more improvement in $\hat{\mathbf{\Sigma}}_G$ than joint estimation ($\mathcal{P}_d$), suggesting that the three-dimensional search had some problems.

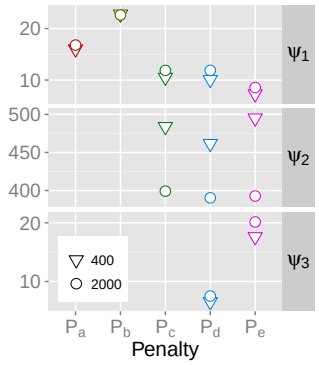


Figure 2. Tuning factors (s_2 = 400 and 2000)

DISCUSSION

It has been shown that a generalization of the inverse Wishart distribution can be utilised to derive a penalty for penalized REML estimation of covariance components which allows differential shrinkage to be applied to different blocks of the covariance matrices to be estimated. A simulation study has been used to demonstrate that this can improve estimates more than non-differential penalties when there are substantially different numbers of records for different subsets of traits, especially those of residual covariances. However, this requires separate tuning factors to be determined. While not shown here, this can increase the complexity of analysis and computational burden markedly. The differential penalty employed utilizes sequential, conditional distributions of subsets of traits. Results suggest that estimation of tuning factors in an analogous fashion is advantageous.

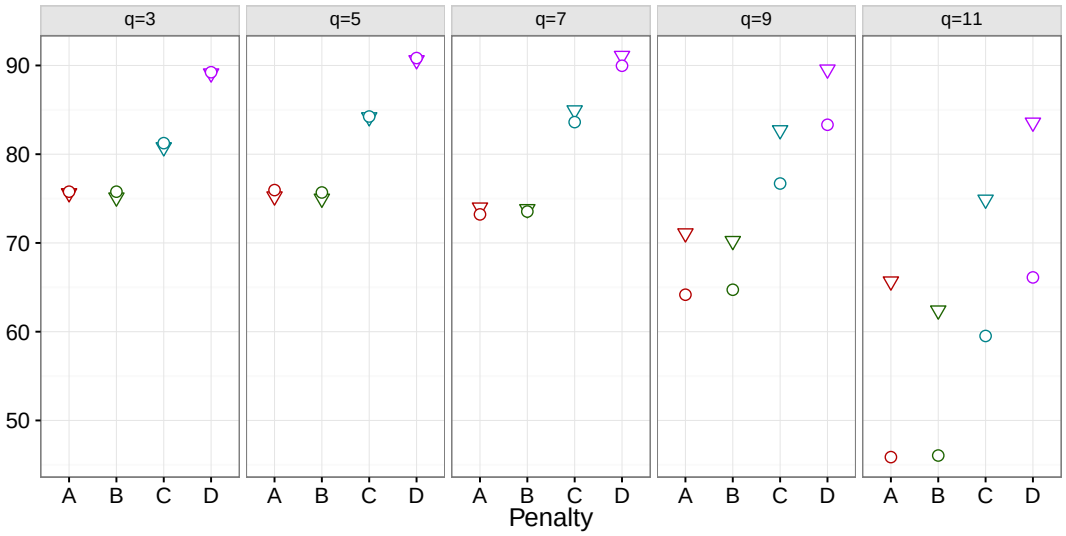


Figure 3. Mean percent reduction in average loss (PRIAL) for estimates of covariance matrices and corresponding change in log likelihood ($\Delta\log\mathcal{L}$) for $s_2 = 400$ ($\triangledown$) and $s_2 = 2000$ ($\circ$)

Simulations results given represent a 'best possible' scenario as tuning factors were obtained utilizing the population values. Even so, additional improvements in estimates of genetic covariances, over and above those achieved by a simpler, non-differential penalty ($\mathcal{P}_a$), were moderate. Additional investigations (not shown) indicated that these decreased with the size of the subset of data with records for all traits. Somewhat surprisingly, benefits of penalties $\mathcal{P}_c$, $\mathcal{P}_d$ and $\mathcal{P}_e$ were most pronounced for the residual covariances. Whether in practice the extra gains possible warrant the additional effort required depends on how well multiple tuning factors can be estimated from data at hand. Future work should address this question. In the meantime, it is reassuring that the simple, non-differential penalty appears to be fairly robust against marked differences in information available for different traits, and can achieve a substantial proportion of the improvements feasible.

CONCLUSIONS

Differential shrinkage of different blocks of covariance matrices to be estimated is feasible, employing a penalty based on the generalised inverse Wishart distribution. However, this requires considerable effort to determine appropriate, multiple tuning factors whilst additional improvements in estimates of genetic covariances achievable appear quite moderate.

ACKNOWLEDGEMENTS

This work was supported by Meat and Livestock Australia under grant B.BFG.0050.

REFERENCES

Brown P.J. (2006) In *Encyclopedia of Environmetrics*. John Wiley and Sons, Ltd, pp. 1079–1082.
Meyer K. (2011) *Genet. Sel. Evol.* **43**:39.
Meyer K., Kirkpatrick M. and Gianola D. (2011) *Proc. Ass. Advan. Anim. Breed. Genet.* **19**:87.
Munilla S. and Cantet R.J.C. (2012) *J. Anim. Breed. Genet.* **129**:173.
Powell M. (2008) *IMA J. Numer. Anal.* **28**:649.

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:432-434

AN APPROACH TO CONNECT MULTI-TRAIT MIXED MODEL AND PRINCIPAL COMPONENT ANALYSIS FOR DESCRIBING VARIATION IN CARCASS QUALITY OF CROSSBRED CATTLE

H.R.Mirzaei[1,2] and W.S. Pitchford[2]

[1]Department of Animal Science, University of Zabol, Iran
[2]School of Animal and Veterinary Sciences, University of Adelaide, Roseworthy SA 5371 Australia

SUMMARY

A principal component analysis of the 4×4 sire, maternal, management and environmental (co)variance matrices derived from a multi-trait sire model was conducted to describe variability in four economically important carcass traits. Carcass weight (HCWt), P8 fat (P8), eye muscle area (EMA) and intramuscular fat (IMF) collected from 1144 heifers and steers calves from seven sire breeds: Angus, Belgian Blue, Hereford, Jersey, Limousin, South Devon and Wagyu, born over a 4-year period. The first two principal components (PC1, PC2) accounted for 90% of the total variance in the considered variables, except for the maternal component, where PC1 and PC2 accounted for 83% of the total variance. The largest and the least variations attributed to the management (99%) and maternal (83%) components, respectively. Sire and environment components showed similar patterns of eigenvector coefficients for the first two vectors. The first and second eigenvectors have large loadings for P8 fat and IMF, respectively. The third orthogonal vector had a large coefficient for the HCWt and EMA but not other traits. For the maternal component, which is a small component of overall variation, P8 fat in contrast to IMF had a significant relationship with the PC1. PC1 could be defined as a fat distribution component. PC2 respects mean values for carcass traits with less attention to EMA, presenting market suitability. For management component as the largest component of overall variation, PC1 could be interpreted as a weighted mean with much more emphasis on the IMF. PC2 accounting, for 25.78% of the total variance, indicated a major contrast between P8 and IMF, consequently it can be interpreted as a fat distribution component.

Keywords: Principal component analysis, Multivariate, Sire model, Carcass traits

INTRODUCTION

In two components papers by Mirzaei *et al.* (2009) multi-trait mixed model and principal component analysis (PCA) have been conducted to examine of variation in carcass traits. PCA of raw data is a useful exploratory tool but lacks adjustment for fixed effects (e.g., breed and sex). Thus, the correlation structure and variation involving these traits should be regarded with caution. Hence, it is worthwhile conducting principal component analysis on estimated (co)variance structures for sire, maternal, management and environment obtained from multi-trait mixed model with the hope of obtaining quality information. The aim of this paper is to investigate the decomposition of a square matrix (4×4) of the sire, maternal, management and environment into eigenvalues and eigenvectors (PCA) for the four carcass traits obtained from multi-trait mixed model. The determination of the eigenvectors and eigenvalues of those components aids in understanding the important sources of variation in carcass quality traits and to realize that how fitting fixed factors affect linear combinations of the original variables.

MATERIALS AND METHODS

Data were obtained from the Southern Crossbreeding Project which was designed for meeting a range of market specifications. Mature Hereford cows (581) were mated to semen from 97 sires

from seven breeds (Angus, Belgian Blue, Hereford, Jersey, Limousin, South Devon and Wagyu), resulting in 1144 live calves born over 4 years (1994-97). All statistical analyses were conducted using PROC PRINCOMP (SAS Institute Inc. 1999).

A multi-variate sire model was fitted using ASREML (Gilmour *et al.*, 2000), estimating multi-trait (co)variance components including genetic and non-genetic parameters of carcass quality traits (Table 1). Variation in the four carcass quality traits (HCWt, P8, EMA and IMF) was considered in terms of the same fixed and random factors as the growth models. Principal component analysis of the 4×4 (co)variance matrices derived from the above multi-trait carcass model for sire (¼ additive genetic as effectively nested within breed), maternal (¼ additive genetic + maternal genetic + dam permanent environmental effect), management group (combination of sex, year and pre- and post-weaning cattle management group) and environmental (residual). The model is : $\mathbf{PC_n = X\tau + Zu + e}$ where $\boldsymbol{\tau}$ is the vector of fixed effects, $\mathbf{u}$= vector of random effect, $\mathbf{e}$ = vector of random residual effect (temporary environmental effect or measurement error), $\mathbf{NID\ (0, \sigma^2)}$.

RESULTS

PCA results herein permit a description of the simultaneous or multivariate patterns of covariation among the various carcass quality traits within each variance components. These eigenvectors were orthogonally rotated to facilitate more interpretable results, i.e. statistically independent vectors exhibiting either high or low eigenvector coefficients or few intermediate values. The four patterns of co-variation (eigenvectors) summarize the common information among these four carcass quality traits. In general, sire correlations between carcass traits were variable, dam and management correlations were high and environmental (residual) correlations were low.

The first two principal components accounted for the major proportions of the total variation in four components (83-99%, Table 1).

For sire, management and environment (residual), PC1 was related to fatness with the eigenvector was positive for both P8 fat depth and IMF (Figure 1). PC2 was related to fat distribution with opposite weightings for P8 fat and IMF. The maternal component was quite different in that PC1 could be described as fat distribution and PC2 as growth since the Eigen vector had positive weightings for all four carcass traits. Coefficients for the remaining traits were small and contribute little to those eigenvectors.

DISCUSSION

Overall, the first two principal components accounted for much variation in carcass traits and quite obviously correlated with fat traits (P8 and IMF) reflecting relatively high correlations between the four carcass traits. While the results herein are scientifically interesting, for four traits it is difficult to see large benefit in using principal component analysis. However, it could be beneficial for summarizing larger numbers of traits and potentially for describing bull “types” which is common in stud sale catalogues.

REFERENCES

Gilmour. A.R., Bulls. B.R., Welham. S.J. and Thompson. R. (2000). Orange Agricultural Institute. Forest Road. Orange 2800 NSW Australia.

Mirzaei. H.R., Verbyla. A.P., Deland. M.P. and Pitchford. W.S. (2009) *Pak J Biol Sci.* **12:** 222.

SAS, (1999). SAS/STAT User’s Guide. Fourth edition, Cary, N.C. : SAS Institute

Burrow. H.M. (2012). *Animal.* **6:** 729.

Table 1. The eigenvalues and eigenvectors (PCs) of the sire, maternal, management and environmental correlation matrices for carcass traits

	PC1	PC2	PC3	PC4
Sire				
HCWt	-0.05	-0.03	**0.79**	**0.61**
P8	**0.94**	-0.33	0.08	-0.05
EMA	-0.13	-0.11	**0.60**	**-0.78**
IMF	0.32	**0.94**	0.12	-0.09
%variance	65	27	7	1
Maternal				
HCWt	0.12	**0.54**	**0.50**	**0.66**
P8	**0.59**	**0.61**	**-0.46**	-0.26
EMA	0.15	0.12	**0.72**	**-0.67**
IMF	**-0.79**	**0.56**	-0.13	-0.22
%variance	51	32	13	4
Management				
HCWt	0.16	-0.13	-0.05	**0.98**
P8	0.22	**-0.93**	-0.25	-0.17
EMA	0.15	-0.22	**0.96**	0.00
IMF	**0.95**	0.27	-0.09	-0.13
%variance	74	26	0	0
Environment				
HCWt	0.07	0.01	-0.42	**0.91**
P8	**0.94**	-0.34	0.04	-0.05
EMA	0.02	-0.01	**-0.91**	-0.42
IMF	0.34	**0.94**	0.01	-0.03
%variance	53	39	7	2

Values in bold indicate high loading values

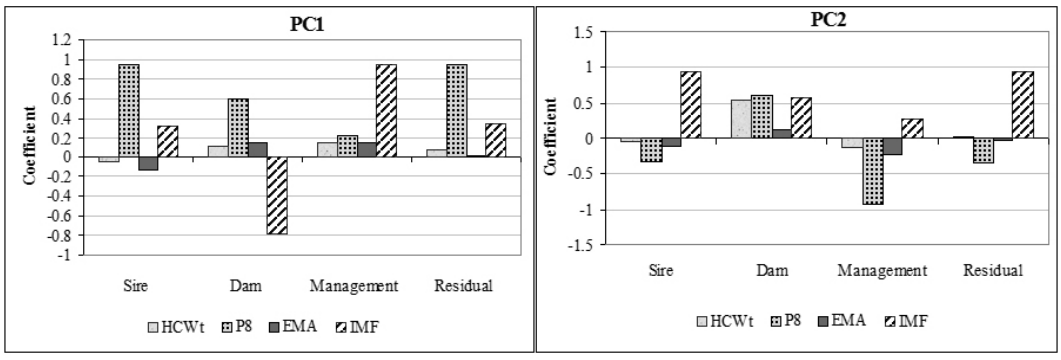


Figure 1. Loading comparisons of the first two principal components

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:435-438

THE MILK FATTY ACID COMPOSITION AND CONJUGATED LINOLEIC ACID CONTENT OF JERSEY AND FLECKVIEH X JERSEY COW MILK IN A PASTURE BASED FEEDING SYSTEM

C.J.C. Muller[1], B. Sasanti[2,3], S. Abel[3] and A. Schmulian[2]

[1]Research & Technology Development Services, Directorate: Animal Sciences, Private bag X1, Elsenburg 7607, South Africa; [2]Cape Peninsula University of Technology, Private bag X8, Wellington 7654, South Africa; [3]PROMEC Unit, Medical Research Council, P.O. Box 19070, Tygerberg 7505, South Africa

SUMMARY

A number of fatty acids like omega-3, omega-6 and conjugated linoleic acid (CLA) present in the milk from dairy cows are considered beneficial nutrients for humans. The aim of the study was to compare the milk fatty acid (FA) content, particularly the CLA, omega-3, omega-6 FA content of the milk fat of Jersey (J) and Fleckvieh x Jersey (FxJ) cows in a pasture-based feeding system. All cows were fed the same diet consisting of kikuyu-ryegrass pasture supplemented with a standard concentrate mixture at 7kg per cow per day. Four to five milk samples were collected every five weeks from 10 days after calving (DIM) up to 175 DIM. In addition, two further samples were collected every five weeks from 240 DIM to the end of the lactation. All milk samples were collected at the evening and the next morning's milking and pooled for each cow. Samples were kept frozen at -20°C until laboratory analysis by gas liquid chromatography. Thirty six FAs were detected and concentration levels determined. All milk samples (128 for J and 239 for FxJ) were used to compare breeds for FA content. Total omega-6 and total CLA differed ($P<0.01$) between breeds being 1.571±0.040 and 1.754±0.029 and 0.630±0.023 and 0.740±0.018 g per 100 g milk fat for J and FxJ cows, respectively. For both breeds the CLA content of the milk fat showed a curvilinear increase with lactation stage possibly indicating a standard sampling time to determine cow differences for genetic merit analysis. Further studies are required to determine the milk FA composition in different milk products.

INTRODUCTION

The fat component of milk has for many years been regarded as unhealthy because of its affect on heart diseases in humans (Salter 2005). Health practitioners recommend that the fat content of the human diet be reduced for protection against cardiovascular diseases and some forms of cancer. This has resulted in the popularity of fat free and low fat milk (0 and 2% fat respectively) as well as low fat cheese and yoghurt products. However, the fat in milk is made up by a large number of saturated and unsaturated FAs each contributing differently to the health of people. Bovine milk is increasingly being recognized as an important source of energy, high-quality protein, and essential minerals and vitamins (Heaney 2000 and Neuman *et al*. 2003). The fat in milk has recently acquired an improved status as new research has shown that some FAs have a beneficial effect on the health status of people. It is especially omega-3 FA and conjugated linoleic acid (CLA) that have anticarcinogenic, antidiabetic and antidipogenic effects. The amount of CLA in cows' milk is affected mostly by their diet and healthy FAs increase when cows are on pasture (Mitchell and McLeod 2008) or when feeds such as extracted soy beans and cottonseed are fed (Collomb *et al*. 2006). While diet has a major influence on milk fat CLA (Chilliard *et al*. 2001), the effects of factors such as breed, stage of lactation and parity on the CLA content in milk fat have received little attention (Kelsey *et al*. 2003). Some studies indicated breed differences in CLA content (Lawless *et al*. 1999) with Montbéliarde having 13% greater CLA content in milk fat

in comparison to Irish Holstein/Friesian, Dutch Holstein/Friesian and Normande. Large differences are observed among individual cows receiving the same diet (Kelsey *et al*. 2003). Crossbreeding is a means to overcome some breeding problems like fertility and longevity in some dairy breeds (Funk 2006). Recently, attention has been given towards using dual-purpose breeds in crossbreeding programmes to increase the beef production of crossbred animals while maintaining the milk yield of cows. The Fleckvieh, a Simmental-derived breed from Germany is one such breed. A study in Canada has shown that the milk from Fleckvieh x Holstein cows produced more CLA than purebred Holsteins under similar feeding conditions (Patrick *et al*. 2000, Lock & Bauman 2004). The aim of the paper was to compare the milk FA content of the milk of J and FxJ cows in a pasture-based feeding system.

MATERIAL AND METHODS

Location and Animals. This paper was based on an on-going breed comparison at the Elsenburg Research Farm of the Western Cape Department of Agriculture (Muller *et al*. 2009). 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 hot summers. To create two comparative pure- and crossbred dairy herds, all available J cows (n=46) were divided into two groups according to estimated breeding value for milk yield. Groups were randomly allocated to be inseminated by J or F bulls. During the following lactation period cows were inseminated with the alternative sire breed. The progeny born from the J and F sires were further inseminated with the same breed. Subsequently, the production performance of J (n=56) and FxJ (n=64) cows and their progeny was compared in a partly pasture-based feeding system. This consisted of mostly kikuyu pasture supplemented with a commercial concentrate mixture at 7 kg per cow per day regardless of milk yield and lactation stage. During winter the pasture was supplemented with a mixture of oats and lucerne hay. Fresh drinking water was freely available at all times.

Milk sample collection and analysis. Milk samples for FA analysis were collected and recorded every five weeks according to milk recording procedures. At each milk recording event, milk samples were collected from cows of both breeds. Milk was sampled from 10 days after calving (DIM) to about 175 DIM (milk tests 1 to 5) and thereafter from 240 DIM (milk tests 7 to 8). Milk samples were collected at the evening and next morning's milking session and combined. Milk samples were kept frozen at -20°C until laboratory analysis. Fatty acid composition of milk samples was obtained by gas liquid chromatography at the PROMEC Unit of the Medical Research Council. Thirty six FA were detected and concentration levels determined.

Statistical analyses. All milk samples (128 for J and 239 for FxJ) were analysed for 36 FAs. In the current study only the major FAs were presented. FAs were compared between breeds by analysis of variance using samples of all cows within breed as replicates using the GLM procedure (SAS Institute Inc.).

RESULTS AND DISCUSSION

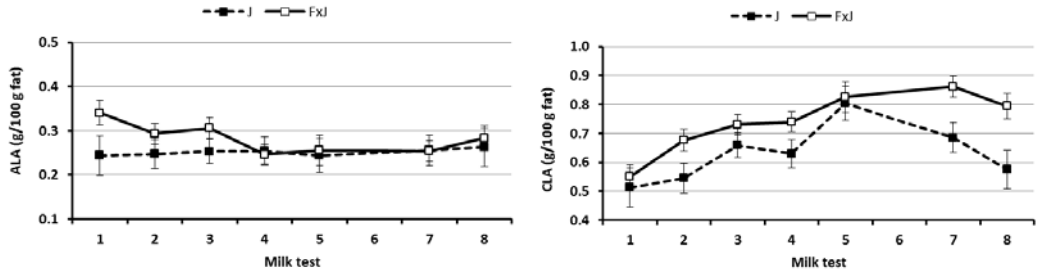
Some FAs differed ($P<0.05$) between breeds being 1.533±0.032 and 1.664±0.025 for omega-6, and 0.621±0.021 and 0.725±0.015 g/100 g fat for total CLA content for J and FxJ cows, respectively (Table 1). The specific FAs trans 18:1, 18:2n-6 (LA, linoleic acid) and the main CLA isomer, C9 T11 18:2, also differed ($P<0.05$) between breeds. No significant difference in the omega-3 FA, α-linolenic acid (ALA, 18:3n-3) was observed. Maurice-Van Eijndhoven *et al*. (2011) compared 4 cattle breeds in the Netherlands, showing breed differences although results were confounded with breed-herd effects as only one breed per farm was sampled. Grazing- or non-grazing-based feeding systems largely influences milk FA composition (Palmquist *et al*. 1993). Kelsey *et al*. (2003) compared Holstein and Brown Swiss cows being fed a single diet and

milk sampled on the same day to avoid confounding effects of diets and season. However, only minor differences between these two breeds were found. The CLA content of milk fat varied over threefold among individual cows. In the present study cows from both breeds were under similar feeding and management conditions.

Table 1. The mean±se fatty acid content (g/100 g fat) of the milk Jersey (J) and Fleckvieh x Jersey (FxJ) cows in a partly pasture-based feeding system (LA = linoleic acid; CLA = conjugated linoleic acid; ALA = α-linolenic acid)

	Breeds		P-values		
Fatty acids	J	FxJ	Breeds	Test	Breed x Test
Trans 18:1	0.918±0.026	1.018±0.020	0.003	0.019	0.723
LA n-6, 18:2	1.356±0.036	1.509±0.027	0.001	0.003	0.277
CLA (C9, T11, 18:2)	0.589±0.022	0.690±0.017	0.001	0.001	0.376
ALA n-3, 18:3	0.252±0.014	0.283±0.011	0.083	0.807	0.697
Total n-6	1.571±0.040	1.754±0.029	0.001	0.007	0.302
Total n-3	0.314±0.016	0.350±0.012	0.070	0.859	0.554
Ratio n-6/n-3	5.517±0.172	5.590±0.130	0.738	0.085	0.879
Total CLA	0.630±0.023	0.740±0.018	0.001	0.001	0.456

The results of milk recording test as affected by lactation stage or DIM on the content of ALA and total CLA is presented in Figure 1. While the level of ALA was not affected ($P>0.05$) by milk test event based on increasing DIM, for both breeds the CLA content in the milk fat increased potentially following a curvilinear trend ($R^2 = 0.74$ and $R^2 = 0.88$ for J and FxJ, respectively). This would suggest that for CLA, a standard sampling time should be considered to determine cow differences for genetic merit analysis or that results should be adjusted for lactation stage or DIM. The CLA content of both J and FxJ milk increased ($P<0.05$) by more than 40% from early in the lactation (<40 DIM) to later in lactation (>140 DIM). Similar trends were not observed for other FAs. However, Kelsey *et al*. (2003) found that lactation stage (DIM) had little effect (<2.0% of the total variation) on the CLA content of the milk fat of Holstein cows consuming a total mixed ration. Auldist *et al*. (1998) also found a small increase from 7.9 mg/g in early lactation (~30 DIM) to 9.7 mg/g FA in late lactation (~210 DIM). According to Stanton *et al*. (1997) lactation stage had no effect on CLA levels in milk fat, however, these studies were limited in scope, i.e. 36 cows ranging from 12 to 93 DIM. Frelich *et al*. (2009) found significant differences ($P<0.05$) between farms in the concentration of five FAs while 16 FAs of milk fat differed ($P<0.01$) between the indoor and the grazing period indicating the effect of pasture on FA content. The content of long-chain (>C16), mono- and poly-unsaturated FAs and CLA in the milk fat was higher in the grazing period. These results indicated a positive influence of seasonal grazing on the FA profile of cow milk fat as regards to its potential health effects for consumers (Frelich *et al*. 2009).



(a) (b)

Figure 1. The (a) ALA and (b) total CLA (b) content of the milk of Jersey (■) and Fleckvieh x Jersey (□) cows as affected by milk test

CONCLUSION

Some FAs differed between breeds although not all differences were significant. To demonstrate breed differences requires a significant number of animals from each breed. Milk test combined within DIM as per standard milk recording affected the CLA content of both J and FxJ milk which increased by more than 40% from early- to mid-lactation. Further studies are required to determine the FA composition in different milk products.

REFERENCES

Auldist M.J, Walsh B.J. and Thompson N.A. (1998) *J. Dairy Res*. **65**: 401.
Chilliard Y., Ferlay, A. and Doreau, M. (2001) *Livest. Prod. Sci*. **70**: 31.
Collomb M., Schmid A., Sieber R. and Wechsler D. (2006) *Intern. Dairy J*. **16**: 1347.
Frelich J., Šlachta M., Hanuš O., Špišcka J and Samková E. (2009) *Czech J. Anim. Sci*. **54**: 532.
Funk D.A. (2006) *J. Dairy Sci*. **89**: 1362.
Heaney R.P. (2003) *J. Am. Coll. Nutr*. **19**: 83S.
Kelsey J.A., Corl B.A., Collier R.J. and Bauman D.E. (2003) *J. Dairy Sci.* **86**: 2588.
Lawless F., Stanton C., L'Escop P., Devery R., Dillon P. and Murphy J.J. (1999) *Livest. Prod. Sci*. **62**: 4.
Lock A.I. and Bauman D.E. (2004) University of Wisconsin-Madison.
Maurice-Van Eijndhoven M.H.T., Hiemstra S.J. and Calus M.P.L. (2011) *J. Dairy Sci*. **94**: 1021.
Michell L. and McLeod S. (2008) *Minireview J*. **31**: 1253.
Muller C.J.C., Botha, J.A., Cloete, S.W.P. and Potgieter, J.P. (2009). *Proc. Assoc. Advmt. Anim. Breed. Genet*. **18**: 616.
Neuman C.G., Bwibo N.O., Murphy S.P., Sigman M., Whaley S., Allen L.H., Guthrie D., Weiss R.E. and Demment M.W. (2003) *J. Nutr*. **133**: 3941S.
Palmquist D.L., Beaulieu A.D. and Barbano, D.M. (1993) *J. Dairy Sci*. **76**: 1753.
Patrick R., O'Quinn J.J., Robert D. and Michael D.T. (2000) Kansas University. *www.journals.cambridge.org*
Salter A.M. (2005) *Proc. Cornell Nutritional Conf*. **2005**: 127.
Stanton C., Lawless F., Kjellmer G., Harrington D., Devery R., Connolly, J.F and Murphy J. (1997) *J. Food Sci*. **62**: 1083.

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:439-443

REPRODUCTIVE PERFORMANCE OF HOLSTEIN AND FLECKVIEH X HOLSTEIN HEIFERS AND COWS IN A TOTAL MIXED RATION FEEDING SYSTEM

C.J.C. Muller[1], J.P. Potgieter[2], S.W.P Cloete[1, 2] and J.A. Botha[1]

[1]Research & Technology Development Services, Directorate: Animal Sciences, Private bag X1, Elsenburg 7607, South Africa; [2]Department of Animal Sciences, University of Stellenbosch, Stellenbosch 7600, South Africa

SUMMARY

The fertility in dairy herds is becoming a major issue as several studies indicate a decline in the reproductive performance of dairy cows. Crossbreeding is regarded as a way to overcome this. In this paper, preliminary results of the reproductive performance of Holstein (H) and Fleckvieh x Holstein (FxH) heifers and lactating cows are presented. Heifers and cows were in an on-going breed comparison study in a total mixed ration (TMR) feeding system. Reproductive traits were derived from interval traits between birth and artificial insemination (AI) dates for heifers and calving and AI dates for cows. Means±sd for the interval from calving to first insemination (CFS) were 91±31 and 85±31 days (P=0.10) for H and FxH cows respectively. The proportion of cows having a first insemination within 80 days post partum (FS<80d), and confirmed pregnant within 100 days post partum (PD100d) for H and FxH cows was 0.41 and 0.51 (P=0.09) and 0.29 and 0.45 (P=0.01) respectively. Age at first service was lower and the proportion of heifers inseminated by 14 months of age was higher (P<0.05) in FxH in comparison to H heifers. While crossbred heifers and cows showed improved absolute reproduction compared to purebred animals, differences between breeds were not significant in all instances. As reproduction management strongly affects the performance of dairy cows, a larger data set and possibly records from other herds might reduce variability in fertility traits.

INTRODUCTION

Breeding and selection programmes in dairy herds in South Africa are mainly focused on the improvement of milk yield and conformation traits. Although the reproductive performance of dairy cows affects herd profitability, little emphasis is put on the genetic improvement of fertility. Cows may have repeated failed inseminations followed by hormonal treatment and eventually natural service. At best, non-pregnant cows are culled. In South African Holsteins, calving interval (CI) increased from 386 days in 1986 to 412 days in 2004 (Makgahlela 2008). Little local research has been done on the genetic improvement of fertility in dairy cows. Recently, Mostert *et al*. (2010) reported on the genetic parameters for CI for the four major dairy breeds in South Africa.

Because of increasingly poor reproductive performance in dairy herds, farmers are considering crossbreeding as a possible solution, as fertility traits are lowly heritable and should benefit from heterosis. While crossbreeding is applied in some herds, no research has been conducted locally to provide scientific support for it. Furthermore, crossbreeding in dairy herds is very contentious and regarded by breed societies as a poor way to overcome breeding and/or management problems. Crossbreeding is, nevertheless, increasingly being considered by global dairy producers because of their concerns about fertility, cow health and calf survival in the Holstein breed in particular (Funk 2006). Dairy breeds used mostly in crossbreeding studies include Jerseys and Ayrshires (Heins *et al*. 2008). McAllister (2002) compared Jersey x Holstein, Ayrshire x Holstein crossbreds while Touchberry (1992) compared Guernsey x Holstein crossbreds to pure Holsteins generally showing improved performances with crossbreds.

Dual-purpose breeds such as the Fleckvieh, a Simmental-derived breed, have not been seriously considered in crossbreeding programmes. True dual-purpose breeds have high milk

yields and milk quality traits while in some countries it is primarily used for beef production (Grogan *et al*. 2005). 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). Heins & Hansen (2012) showed that Normande x Holstein, Montbéliarde x Holstein cows had fewer (P<0.01) days to first breeding, better first-service conception rates (P<0.10), fewer days open (P<0.01) than Holstein cows. Recently Walsh *et al*. (2008) found that Holstein-Friesian cows had lower (P<0.05) submission rates and overall pregnancy rates in comparison to Montbéliarde, Normande, Norwegian Red, Montbéliarde x Holstein-Friesian and Normande x Holstein-Friesian cows. In some parts of Germany and Holland, crossbreeding of Holsteins is underway to improve beef production, fertility and productive life of dairy cows (Swalve, 2007). The aim of this paper is therefore to compare the reproductive performance of H and FxH heifers and cows in a total mixed ration feeding system.

MATERIAL AND METHODS

Location and Animals. This paper was based on an on-going breed comparison study being conducted at the Elsenburg Research Farm of the Western Cape Department of Agriculture (Muller *et al*. 2009). 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 and hot summers. Holstein (n=24) and FxH heifers (n=24) were initially sourced from a commercial H dairy herd and reared at Elsenburg until first calving. Subsequently, the production performance of these H and FxH cows and their progeny was compared in a total mixed ration (TMR) feeding system. Records from the Elsenburg Holstein herd (36 Holstein cows and 28 heifers) were also included in the study. Pure- and crossbred heifers were reared similarly to first calving. After calving, all cows received a TMR, providing 17% CP and 11 MJ ME/kg DM consisting of alfalfa hay, oat silage, wheat straw and a commercial concentrate mixture in open camps with fence-line feeding troughs. The TMR was fed twice a day at levels ensuring an *ad libitum* feed intake. Fresh drinking water was freely available at all times. Cows were machine-milked twice a day in a milking parlour, approximately 500 m from the open camps.

Data recording. Cows were routinely checked within the first 10 days after each calving and treated by a veterinarian for retained placentas and uterine infections. From 40 days after calving, a tail-marker was put on each cow to enable oestrus detection. Cows not showing signs of reproduction activity were treated according to a standard hormonal programme. Oetrus detection was done daily pre-milking. Cows were artificially inseminated (AI) from about 60 days after calving when showing standing oestrus. At 13 months of age, heifers were put in an AI-service group after being checked by a veterinarian for reproductive activity. Heifers were artificially inseminated when showing standing oestrus. The reproductive performance of heifers and cows was determined based on AI dates and the result of pregnancy diagnosis by rectal palpation by a veterinarian at least 45 days after the last insemination. Reproductive parameters determined for cows were the interval (number of days) from calving to first insemination (CFS), number of inseminations per conception (SPC), interval from calving to conception (DO), whether first insemination occurred within 80 days post partum (FS<80d), whether cows became pregnant from first insemination (PDFS) or within 100 (PD100d) or 200 days (PD200d) after calving. Reproduction parameters determined for heifers were age at fist insemination (AFS), whether first insemination of heifers was before 14 and 17 months of age, conception age of heifers and whether heifers became pregnant before 14 months of age as well as age at first calving (AFC). Categorical traits were scored as 1 for yes and 0 for no.

Statistical analyses. Binomial fertility traits (1 or 0) were compared between breeds within the production system using frequency tables with Chi-square tests for categorical records and

analysis of variance for continuous records using cows within breed as replicates. Breed means and probabilities of differences are provided.

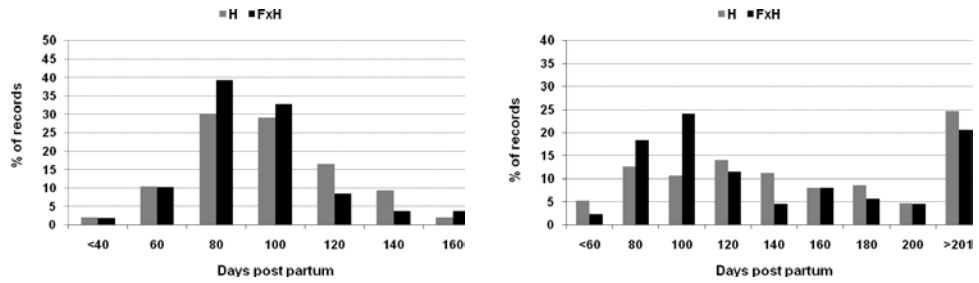
RESULTS AND DISCUSSION

Fleckvieh x Holstein heifers were inseminated earlier (P<0.05) than H heifers, i.e. 15.3±1.8 and 16.0±2.1 months of age respectively (Table 1). This resulted in more (P=0.05) FxH heifers being inseminated for the first time by 14 months of age. Age at first calving was, however, similar (P>0.05) for both breeds, i.e. 26.4 vs. 26.3 months with first service success rate higher (P<0.05) for H heifers. Fleckvieh x Holstein heifers showed oestrus more regularly, as indicated by the larger absolute number of SPC, i.e. 2.33±1.45 vs. 1.86±1.21 for H heifers. Haile-Mariam *et al. (*2004) reported a SPC of 1.84 for Holstein cows in Australia.

Table 1. The reproductive performance of Holstein (H) and Fleckvieh x Holstein (FxH) heifers and cows in a total mixed ration feeding system (CFS = interval calving to first service; DO = interval calving to conception; DIM = days in milk)

	Heifers			Cows	
Variables	H	FxH	Variables	H	FxH
Number of records	115	53	Number of records	201	108
Age first service (m)	16.0[a]±2.1	15.3[b]±1.8	Lactation number	1.83±0.98	1.97±1.03
First service <14m	0.14[a]	0.26[b]	Interval CFS (days)	91[a]±31	85[b]±31
First service <17m	0.75	0.85	First service <80DIM	0.41[a]	0.51[b]
Services per conception	1.86±1.21	2.33±1.45	Services/conception	2.33±1.51	2.34±1.68
Pregnant first service	0.56[a]	0.35[b]	Pregnant first service	0.37	0.40
Conception age (m)	17.2±2.4	17.1±2.3	Interval DO (days)	149±72	137±71
Pregnant <14m	0.21	0.23	Pregnant <100DIM	0.29[a]	0.45[b]
Age at first calving (m)	26.4±2.4	26.3±2.3	Pregnant <200DIM	0.57	0.66

[a,b]Values with different superscripts differ at P<0.10



(a) (b)

Figure 1. The distribution of the number of records for (a) interval from calving to first service (CFS) and (b) interval from calving to conception (DO) for Holstein (H) and Fleckvieh x Holstein (FxH) cows

Although average values for some traits for cows were acceptable, large variations were observed as indicated by high standard deviations. The coefficient of variation for interval traits ranged from 0.34 to 0.52 for CFS and DO respectively. The CFS interval for H and FxH tended to

differ (P=0.10) while proportion of first services within 80 days after calving was 0.41 and 0.51 respectively.

While the first service success rate did not differ (P>0.05) between breeds, the number of cows confirmed pregnant PD100d was higher for FxH in comparison to H cows, 0.45 vs. 0.29 respectively (Table 1). Only 57 and 66% of all cows were confirmed pregnant within 200 days postpartum. According to an Australian survey (Little, 2003), this level of performance would indicate reproductive problems in a herd. Mackey *et al*. (2007) reported that in 19 Holstein-Friesian dairy herds in Ireland, fertility performance was generally poor with the interval to first service being 84.4±35.4 days and the first insemination success rate 40.6±0.68%. The 100-day in-calf rate was 46.0±0.68% and CI 404±65 days. The major causes of the poor reproductive performance in these herds were the prolonged interval to first service and the poor AI success rate at first AI.

CONCLUSION

Absolute differences in reproductive performance in favour of FxH cows and heifers were observed for a number of fertility parameters. While first insemination was earlier for FxH heifers, age at first calving did not differ between FxH and H heifers because of a higher first insemination success rate in H heifers. Similarly, FxH cows were inseminated earlier after calving than H cows with a larger proportion pregnant by 100 days in milk. This advantage, however, did not result in a shorter interval (number of days) between calving and conception. While in this study crossbred heifers and cows showed better absolute reproductive values in comparison to purebred animals, differences between breeds were not significant in all instances. As reproduction management strongly affects the performance of dairy cows, a larger data set and possibly records from other herds might reduce variability in fertility traits.

REFERENCES

Funk D.A. (2006) *J. Dairy Sci.* **89**: 1362.
Grogan A., Wickham B., Olori V. and Cromie A. (2005) *Stočarstvo* **59**: 339.
Haile-Mariam M., Bowman P.J. and Goddard M.E. (2004) *Austr. J. Agric. Res.* **55**:77.
Heins B.J. and Hansen L.B. (2012) *J. Dairy Sci*. **95**: 918.
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.
Little S. 2003. The InCalf Book for dairy farmers. Dairy Australia.
Mackey D.R., Gordon A.W., McCoy M.A., Verner M. and Mayne C.S. (2007) *Animal* **1**: 29.
Makgahlela L. (2008) National Milk Recording Scheme. Newsletter No 13. p: 20.
McAllister A.J. (2002) *J. Dairy Sci*. **85**: 2352.
Mostert B.E., Van der Westhuizen R.R. and Theron H. (2010). *SA J. Anim. Sci*. **40**: 156.
Muller C.J.C., Botha J.A., Cloete S.W.P. and Potgieter J.P. (2009). *Proc. Assoc. Advmt. Anim. Breed. Genet*. **18**: 616.
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**: 4401.

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:443-446

THE BEEF PRODUCTION OF A JERSEY HERD AS AFFECTED BY CROSSBREEDING USING FLECKVIEH SIRES

C.J.C. Muller[1], S. Goni[2], K. Dzama[2] and J.A. Botha[1]

[1]Research & Technology Development Services, Directorate: Animal Sciences, Private bag X1, Elsenburg 7607, South Africa; [2]Department of Animal Sciences, University of Stellenbosch, Stellenbosch 7600, South Africa

SUMMARY

Beef production is a natural possibility in a dairy herd through cull cows and bull calves. This is not always exploited fully probably because of its relatively small contribution to farm income. While producing high quality beef, the growth rate of Jersey (J) bull calves for veal and beef is low in comparison to other dairy breeds. This could be improved by crossbreeding with beef breeds. In this paper the beef production of purebred J and Fleckvieh x Jersey (FxJ) bull calves was compared. Bull calves were reared similarly for veal, i.e. a carcass weight not exceeding 100 kg, or as steers for beef to 21 months of age. In the veal production system, for J and FxJ bull calves, respectively, the mean±se birth weight of 27.5±1.2 and 31.9±0.8 kg, live weight at 6 months of age of 166.2±10.4 and 190.0±20.1 kg, average daily gain (ADG) of 0.754±0.013 and 0.865±0.017 kg and marketing age at 7.3±0.1 and 6.2±1.2 months differed ($P<0.01$). In the beef production system, for J and FxJ bull calves, respectively, the mean±se birth weight of 26.5±1.0 and 33.4±1.1kg, end live weight at 21 months of age of 324.4±10.2 and 433.0±13.3 kg, ADG of 0.465±0.016 and 0.624±0.021 kg differed ($P<0.01$). Results indicate a potentially higher beef income for crossbred veal calves and steers. Further studies are required to determine an optimal feeding programme and marketing age as well as its effect on beef quality characteristics.

INTRODUCTION

In South Africa, the beef potential of dairy herds is not always exploited fully. As most dairy farmers are not bull breeders, bull calves could be reared for veal or beef. However, J bull calves are regarded as unwanted animals and are sold at low prices. Even though the beef quality of J steers is high in terms of tenderness (Koch *et al*. 1976) and meat:bone ratio (Purchas *et al*. 2003), their growth potential is low in comparison to other dairy breeds (Morgan *et al*. 1969, McIvor, 2004). Specialization of farming systems have resulted in most dairy herds becoming a purely milk production system in contrast to past systems. In the 1980's, a major portion of the beef animals in the United Kingdom were born in dairy herds and were reared for beef production. Breeding and selection programmes towards increased milk yields have resulted in cows showing more dairy character or "sharpness" (Hansen, 2003) with cows having a lower beef potential in comparison to the earlier British Friesian type dairy cows. Kempster *et al*. (1988) found that Canadian Holsteins slaughtered either at 16 and 24 months of age, had a lower ($P<0.05$) carcass weight and conformation score in comparison to British Friesian steers. The growth in the Jersey breed replacing Friesian or Holstein herds has further reduced the beef potential of the dairy industry. Culling of cows not becoming pregnant to maintain a strict seasonal calving system have in some countries like Ireland resulted in fertile cows requiring low replacement rates (<20%) to maintain herd sizes (Downing 2013). This provides the opportunity to inseminate a considerable portion of the herd with beef semen to increase the beef potential of dairy herds. In South Africa this practice is not always possible as the internal herd growth of most dairy herds is questionable because of high culling rates of cows and poor success rate of heifer rearing.

Crossbreeding has become a system to overcome some breeding problems like fertility and longevity in some dairy breeds (Funk 2006). Little attention has been given towards using dual-

purpose breeds in crossbreeding programmes which provides the opportunity to maintain the milk yield of cows while increasing the beef production of crossbred animals. One such a breed to consider is the Fleckvieh (F), a Simmental-derived breed from Germany. This is a dual-purpose breed with medium to high (in comparison to Holstein cows) milk yield levels and milk components while also having a high beef production potential. The aim of the paper is to compare the beef production of J and FxJ bull calves reared intensively for veal and for beef in a partly pasture-based feeding system.

MATERIAL AND METHODS

Location and Animals. This paper was based on an on-going breed-comparison study being conducted at the Elsenburg Research Farm of the Western Cape Department of Agriculture (Muller *et al*. 2009). 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. To create two comparative pure- and crossbred dairy herds, all available J cows were divided into two groups according to estimated breeding value for milk yield. Groups were randomly allocated to be inseminated by J or F bulls. The following lactation cows were inseminated with the alternative sire breed. The progeny born from J and F sires were subsequently inseminated with the same breed. Pure- and all crossbred (comprising 50 and 75% F) bull calves born were used in the beef or veal production system. Bull calves born within 7 days from each other were allocated to the beef production system while all other bull calves were used in the veal production system. For the veal production system calves were fed intensively using a commercial calf starter meal to 2 months of age and a calf growth meal to marketing, viz. a carcass weight not exceeding 100 kg. For the beef production system, J and FxJ bull calves were castrated at about 2 months of age and reared similarly as the veal production system to 3 months of age after which they were put on kikuyu pasture supplemented with about 2 kg of a calf growth meal to 6 months of age. After this stage, they were kept on natural pasture, i.e. pasture was rain-fed and no fertilizers were used. During summer droughts, pasture was supplemented with oats hay. Fresh drinking water was freely available at all times.

Data recording. Birth weights were recorded when bull calves were removed from their dams to be put into individual crates at two days of age. Thereafter calves were weighed once a month. On reaching a live weight of about 180 kg, calves reared for veal, were weighed once a week on a Thursday. When a live weight of approximately 195 kg was reached, bull calves were marketed the following Tuesday. Calves were weighed before leaving to the abattoir (end live weight) and hot and cold carcass weights were recorded after slaughter. Bull calves reared for beef were grouped according to calving date which had to be within 7 days of each other for both breeds. This was to ensure that animals from both breeds were exposed to similar environmental conditions over the 21-month growing-out period. Similarly, bull calves were weighed at birth and thereafter once a month until marketing at 21 months of age when they were transported to the abattoir.

Statistical analyses. Growth traits were compared between breeds within production system by analysis of variance using cows within breed as replicates. Data on all crossbred combinations (50 and 75% F) were grouped together. Breed means and probabilities of differences are provided.

RESULTS AND DISCUSSION

The birth weight of J and FxJ bull calves reared for veal differed ($P<0.01$) being 27.5±1.2 and 31.9±0.8 kg respectively (Table 1). Crossbred bull calves had a higher ($P<0.01$) average daily

gain (ADG) thus reaching the required live weight for marketing as veal earlier (P<0.01) than purebred J, i.e. 6.2±0.1 and 7.3±0.1 months of age respectively. The ADG of FxJ and J veal calves differed (P<0.01) being 0.865±0.017 and 0.754±0.013 kg. The birth weight of J and FxJ bull calves reared for beef differed (P<0.01) being 26.4±1.0 and 33.4±1.1 kg respectively. Crossbred bull calves had a 34% higher (P<0.01) end live weight at marketing at 21 months of age of 433.0±13.3 kg in comparison to 324.4±10.2 kg for J steers. The ADG for FxJ was higher (P<0.01) than for J steers being 0.624±0.021 and 0.465±0.016 kg respectively.

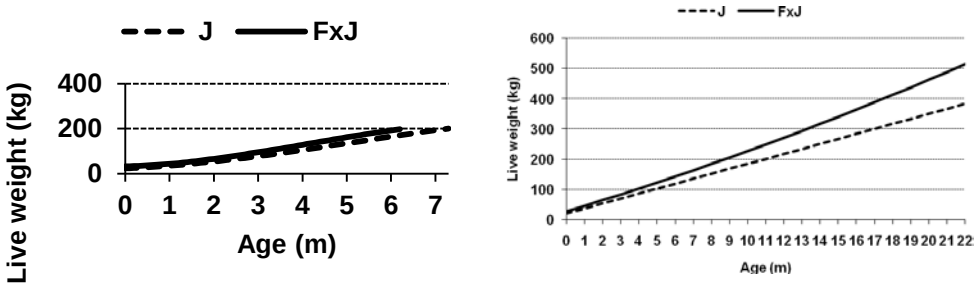
Table 1. The mean±se growth performances of Jersey (J) and Fleckvieh x Jersey (FxJ) bull calves reared intensively for veal or in a partially pasture-based feeding system for beef production ([1]50% F: n=22, 75% F: n=17; [2]50% F: n=17, 75% F: n=8)

	Veal production system		Beef production system	
Variables	J	FxJ	J	FxJ
Number of records	22	39[1]	22	25[2]
Birth weight (kg)	27.5[a]±1.2	31.9[b]±0.8	26.4[a]±1.0	33.4[b]±1.1
End live weight (kg)	193.6±2.0	194.4±2.5	324.4[a]±10.2	433.0[b]±13.3
Marketing age (m)	7.27[a]±0.12	6.21[b]±0.08	21.06±0.08	21.05±0.08
Average daily gain (kg)	0.754[a]±0.013	0.865[b]±0.017	0.465[a]±0.016	0.624[b]±0.021
Hot carcass weight (kg)	93.2±1.8	97.9±1.3	161.1[a]±7.9	204.4[b]±8.1
Dressing-out (%)	0.48[a]±0.01	0.50[b]±0.01	0.49±0.017	0.47±0.011

[a,b]Values with different superscripts within production system differ at P<0.01

Early work by Naude and Armstrong (1967) in South Africa also found low growth rates and efficiency of gain for purebred Jersey steers in comparison to beef-Jersey crossbred steers. In that study the weight gain of J bulls was improved by 39% by crossbreeding with Simmental bulls. Morgan *et al.* (1969) and Barton *et al.* (1994) also found that the disadvantages of pure J cattle are greatly reduced by crossbreeding with beef breeds.

The live weight of bull calves reared as veal or steers reared as beef is presented in Figure 1 demonstrating the earlier age of marketing for veal FxJ calves as well as the higher live weight of FxJ steers at the same marketing age in comparison to J calves and steers respectively.



(a) (b)

Figure 1. The live weight of Jersey (J) and Fleckvieh x Jersey (FxJ) bull calves reared as (a) veal to 100 kg carcass weight and (b) as steers for beef to 21 months of age

The Jersey breed is becoming increasingly popular, especially for pasture-based dairy farming systems. Some research should be conducted to determine the effect of different beef breeds on the beef potential of bull calves which are usually culled soon after birth. To include a substantial beef production option in a dairy herd is, however, only possible when the culling rate of cows in the herd is low requiring a low replacement rate. This causes a strong internal herd growth rate

resulting in surplus dairy heifers which could be sold as breeding animals especially when herd expansion is limited. Alternatively, when the market for such heifers is poor, a significant portion of the cows in the herd could be inseminated with beef semen to create a beef option for the dairy herd. This requires further research to determine the best beef breeds suitable to be used in such a production system. Earlier work by Morris *et al*. (1995) showed that in New Zealand beef production could be increased through higher dressing-out percentages and meat yield by using suitable beef breeds, i.e. Piedmontese and Belgian Blue sires on Friesian cows. Arpacik *et al*. (1993) showed the potential of Jersey cows in crossbreeding programmes delivering progeny from Belgian Blue and Chianina sires. Birth weights of calves from these sires were on average 34.7 and 35.0 kg respectively with no dystocia in either group of cows. The growth rate of crossbred steers was higher ($P<0.05$) than that of purebred Jerseys bulls.

CONCLUSION

In this study a breed comparison was conducted using production systems generally used by dairy farmers. Higher growth rates for FxJ in comparison to purebred J bull calves reared for either veal or beef under similar feeding conditions were observed. Crossbred bull calves reached the required live weight for veal, on average 32 days earlier than J bull calves. The end live weight of FxJ steers reared as beef in a partially pasture-based system was 34% higher than J steers. Although a higher beef production is realized from crossbreeding using a dual-purpose breed, the improvement in milk yield, milk composition and fitness traits would determine the economic value of crossbreeding. Further studies should be conducted to determine the effect of including better quality pasture into the diet of steers reared for beef as only poor quality pasture was available in the present study. This should include the effect of the inclusion of supplementary feeds to increase the performance of crossbred steers as steers being finished on grass could result in too lean carcasses.

REFERENCES

Arpacik R., Alpan O., Bayraktar M. and Cekgul E. (1993) *Lalahan-Bayvancilik-Arastirma-Enstitusu-Dergisi*. **33**: 1.

Barton R.A., Donaldson J.L., Barnes F.R., Jones C.F. and Clifford H.J. (1994) *New Zealand J. Agric. Res.* **37**: 51.

Downing K. (2013) Personal communication. ICBF. Bandon. Ireland.

Funk D.A. (2006) *J. Dairy Sci.* **89**: 1362.

Hansen L.B. (2003) 2nd Bi-Annual W.E. Petersen Symposium. Univ of Minnesota. p. 48.

Kempster A.J., Cook G.L. and Southgate J.R. (1988) Anim. Prod. **46**: 365.

Koch R.M., Dikeman M.E., Allen D.M., May M., Crouse J.D. and Campion D.R. (1976) *J. Anim. Sci*. **43**: 48.

McIvor S. (2004) *Dairy Exporter* **79**: 109.

Morgan J.H.L., Pickering F.S. and Everitt G.C. 1969. *Proc. New Zealand Soc. Anim. Prod*. **29**: 164.

Morris S.T., Parker W.J., Purchas R.W. and McCutcheon S.N. (1995) *Proc. New Zealand Grassl. Assoc*. **54**: 19.

Muller C.J.C., Botha J.A., Cloete, S.W.P. and Potgieter, J.P. (2009). *Proc. Assoc. Advmt. Anim. Breed. Genet*. **18**: 616.

Naude R.T. & Armstrong C.W.B. (1967) *Proc. S. Afr. Soc. Anim. Prod*. **1967**: 156.

Purchas R.W., Fisher A.V., Price M.A. and Berg R.T. (2003) *Meat Sci*. **61**: 329.

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:447-450

A PRELIMINARY STUDY ON BREED DIFFERENCES IN SUSCEPTIBILITY OF SHEEP TO *MYCOTOXIN SPORIDESMIN*

S.H. Phua, H. Henry and K.G. Dodds

AgResearch Invermay, Private Bag 50034, Mosgiel 9053, New Zealand

SUMMARY

Sporidesmin is the mycotoxin that causes facial eczema disease (FE) in New Zealand (NZ) livestock. In an artificial sporidesmin dosing test, the introduced Finnish Landrace sheep breed was found to be significantly more tolerant to FE than the introduced Texel breed. This finding enables cross-breeding strategies to improve disease tolerance. Combining published data with the current study, a tentative inference is that Finnish Landrace, Merino and East Friesian are more FE resistant than Romney, Texel and Border Leicester.

INTRODUCTION

Genetic adaptation to environment is a key evolutionary feature of living things. Organisms with a small genome size, such as unicellular bacteria, adapt to new environment by acquiring extra-chromosomal genes through plasmids, which enable them to grow, for instance, in the presence of antibiotics and heavy metals (Dib *et al*. 2013; Dhanarani *et al*. 2009). On the other hand, higher organisms with larger genome sizes have a wider spectrum of genes and variants that can form novel biochemical pathways for adaptation. Plants for example, when faced with herbicide glyphosate challenge, acquire resistance to the xenobiotics by changing their glyphosate metabolism and translocation (González-Torralva *et al*. 2012), gene amplification, and by increasing the enzymatic activity of a specific gene product (Salas *et al*. 2012). Such adaptive changes in animals give rise to genetic differences of sheep breeds developed in different countries.

As an island nation, NZ has an indigenous problem which is facial eczema (FE). In a severe outbreak, the disease costs the NZ sheep industry an estimated $60M. The current methods used to reduce the impact of FE are by zinc prophylactic treatment of animals and through breeding of resistant livestock. Through selection and cross breeding over the years, NZ sheep flocks generally become more tolerant to the disease. Exotic sheep breeds, developed under different foreign conditions, face a new FE challenge following importation into NZ. In this report, two introduced sheep breeds, Finnish Landrace and Texel, were tested for their relative susceptibility to FE in an artificial sporidesmin challenge experiment.

MATERIALS AND METHODS

Animals. Information on the experimental animals is summarized in Table 1. The Finnish Landrace (Finn) and Coopworth (Coop) animals were from the AgResearch Woodlands farm, Texel and Finn x Texel animals were purchased from 2 different commercial farms and housed at Woodlands for a month before the experiment. Nine animals were obtained from each source and were progeny from three sires (with 3 progeny per sire); the exception was Finn which had only 2 sires with 3 and 6 progeny (Table 1). All animals were 10-month old lambs. All Finn and Coopworth animals were females, while all Texel and Finn x Texel were males. Morris *et al.* (1995) observed no sex differences in susceptibility to FE.

Sporidesmin challenge regime. Ethics approval to conduct this work was obtained from AgResearch Animal Ethics Committee (application AEC-P516). The 36 experimental animals were allocated into 3 dosing groups of equal size. Each group had at least 1 progeny from each sire (Table 1). The 0.2, 0.3 and 0.4 dosing groups had sporidesmin dose rates of 0.2, 0.3 and 0.4 mg/kg

live-weight (LWT), respectively.

Animals were weighed and blood sampled a week before dosing: the weights were used for calculating the dosage for each animal, and the blood samples were used to determine the pre-dose levels of liver-specific enzymes, gamma-glutamyl transferase (GGT) and glutamate dehydrogenase (GDH) (in IU/L). After dosing, GGT and GDH were measured weekly for 5 consecutive weeks: GGT0 and GDH0 refer to pre-dose GGT and GDH respectively, GGT1-5 and GDH1-5 refer to GGT and GDH levels at 1- to 5-week post dosing. Levels of GGT and GDH in the blood reflect the severity of liver damage caused by sporidesmin.

Throughout the 5-week period, animals were generally kept outdoors on good pasture. After sporidesmin dosing, animals were monitored thrice daily for clinical signs of photosensitivity. Animals were housed indoors in well-ventilated shed, supplied with feed and water, if they showed any sign of restlessness, stamping of feet, pruritus, shaking or rubbing of the head, swollen eyes/ears/lips, or avoidance of sunlight.

Table 1. Experimental groups and their constituent animals

Breed	Sire[1]	Progeny (n)	Number of animals in each dosing group[2]		
			0.2	0.3	0.4
Finn	Sire A	3	1	1	1
	Sire B	6	2	2	2
Texel	Sire C	3	1	1	1
	Sire D	3	1	0	0
	Sire E	3	1	1	1
Finn x Texel	Sire F	3	1	1	1
	Sire G	3	1	1	1
	Sire H	3	1	1	1
Coopworth	Sire I	3	1	1	1
	Sire J	3	1	1	1
	Sire K	3	1	1	1

[1]Sires A - K denote different sires.
[2]Two Texel progeny (from sire D) died just before the experiment started and no replacements were available.

Statistical analyses. GGT and GDH data were natural log transformed (logGGT & logGDH) and then analysed in a mixed model which included dose rate and breed as fixed effects, and sire within breed as a random effect. The model was fitted by residual maximum likelihood (REML). The sire within breed variance was used to test for differences between breed. Dose rates by breed interactions were tested in the initial models and were dropped because none was significant. The occurrence of clinical signs was analysed in a logistic mixed model, with the same model terms as above. Correlations were calculated between residuals from the analyses at different time points.

RESULTS AND DISCUSSION

The background ranges of logGGT and logGDH for sheep are 0-4.0 and 0-2.6, respectively. With reference to these ranges, the Finn animals did not react to the sporidesmin challenge (Table 2). The Texel animals were significantly more susceptible to the toxin than Finn, with ~29% developing clinical photosensitivity. As expected the Finn x Texel animals showed an intermediate

toxin response, between that of Finn and Texel (Table 2). Nonetheless, some of the breed differences observed could be attributed to other factors associated with the different sources of animals; however these factors would be minor. For example, a major factor like rearing rank was found to have no effect on FE resistance (Morris *et al.* 2001).

The Coopworth result showed this group of animals to be at least as FE sensitive as Texel (Table 2). They were from a South Island FE-free region, where animals tend to be more FE-susceptible than their counterparts in North Island. However, Coopworths are genetically diverse, with grade-up from other breeds allowed (http://coopworthgenetics.co.nz/). Hence the group tested here may not be a representation of the whole NZ Coopworth population.

In a toxicological view that "dosage determines poison", it is intriguing to observe that increasing sporidesmin dose rates from 0.2 to 0.4 mg/kg LWT did not increase the numbers of reactant animals within breed nor between dosing groups. No explanation could be forwarded to account for this observation.

Table 2. Summary of least square means of logGGT and logGDH, and clinical cases of animals after sporidesmin challenge

Trait[1]	Breed least squares means[2]					Dose least squares means[3]			
	Finn	Texel	Finn x Texel	Coop	Mean SED	0.2	0.3	0.4	Mean SED
LogGGT0	3.3^{a}	3.42^{a}	3.64^{a}	3.55^{a}	0.16	3.47	3.48	3.48	0.08
LogGGT1	3.51^{a}	3.83^{b}	3.82^{b}	3.75^{b}	0.09	3.68	3.78	3.73	0.08
LogGGT2	3.3^{a}	4.48^{b}	4.19^{a,b}	4.90^{b}	0.38	4.51	4.23	3.95	0.32
LogGGT3	3.47^{a}	5.02^{b,c}	4.66^{a,b}	6.03^{c}	0.51	4.83	4.96	4.59	0.36
LogGGT4	3.43^{a}	5.22^{b,c}	4.9^{b}	6.18^{c}	0.57	4.91	5.02	4.89	0.31
LogGGT5	3.42^{a}	5.23^{b}	4.93^{b}	6.04^{b}	0.58	4.89	4.94	4.90	0.32
LogGDH0	0.33^{a}	1.62^{a,b}	1.35^{a,b}	2.48^{b}	0.48	2.28	1.73	1.57	0.42
LogGDH1	0.34^{a}	2.60^{b}	2.27^{b}	2.20^{b}	0.45	2.06	2.05	1.94	0.30
LogGDH2	0.46^{a}	4.25^{b}	3.13^{b}	4.44^{b}	0.60	3.60	3.16	3.26	0.34
LogGDH3	0.37^{a}	4.71^{b,c}	3.66^{b}	5.63^{c}	0.54	3.88	4.01	4.03	0.46
LogGDH4	0.34^{a}	5.20^{b,c}	4.38^{b}	5.96^{c}	0.50	4.45	4.17	4.43	0.43
LogGDH5	0.36^{a}	5.49^{b,c}	4.61^{b}	5.83^{c}	0.52	4.65	4.40	4.44	0.43
% Clinical	0^{a}	28.6^{a}	22.2^{a}	55.6^{a}		33.3	27.3	18.2	

[1]LogGGT0/logGDH0 refer to pre-dose logGGT/logGDH, logGGT1-5/logGDH1-5 refer to logGGT/logGDH at 1-5 weeks after dosing. The % Clinical refers to percentage of animals showing signs of photosensitivity over all dosage rates.

[2]Values are breed least squares means (over all dosing groups). Means with the same superscript do not differ significantly at $P<0.05$ (within trait/row).

[3]These values are dose least squares means (over all breeds); they do not differ significantly at $P<0.05$ (within trait/row).

The observation of decreasing clinical percentages with increasing dose rates was not significant (Table 2). There is evidence (authors' unpublished data) that shows that not all sheep with severe liver damage caused by sporidesmin develop photosensitivity. A current view is that a separate set of genes are involved in other factors required for clinical manifestation.

Correlations between residuals from the models were all positive from week 2 onwards, in which case they ranged between 0.71 and 0.97 for logGGT, and between 0.68 and 0.82 for logGDH. Given this and the fact that the greatest differences observed here were in weeks 3 and 4, there is no reason to revise the industry practice of measurement at 3 weeks post-dose.

Year 1972 saw the first importation into NZ of Finn, East Friesian, German Whiteheaded Mutton and Oxford breeds from the United Kingdom and Ireland (Clarke and Meyer 1977). These animals failed the quarantine due to the suspected occurrence of scrapie in some individuals (Tervit *et al*. 1986). The second importation in 1984 using frozen embryos successfully introduced the Texel and Oxford Down from Denmark, and Finn and Texel from Finland (Tervit *et al*. 1986); the numbers of rams and ewes used to generate the embryos were 11 rams/47 ewes, 14/46, 17/47 and 5/23, respectively. The final numbers of lambs born from the imported embryos were 39 Texel and 19 Oxford Down from Denmark, and 56 Finn and 28 Texel from Finland (Tervit *et al*. 1986). The present day Finn and Texel animals come from these bottlenecks. This report shows that the Finn breed is significantly more FE resistant than the Texel breed.

Based on a sporidesmin dose rate of 2 mg/kg LWT, and the results of liver injury score and clinical cases, Smith *et al*. (1980) determined the Merino breed to be more FE tolerant than Romneys, Border Leicesters and Romney/Border Leicester cross. On a dose rate of 0.15 mg/kg LWT and the resultant logGGT3 data, Morris *et al*. (1995) found Finn to be more resistant to sporidesmin than Romneys, with Finn/Romney cross being intermediate. Using a 0.14 mg/kg LWT dose rate and the logGGT3 results obtained, East Friesians were shown to be more FE tolerant than Romneys (Morris *et al*. 2001). Combining the above results with the current study suggests that the relative FE sensitivity of various sheep breeds is, Finn, Merino and East Friesian are more FE resistant than Romney, Texel and Border Leicester.

CONCLUSION

This report shows that the introduced Finnish Landrace sheep breed is much more tolerant to sporidesmin, hence to FE, than the introduced Texel breed.

ACKNOWLEDGMENTS

Funding for this work was provided to S. Phua by Beef + Lamb New Zealand via Ovita.

REFERENCES

Clarke J.N. and Meyer H.H. (1977) *1977 Ruakura Farmer's Conference Proceedings* p34.

Dhanarani T.S., Shankar C., Park J., Dexilin M., Kumar R.R. and Thamaraiselvi K. (2009) *Poult. Sci.* **88**: 1381.

Dib J.R., Liebl W., Wagenknecht M., Farías M.E. and Meinhardt F. (2013) *Appl. Microbiol. Biotechno.* **97**: 63.

González-Torralva F., Rojano-Delgado A.M., Luque de Castro M.D., Mülleder N. and De Prado R. (2012) *J. Plant Physiol.* **169**: 1673.

Morris C.A., Towers N.R., Wesselink C. and Amyes N.C. (1995) *N. Z. J. Agric. Res.* **37**: 547.

Morris C.A., Wesselink C., Wilson J.A., DeNicolo G. and Towers N.R. (2001) *N. Z. J. Agric. Res.* **44**: 47.

Salas R.A., Dayan F.E., Pan Z., Watson S.B., Dickson J.W., Scott R.C. and Burgos N.R. (2012) *Pest Manag. Sci.* **68**: 1223.

Smith B.L., Stanbridge T.A. and Embling P.P. (1980) *N. Z. Vet. J.* **28**: 35.

Tervit H.R., Baker R.L., Hoff-Jorgenson R., Lintukangas S., MacDiarmid S.C. and Rainio V. (1986) *Proc. N. Z. Soc. Anim. Prod.* **46**: 245.

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:451-454

GENETIC ORIGIN OF ARAPAWA SHEEP AND ADAPTATION TO A FERAL LIFESTYLE

N.K. Pickering[1], E.A. Young[1], J.W. Kijas[2], D.R. Scobie[3] and J.C. McEwan[1]

[1] AgResearch Invermay, Private Bag 50-034, Mosgiel 9053, New Zealand
[2] CSIRO Livestock Industries, Services Road, St Lucia, Brisbane 4067 Queensland, Australia
[3] AgResearch Lincoln, Private Bag 4749, Christchurch 8140, New Zealand

SUMMARY

This work aimed to investigate the population history and patterns of genetic diversity present within the isolated population of New Zealand Arapawa sheep. In order to identify genetic regions associated with reversion to a feral lifestyle, a selection sweep analysis was performed comparing 40 Arapawas to related breeds using Wright's fixation index (F_{ST}). Comparisons were graphed as the moving average of 5 F_{ST} values. A threshold of 0.25 was used to identify significant regions; 8 genomic regions were identified for the Arapawa and Florida Gulf Coast Native, 9 for the Arapawa and Castellana and 3 for the Arapawa and Australian Merino breed pair comparisons. One region on chromosome 2 was identified in all three comparisons with two underlying genes, CFDP2 and NAB1. Other genes identified were RXFP2, IFT88, SLC9A3, HERC2, NIPA1, NIPA2 and DACH2. The current work confirms Arapawa sheep are an important reservoir of unique gene variants available to the New Zealand sheep industry.

INTRODUCTION

The feral sheep of Arapawa Island are thought to be the oldest feral flock in New Zealand. Their isolated island location makes them an excellent example of natural selection accompanying reversion to a feral existence. Anecdotal information suggests the Arapawa was derived from an Australian Merino flock introduced and farmed on the island in 1867. However, recent research suggests the New Zealand feral Arapawa sheep are most closely related to the Gulf Coast Native (GCN) breed, which in turn comprises a significant component of the Castellana (Young *et al.* 2011). Further studies into the GCN have now found two separate lines of the GCN, the Florida and the Louisiana (Kijas *et al.* 2012a).

Phenotypically, Arapawa sheep are unique, differing from domestic sheep in New Zealand as summarised by Orwin and Whitaker (1984). Arapawa sheep have predominantly black skin and wool colouration, with white on the distal part of the tail and a white crown which can extend down the face and throat. Ewes are generally polled, with some growing small scurs; males have large curled horns with approximately 10% being polled. Arapawa sheep are small bodied with long legs and males weigh 51kgs and females 38kg in the wild. Fleece weight is low at approximately 2kgs per year, with shedding occurring in some animals. The wool has high bulk and fibre diameter of approximately 22 µm. Ewes can ovulate throughout the year, lambs are born small with a hairy coat which is later shed. As both the Arapawa and particularly the GCN are reportedly naturally resistant to parasites and footrot, these animals may be a reservoir of unique gene variants for sheep breeds more commonly farmed in New Zealand.

The current work attempts to identify which of the two GCN lines is most closely related to the Arapawa, and gene regions that have been subject to selection sweeps as part of the Arapawa reversion to a feral lifestyle.

MATERIALS AND METHODS

Resource. The 40 Arapawa animals described by Young *et al.* (2011) were sourced from New Zealand flocks. The 56 Australian Merinos, 23 Castellana and 95 GCN (40 Florida and 55

Louisana origin) were sourced from the ovine HapMap project (Kijas *et al.* 2012b). All animals were genotyped using Illumina's OvineSNP50 Beadchips (Kijas *et al*. 2012b). All genotypes were quality checked before analysis. The SNPs were discarded if the minor allele frequency was <0.02 in a population comparison or if the call rate was less than 95%.

STRUCTURE. Model based clustering using SNP genotypes from 132 individuals was performed using the program STRUCTURE (Pritchard *et al.* 2000). Three runs were performed at K = 2 - 4, where K is the number of assumed subpopulations. The admixture model was applied and runs comprised 5000 burn-in replications followed by 5000 run lengths.

Wright's F_{ST}. SNPs not aligned on Ovine genome v3 were discarded. Wright's fixation index (F_{ST}) was calculated for each breed pair as $(H_T - H_S)/ H_T$, where H_T is the expected heterozygosity for the overall breed pair population, and H_S is the expected heterozygosity of the subpopulation. The F_{ST} is the extent of genetic difference between subpopulations. Smoothed estimates were calculated as the moving average with a window of 5 (WIN5) SNPs and plotted for Arapawa versus each of the other breeds (Florida GCN, Castellana and Australian Merinos). A threshold level of a WIN5 F_{ST} value greater than 0.25 was chosen and the identified regions were examined using Ovine genome v3 to identify underlying genes.

RESULTS AND DISCUSSION

Structure and principal components. Model based clustering was used to examine the relationship between three populations: the Arapawa and two lines of the Gulf Coast Native breed, the best solution (K = 3) is shown in Figure 1. The Florida GCN line had the highest proportion of common ancestry with Arapawa, contributing 27% of the dark grey component and only 6% of the black. The light grey component (67%) of the Arapawas has been contributed from elsewhere. The Louisiana native is believed to be derived from animals introduced by explorers from Latin America, whereas the Florida natives are most likely founded from sheep arriving with settlers on the east coast (Kijas *et al.* 2012a). This supports the theory stated in Young *et al.* (2011) that it was possible that Arapawa sheep were introduced to New Zealand by whalers in the early 19th Century.

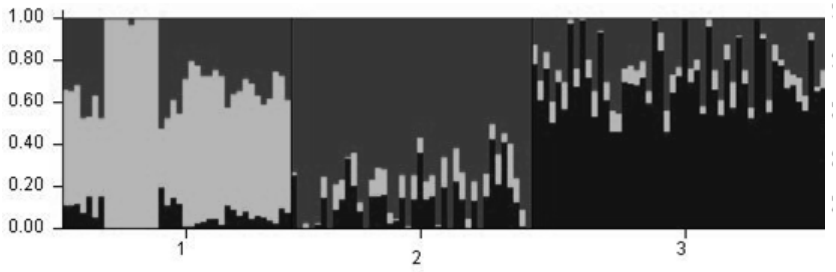


Figure 1. Admixture analysis using 5000 SNP, for Arapawa (1), Florida GCN (2) and Louisiana GCN (3). For each animal, the proportion of 3 genomic components (light grey, dark grey and black) is given on the Y axis.

Selection sweep. Genotypes from breed pairs were used to search for genomic regions with signatures of selection (Arapawa versus Florida GCN, Castellana or Merino). A number of significant regions identified in the comparisons with Arapawa e.g. Figure 2 shows the moving window of 5 (WIN5) F_{ST} values across the genome for the Arapawa versus Florida GCN comparison. This comparison identified 8 significant peaks (> 0.25 WIN5F_{ST} value) on chromosomes 1, 2, 9, 10 and X. Table 1 lists significant regions from all the comparisons and the major genes identified within.

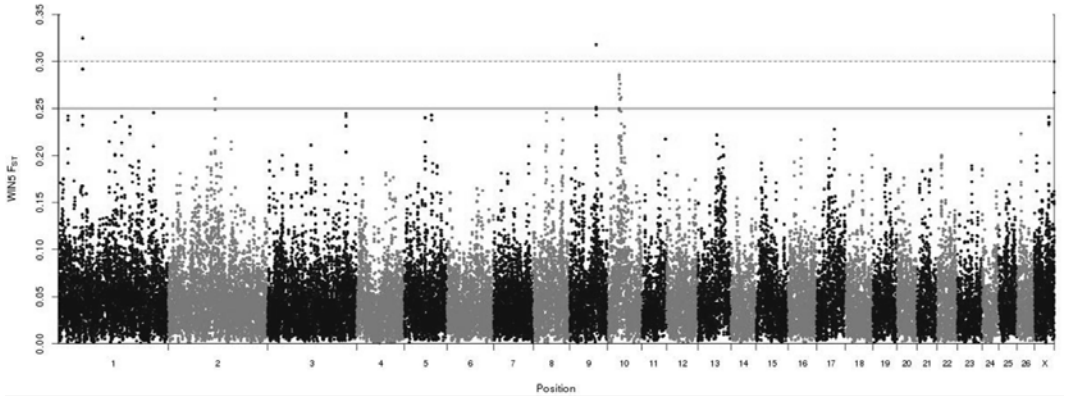


Figure 2. Manhattan plot of the moving window of 5 (WIN5) F_{ST} values between Arapawa and Florida GCN. Ordered on Ovine genome v3, WIN5F_{ST} = 0.25 (solid line), WIN5F_{ST} = 0.30 (dash line).

Table 1. The number of regions, chromosomes and significant known genes found under the peaks with WIN5 F_{ST} values > 0.25 for each breed comparison with Arapawa.

Comparison	Regions	Chromosomes	Genes
Florida GCN	8	1,2,9,10,X	NAB1, CFDP2, RXFP2
Castellana	9	1,2,4,6,9,10,13	HERC2, NIPA1, NIPA2, NAB1, CFDP2, RXFP2, IFT88
Merino	3	2,16	NAB1, CFDP2, SLC9A3

The polled/horns gene relaxin/insulin-like family peptide receptor 2 (RXFP2) was identified in both the comparisons of Arapawa to Florida GCN and Castellana. This region acts as a positive control, as this gene is known to be associated with polledness in sheep (Kijas *et al.* 2012b). Most Arapawa and Merino rams are horned and the GCN Florida and Castellana breeds are predominantly polled. The results suggested that there had been natural selection for horns once animals were introduced to Arapawa Island, as the selection sweep was notable, based on diversity reduction.

The gene SLC9A3 (solute carrier family 9, subfamily A, member 3) has been associated with pH regulation in mice. (Schultheis *et al.* 1998). This gene was identified in the Arapawa/ Merino comparison, with reduced genetic diversity in Merinos. Merino meat has a higher ultimate pH than crossbreds (Young *et al.* 1993) and high pH has been associated with undesirable flavours (Hopkins and Fogarty 1998). It would be interesting to assess the meat quality of the Arapawa as it may be leaner like the Merino, yet with a lower pH.

The genes: HECT and RLD domain containing E3 ubiquitin protein ligase (HERC2), non-imprinted in Prader-Willi/Angelman syndrome 1 and 2 (NIPA1 and NIPA2) are within the region associated with the paternally imprinted Prader-Willi syndrome and the maternally imprinted Angelman syndrome (Cassidy *et al.* 2012). In this analysis, selection sweep was towards homozygosity in the Arapawas. HERC2 has also been associated with hair colour in cattle (Han *et al.* 2008) and in a long haplotype block in Spanish Churra sheep (Garcia-Gamez *et al.* 2012).

The intraflagellar transport 88 (IFT88) gene located on chromosome 10 is a homolog to mouse TG737 associated with recessive polycystic kidney disease (Moyer *et al.* 1994). Reduced genetic diversity was observed for Arapawa in this region.

One region on chromosome 2 was identified in all three breed pair comparisons, with reduced diversity observed for Merino and Castellana. Two genes were identified in this region. The gene craniofacial development protein 2 (CFDP2/*p97bcnt*) created by gene duplication of *bcnt/cfdp1* (Iwashita *et al*. 2006). There are 48 copies of CFDP2 in Ovine genome v3 and association of intra-muscular fat with a copy on chromosome 14 was found in a genomic selection study of carcass and meat quality traits in Australian sheep (Daetwyler *et al.* 2012). The second gene in this region, NGFI-A binding protein (NAB1), is highly expressed in cardiac muscle and is implicated as a regulator of pathological cardiac growth (Buitrago *et al.*, 2005).

CONCLUSION

The Arapawa shared the highest proportion of common ancestry with the GCN Florida, supporting previous evidence that they were introduced to New Zealand by Whalers in the 19th century. Nine genes were identified as under selection from the F_{ST} analysis, seven are described above. The phenotypes associated with reversion to a feral lifestyle are unknown for most genes identified as under selection. The exceptions are horns and coat colour. However, this study provides a list of candidate genes for future studies of domestication.

REFERENCES

Buitrago M., Lorenz K., Maass A. H., Oberdorf-Maass S., Keller U., Schmitteckert E. M., Ivashchenko Y., Lohse M. J. and Engelhardt S. (2005) *Nat. Med.* **11**: 837.

Cassidy S. B., Schwartz S., Miller J. L. and Driscoll D. J. (2012) *Genet. Med.* **14**: 10.

Daetwyler H. D., Swan A. A., van der Werf J. H. and Hayes B. J. (2012) *Genet. Sel. Evol.***44**: 1.

Garcia-Gamez E., Sahana G., Gutierrez-Gil B. and Arranz J.-J. (2012) *BMC Genet.***13**: 43.

Han J., Kraft P., Nan H., Guo Q., Chen C., Qureshi A., Hankinson S. E., Hu F. B., Duffy D. L., Zhao Z. Z., Martin N. G., Montgomery G. W., Hayward N. K., Thomas G., Hoover R. N., Chanock S. and Hunter D. J. (2008) *PLoS Genet.* **4**: e1000074.

Hopkins D. L. and Fogarty N. M. (1998) *Meat Sci.* **49**: 477.

Iwashita S., Ueno S., Nakashima K., Song S.-Y., Ohshima K., Tanaka K., Endo H., Kimura J., Kurohmaru M., Fukuta K., David L. and Osada N. (2006) *Mol. Biol. Evol.* **23**: 798.

Kijas J. W., Miller J. E., Hadfield T., McCulloch R., Garcia-Gamez E., Porto Neto L. R. and Cockett N. (2012a) *PLoS One*. **7**: e41508.

Kijas J. W., Lenstra J. A., Hayes B., Boitard S., Porto Neto L. R., San Cristobal M., Servin B., McCulloch R., Whan V., Gietzen K., Paiva S., Barendse W., Ciani E., Raadsma H., McEwan J., Dalrymple B. and other members of the International Sheep Genomics Consortium. (2012b) *PLoS Biol.* **10**: e1001258.

Moyer J. H., Lee-Tischler M. J., Kwon H., Schrick J. J., Avner E. D., Sweeney W. E., Godfrey V. L., Cacheiro N. L. A., Wilkinson J. E. and Woychik R. P. (1994) *Science* **264**: 1329.

Orwin D. F. G. and Whitaker A. H. (1984) *N. Z. J. Zool.* **11**: 201.

Pritchard J. K., Stephens M. and Donnelly P. (2000) *Genetics*. **155**: 945.

Schultheis P. J., Clarke L. L., Meneton P., Miller M. L., Soleimani M., Gawenis L. R., Riddle T. M., Duffy J. J., Doetschman T., Wang T., Giebisch G., Aronson P. S., Lorenz J. N. and Shull G. E. (1998) *Nat. Genet.* **19**: 282.

Young E. A., Kijas J. W., McCulloch R., Scobie D. R., McRae K. M., Pickering N. K., Dodds K. G. and McEwan J. C. (2011) *Proc. N. Z. Soc. Anim. Prod.* **71**: 248.

Young O. A., Reid D. H. and Scales G. H. (1993) *N. Z. J. Agric. Res.* **36**: 363.

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:455-458

COMPOSITE SIGNATURES OF DIRECTIONAL SELECTION IDENTIFIED MULTIPLE GENES FOR STATURE ON BOVINE CHROMOSOME 13 AND 14

I.A.S. Randhawa, M.S. Khatkar, P.C. Thomson and H.W. Raadsma

ReproGen - Animal Bioscience Group, Faculty of Veterinary Science, University of Sydney, Camden, NSW 2570, Australia

SUMMARY

Advances in genomic tools have made it possible to identify signatures of positive selection for complex traits in non-inbred populations. We investigated the evidence of selective sweeps for stature in 9 breeds of European *Bos taurus* by using 34,857 SNPs genotyped with an Illumina BovineSNP50 chip assay. The genotypic data were grouped in two phenotypic categories according to body size of the breeds (small-medium and medium-large). We implemented our recently developed composite index of multiple selection tests called MFR (mean fractional rank) that combines the rank distribution of three complementary test statistics to capture signatures of selection. Two strong selective sweeps were detected at loci that harbour UQCC-GDF5 and PLAG1-CHCHD7 gene pairs on chromosome 13 and 14, respectively. The two loci have previously been associated with height in humans, while PLAG1-CHCHD7 has also been reported for stature in cattle. Further investigations of the several variants in newly identified genes may help to explain the biological function of causative mutations in the diversity of bovine stature.

INTRODUCTION

Recent advances in genomic tools have facilitated studies on diverse genetic models and complex modes of their underlying inheritance in many species. Understanding the role of genetic variants in phenotypic diversity has always been challenging, and requires specific resources, tools, costs and time. Recently we developed a new method that combines multiple pieces of evidence of trait-specific selection signatures, by using the rank distribution of single nucleotide polymorphism (SNP) and haplotype-based selection tests (Randhawa *et al.* 2013).This method can be used to expand our knowledge about the genomic regions and genes controlling the diverse functions of complex traits in domestic species.

Height is a polygenic trait with high heritability in many species including cattle (Kemper and Goddard 2012). Genetic architecture of human height has been extensively investigated to find variants with major effects across the genome (Lettre *et al.* 2008; Sanna *et al.* 2008). In cattle, to date, only a few genes responsible for stature (body size) have been reported from genome-wide association studies (Pryce *et al.* 2011; Visscher and Goddard 2011; Nishimura *et al.* 2012). The known genes explain only a small proportion of the existing phenotypic variation in bovine stature (Kemper and Goddard 2012). Hence, further studies implementing new genomic tools are required to improve understanding of the genetic control of stature. To find undiscovered genetic factors, we investigated several breeds of cattle for their diversity in body size in this study.

MATERIALS AND METHODS

Data on stature in 241 animals representing nine breeds of European *Bos taurus* (Decker *et al.* 2009; Gautier *et al.* 2010) were used for this study. These breeds were selected based on the availability of the precise information on stature. The animals were genotyped with an Illumina BovineSNP50 chip assay. After quality control (MAF > 0.05) 34,857 SNPs were retained for further analysis. The animals were grouped in two phenotypic categories according to body size of their breeds (small-medium and medium-large). Breeds (sample size) selected for the small-medium group were Angus (44), Hereford (31), Limousin (35) and Romosinuano (8). Breeds

(sample size) of the medium-large sized group were Charolais (55), Chianina (8), Piedmontese (26), Romagnola (24) and Simmental (10). Imputation of missing genotypes and haplotype phasing were performed with BEAGLE 3.3 (Browning and Browning 2007). All the SNPs were mapped on UMD3.1 bovine assembly. Ancestral and derived allelic phases of these SNPs were acquired from Decker *et al.* (2009) and Matukumalli *et al.* (2009).

The analysis was performed with the mean fractional rank (MFR) method, explained in the companion paper (Randhawa *et al.* 2013), in which we combined results from commonly used 3 tests i.e., population differentiation (F_{ST}), change in derived allele frequency (ΔDAF) and across population extended haplotype homozygosity (XP-EHH) to capture evidence for selection from SNP data across multiple populations. The $-\log_{10}(p\text{-value})$ of MFR statistics were smoothed by averaging over SNPs within 1 Mb sliding windows centered at each SNP and their genome-wide top 0.1% of were used to declare the SNPs as significant. Clusters of significant SNPs were identified as the genomic regions under selection and their positions (± 0.5 Mb) were investigated to report the candidate genes under selection.

RESULTS AND DISCUSSION

Figure 1 shows the genome-wide map of the smoothed MFR scores from comparing a panel of small-medium against medium-large body size cattle breeds. Two regions of strong selective sweeps were detected which harbour multiple gene pairs on *Bos taurus* autosomes (BTA) 13 and 14 (Figure 1, Table 1). The two regions show an enrichment of high scores based on F_{ST} and XP-EHH as depicted in Figure 2. Simultaneously, an additional prominent peak at BTA1 – which is close to the significance threshold (Figure 1) – is localized at the *POLL* locus (Allais-Bonnet *et al.* 2013). This can be explained by the existence of strong secondary phenotype diversity for polledness across two breed groups, see Randhawa *et al.* (2013) for polled against horned breeds panel analysis. MFR analyses of individual breed pair data ($n \geq 24$) with contrasting body size confirmed both candidate loci on BTA13 and BTA14, however, these identified a higher number of additional peaks, likely breed-specific or spurious, than combined panels (results not shown).

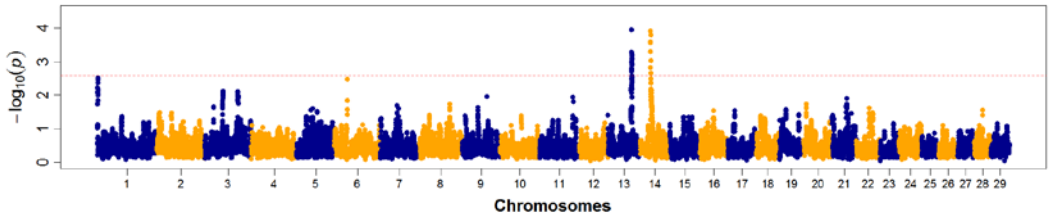


Figure 1: Genome-wide Manhattan plot of smooth $-\log_{10}(p\text{-value})$ of the Mean Fractional Ranks (MFR). Dashed (red) line indicates the top 0.1% threshold of significance.

UQCC-GDF5 locus. On BTA13, a 1.8 Mb selective sweep was localized where the ubiquinol-cytochrome c reductase complex chaperone (UQCC) and growth differentiation factor 5 (GDF5) genes are located at 65.233–65.344 Mb positions on UMD3.1 assembly of cattle (Table 1). UQCC is involved in growth control network in a number of mammalian species; along with several other genes it initiates and promotes morphogenesis and skeletal growth. GDF5 is involved in bone growth and its mutations are associated with several disorders in human skeletal development. Common variants in these two genes have been associated with variation in human height (Sanna *et al.* 2008) and strong signals of recent selection have also been identified at the GDF5 locus in European and East Asian human populations (Voight *et al.* 2006). Ensembl searches show that

UQCC and GDF5 genes have three and two mis-sense mutations, respectively (Table 1). The functional role of the putative variants underlying UQCC-GDF5 locus is unknown in cattle.

PLAG1-CHCHD7 locus. On BTA14 a 1.0 Mb selective sweep was localized where the pleiomorphic adenoma gene 1 (PLAG1) and coiled-coil-helix-coiled-coil-helix domain containing 7 (CHCHD7) genes are located at 25.007–25.059 Mb positions in cattle (Table 1). PLAG1 is consistently rearranged in salivary gland adenomas and its activation results in up regulation of target genes. CHCHD7 has no known function. Both genes have less obvious connections to body size, however, they have been considered either being in strong linkage disequilibrium with the actual causal alleles in other genes or they might indirectly regulate height via different pathways (Lettre *et al.* 2008). Previously, these two genes have been associated with height in humans (Lettre *et al.* 2008) and stature in cattle (Karim *et al.* 2011; Pryce *et al.* 2011; Nishimura *et al.* 2012). Ensemble reports detailed only two synonymous variants in PLAG1. Additional exonic variants propagating at low frequency or that have been fixed in some breeds can be identified by sequencing diverse breeds. Exploring gene networks involving PLAG1-CHCHD7 locus can further help understand the (direct / indirect) role of these genes in the diversity of stature in cattle.

Table 1: Summary of selection regions and number of genetic variants in candidate genes

BTA: region (Mb)	Candidate genes		Illumina 50K SNPs (n)	Genetic variants (n) from Ensembl data			
	Gene ID	Location (Mb)		5´UTR	Intronic	Exonic	3´UTR
13:63.9-65.7	UQCC	65.233–65.327	3 (intronic)	2	$273+1^{SR}$	$3^{MS}+1^{SN}$	6
	GDF5	65.340–65.344	-	-	10	$2^{MS}+2^{SN}$	-
14:24.4-25.4	PLAG1	25.007–25.009	-	-	4	2^{SN}	-
	CHCHD7	25.052–25.059	-	-	$21+1^{SD}$	-	-

UTR: Untranslated region, SR: Splice region, MS: Mis-sense, SN: Synonymous, SD: Splice donor

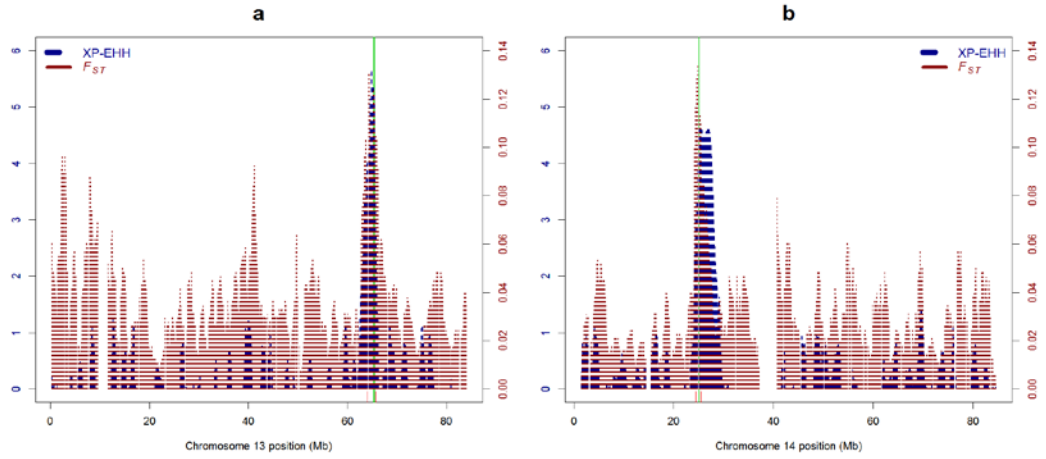


Figure 2: Plot of averaged population differentiation (F_{ST}) and across population extended haplotype homozygosity (XP-EHH) tests between the groups of small-medium and medium-large body sized breeds on a) BTA13 and b) BTA14. Vertical green lines show genic locations and red bars at bottom show candidate regions of significant Mean Fractional Ranks (MFR).

CONCLUSION

By implementing new tools for discovering selection signatures, we demonstrated the localization of candidate genes of major effects on development, skeletal growth and stature in cattle. Our results showed that the complementary signals from constituent statistics of MFR at candidate loci notably improved the resolution of MFR signals in the candidate regions. In addition, the strategy of using multi-breed panels has also contributed towards minimizing the breed-specific unique patterns of diversity in the SNP data. Further investigations of the several non-synonymous variants in the newly identified genes may help to explain the biological function of these mutations in the diversity of bovine stature. Combining selection signature analyses with genome-wide association studies can further improve the fine-mapping of causal mutations controlling stature.

REFERENCES

Allais-Bonnet A., Grohs C., Medugorac I., Krebs S., Djari A., Graf A., Fritz S., Seichter D., Baur A., Russ I., Bouet S., Rothammer S., Wahlberg P., Esquerré D., Hoze C., Boussaha M., Weiss B., Thépot D., Fouilloux M.-N., Rossignol M.-N., van Marle-Köster E., Hreiðarsdóttir G.E., Barbey S., Dozias D., Cobo E., Reversé P., Catros O. *et al*. (2013) *PLoS ONE.* **8**:e63512.

Browning S.R. and Browning B.L. (2007) *Am. J. Hum. Genet.* **81**:1084.

Decker J.E., Pires J.C., Conant G.C., McKay S.D., Heaton M.P., Chen K., Cooper A., Vilkki J., Seabury C.M., Caetano A.R., Johnson G.S., Brenneman R.A., Hanotte O., Eggert L.S., Wiener P., Kim J.-J., Kim K.S., Sonstegard T.S., Tassell C.P.V., Neibergs H.L., McEwan J.C., Brauning R., Coutinho L.L., Babar M.E., Wilson G.A., McClure M.C., Rolf M.M., Kim J., Schnabel R.D. and Taylor J.F. (2009) *Proc. Natl. Acad. Sci.* **106**:18644.

Gautier M., Laloë D. and Moazami-Goudarzi K. (2010) *PLoS ONE.* **5**:e13038.

Karim L., Takeda H., Lin L., Druet T., Arias J.A.C., Baurain D., Cambisano N., Davis S.R., Farnir F., Grisart B., Harris B.L., Keehan M.D., Littlejohn M.D., Spelman R.J., Georges M. and Coppieters W. (2011) *Nat. Genet.* **43**:405.

Kemper K.E. and Goddard M.E. (2012) *Hum. Mol. Genet.* **21**:R45.

Lettre G., Jackson A.U., Gieger C., Schumacher F.R., Berndt S.I., Sanna S., Eyheramendy S., Voight B.F., Butler J.L., Guiducci C., Illig T., Hackett R., Heid I.M., Jacobs K.B., Lyssenko V., Uda M., Boehnke M., Chanock S.J., Groop L.C., Hu F.B., Isomaa B., Kraft P., Peltonen L., Salomaa V., Schlessinger D., Hunter D.J., Hayes R.B., Abecasis G.R., Wichmann H.E., Mohlke K.L. and Hirschhorn J.N. (2008) *Nat. Genet.* **40**:584.

Matukumalli L.K., Lawley C.T., Schnabel R.D., Taylor J.F., Allan M.F., Heaton M.P., O'Connell J., Moore S.S., Smith T.P.L., Sonstegard T.S. and Tassell C.P.V. (2009) *PLoS ONE.* **4**:e5350.

Nishimura S., Watanabe T., Mizoshita K., Tatsuda K., Fujita T., Watanabe N., Sugimoto Y. and Takasuga A. (2012) *BMC Genet.* **13**:40.

Pryce J.E., Hayes B.J., Bolormaa S. and Goddard M.E. (2011) *Genetics.* **187**:981.

Randhawa I.A.S., Thomson P.C., Khatkar M.S. and Raadsma H.W. (2013) *Proc. Assoc. Advmt. Anim. Breed. Genet.* **20**:(in press).

Sanna S., Jackson A.U., Nagaraja R., Willer C.J., Chen W.-M., Bonnycastle L.L., Shen H., Timpson N., Lettre G., Usala G., Chines P.S., Stringham H.M., Scott L.J., Dei M., Lai S., Albai G., Crisponi L., Naitza S., Doheny K.F., Pugh E.W., Ben-Shlomo Y., Ebrahim S., Lawlor D.A., Bergman R.N., Watanabe R.M., Uda M., Tuomilehto J., Coresh J., Hirschhorn J.N., Shuldiner A.R., Schlessinger D., Collins F.S., Smith G.D., Boerwinkle E., Cao A., Boehnke M., Abecasis G.R. and Mohlke K.L. (2008) *Nat. Genet.* **40**:198.

Visscher P.M. and Goddard M.E. (2011) *Nat. Genet.* **43**:397.

Voight B.F., Kudaravalli S, Wen X and Pritchard J. K. (2006) *PLoS Biol.* **4**:e72.

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:459-462

GENETIC ANALYSIS OF ABSENCE OF BREECH STRIKE AND BREECH STRIKE INDICATOR TRAITS IN SOUTH AFRICAN MERINO SHEEP

A.J. Scholtz[1], S.W.P. Cloete[1,2], J.B. van Wyk[3] and E. du Toit[4]

[1]Directorate: Animal Sciences, Elsenburg 7607. South Africa
[2]Department of Animal Sciences, University of Stellenbosch, Matieland 7602, South Africa
[3]Department of Animal Sciences, University of the Free State, Bloemfontein 9300, South Africa
[4]Directorate: Animal Sciences, Riviersonderend 7250, South Africa

SUMMARY

The sheep blowfly is an economically important ectoparasite of sheep, and impacts on animal health and welfare under pastoral conditions in South Africa. The absence of flystrike in the breech (ABS) was recorded in 2198 Merino hoggets on the Tygerhoek Research Farm. This trait was analysed together with dag score (DS; n=1623); neck wrinkle score (NWS; n=2162); midrib wrinkle score (MWS; n=2162) and breech wrinkle score (BWS; n=2162) in a five-trait threshold model, using Bayesian inference and Gibbs sampling. ABS occurred at a frequency of 0.89. Estimates of direct heritability (s.e.) amounted to 0.20 (0.06) for ABS, 0.24 (0.07) for DS, 0.31 (0.08) for NWS, 0.34 (0.06) for MWS and 0.36 (0.08) for BWS. The genetic correlation of ABS with DS was favourable at -0.47 (0.20) and significant ($P < 0.05$). The genetic correlations of ABS with NWS, MBS and BWS were all favourable at -0.78 (0.21), -0.74 (0.16) and -0.78 (0.18) respectively (all $P < 0.05$). Genetic correlations among wrinkle scores (WS) were all high (> 0.88). Selection for a reduction in DS and in wrinkles on any part of the body will benefit ABS.

INTRODUCTION

Mulesing in South Africa was officially banned in 2009 as a method to control breech strike in woolled sheep when the South African National Wool Grower's Association (NWGA) and the National Society for the Prevention of Cruelty to Animals (NSPCA) jointly announced: *'The practice of mulesing is cruel and causes pain and stress to the animal and is a contravention of the Animal Protection Act no. 71 of 1962'* (NWGA 2009). With the cessation of mulesing world-wide as a method to control breech strike, studies have been increasingly directed towards genetic alternatives for the prevention of breech strike (see references in Table 3). The susceptibility of a flock to breech strike is strongly associated with the degree of breech wrinkles and/or dags. It is conceded that emphasis on each of the latter traits could vary dependent upon the environment and type of sheep. Other traits, such as breech cover scores, urine stain and wool colour scores were also identified as potential indicator traits for the prevention of breech strike. Genetic parameters for these indicator traits are available in the literature; indicating that they do exhibit additive genetic variation (see Table 3 for recent estimates on dag and wrinkle scores). This paper reports on genetic parameters for the latter breech strike indicator traits under South African conditions.

MATERIALS AND METHODS

Records were obtained from Merino sheep maintained on the Tygerhoek Research Farm, near Riviersonderend in the Western Cape Province, from 2003 to and including 2010. The climate at this site is Mediterranean, with approximately 60% of the annual rain of 425 mm expected from April to September. Peaks in blowfly activity are expected during October-November and again in February-March. Ewes were mated during October-November to lamb during March-April of the following year, throughout the period of data recording. All progeny born were tail docked at the third palpable joint before they were four weeks old. Lambs were shorn as weaners in August-September of each year, crutched in February and shorn again as hoggets in August-September of

the following year when they had 1 year of wool growth. Weaners were maintained in single flocks (separated on gender) after weaning. All progeny were subjected to visual appraisal of wrinkle score (WS) on the neck (NWS), midrib (MWS) and breech (BWS) at an age of approximately 16 months (Dun and Hamilton 1965). WS was assessed on a 1-6 scale (with 1 being lowest level), which differs from the method of Mortimer *et al*. (2009) and others which only used a 1-5 scale. Progeny were also visually appraised for dag score (DS) just before being shorn as hoggets. The latter trait was scored on a 5-point scale, with 5 being the highest level. DS and BWS are considered as indicator traits for ABS during selection. ABS was recorded in all animals during the 1-year wool growth period between weaner and hogget shearing (Cloete *et al*. 2001). These records were confirmed at hogget shearing in August-September, where needed. Routine management for the prevention of flystrike included the prophylactic treatment of all short-wool animals during November-December. Spot treatment with a long-acting chemical was administered to those sheep suffering from breech strike after the strike had been recorded. Individual sheep were recorded as either having contracted breech strike or not, (i.e. the distribution was binomial). Data were available on a total of 2198 animals. No ABS or WS data were recorded for the progeny group of 2004, and data for this year were excluded. Recording of DS data (n=1623) commenced later, starting with the progeny group of 2006.

Statistical Analysis: A five-trait threshold animal model, applying Bayesian inference and Gibbs sampling in THRGIBBS1F90 software were used (Misztal 2008) to estimate additive genetic variances for each trait. The analysis involved 300000 samples, 50000 of which formed the burn-in period. Every 10^{th} sample of the subsequent 250000 samples was used to calculate posterior means and posterior standard deviations depicting the genetic and environmental (co)variances. Post-Gibbs analysis was done with POSTGIBBSF90 (Misztal 2008).

RESULTS AND DISCUSSION

Distributions for DS and WS approximated normality, but all were treated as threshold traits (Table 1). The highest WS was recorded on the neck of the animals. ABS occurred at a frequency of 0.89.

Table 1. Descriptive statistics for the traits included in the five-trait analysis

Trait	Number	Mean±s.d.	Range	Skewness	Kurtosis
Dag score	1623	2.13 ± 1.17	1-5	0.84	-0.22
Absence of breech strike	2198	0.89 ± 0.31	0-1	4.49	-2.55
Neck wrinkle score	2162	3.73 ± 0.90	1-6	0.23	0.12
Midrib wrinkle score	2162	3.13 ± 0.73	1-6	0.56	0.61
Breech wrinkle score	2162	3.22 ± 0.76	1-6	0.79	1.50

All traits were moderately heritable, estimates ranging from 0.20 for ABS to 0.36 for BWS (Table 2). DS was moderately heritable at 0.24. Estimates from the literature (last 5 years) for DS varied widely, ranging from 0.08 to 0.63 (Table 3), a range that included the present estimate. ABS on the underlying scale was also moderately heritable at 0.20. Preventative chemical treatment and crutching applied to the animals could protect animals against breech strike for up to 12 weeks. Strikes occurring in the remainder of the period, albeit sporadic in some years, were sufficient for genetic variation to be detected. This opens up the possibility of increasing ABS by direct selection, thereby reducing reliance on chemicals for prevention and for spot treatment of strikes, as well as the need for the Mules operation. Estimates for the heritability of breech strike were generally higher than the present estimate, ranging from 0.32 to 0.57 (Table 3). The present heritability estimate for NWS (0.31) was slightly below estimates amounting to 0.42 in the

literature (Table 3). The heritability estimate for MWS (0.34) is comparable to recent estimates of 0.25 to 0.42 (Table 3). Heritability estimates for BWS ranged from 0.35 to 0.69 in the literature (Table 3). The heritability estimate of 0.36 reported in this study is on the lower boundary of this range. Results from this and other studies suggest that heritability estimates for WS on the respective body regions are high enough to support substantial selection gains.

Table 2. Phenotypic variance components (σ^2_P) and (co)variance ratios (±s.e.) for dag score (DS), absence of breech strike (ABS), neck wrinkle score (NWS), midrib wrinkle score (MWS) and breech wrinkle score (BWS)

Trait	DS	ABS	NWS	MWS	BWS
Variance ratios and posterior standard deviations (PSD)					
σ^2_P	1.629	1.260	0.854	0.345	0.609
(Co)variance ratios (heritability in bold on the diagonal, r_g above the diagonal and r_e below the diagonal)					
DS	**0.24 ± 0.07**	-0.47 ± 0.20	0.03 ± 0.15	0.17 ± 0.16	0.09 ± 0.015
ABS	0.18 ± 0.07	**0.20 ± 0.06**	-0.78 ± 0.21	-0.74 ± 0.16	-0.78 ± 0.18
NWS	0.03 ± 0.05	0.01 ± 0.06	**0.31 ± 0.08**	0.95 ± 0.17	0.89 ± 0.15
MWS	-0.05 ± 0.06	0.01 ± 0.07	0.79 ± 0.09	**0.34 ± 0.06**	0.95 ± 0.18
BWS	0.09 ± 0.06	-0.04 ± 0.07	0.62 ± 0.08	0.71 ± 0.10	**0.36 ± 0.08**

Table 3. Heritability estimates (h^2) from the literature (last 5 years) for post-weaning (unless indicated otherwise) dag score (DS), absence/presence of breech strike (ABS), neck wrinkle score (NWS), midrib wrinkle score (MWS) and breech wrinkle score (BWS)

Trait	**h^2**	**Reference**	**Comment**
DS	0.09 ± 0.06	Smith *et al*. (2009)	6 months old (post-weaning stage)
	0.28 ± 0.02	Brown *et al*. (2010)	Late (yearling and hogget age)
	0.31 ± 0.01	Pickering *et al*. (2010)	Lambs (8 months)
	0.08 ± 0.07	Scobie *et al*. (2011)	Yearlings
	0.63 ± 0.08	Greeff *et al*. (2013)	Yearlings
	0.37 ± 0.05	Greeff *et al*. (2013)	Hoggets
ABS	0.46 ± 0.23	Scholtz *et al*. (2010)	15 – 16 month unmulesed animals
	0.57 ± 0.28	Greeff and Karlsson (2009)	Birth – hogget age; threshold trait
	0.32 ± 0.11	Smith *et al*. (2009)	Weaners; continuous trait
	0.51 ± 0.10	Greeff *et al*.(2013)	Birth – hogget age; threshold trait
NWS	0.42 ± 0.01	Mortimer *et al*. (2009)	15 - 16 month old ewes
	0.42 ± 0.12	Scholtz *et al*. (2010)	15 – 16 months unmulesed animals
MWS	0.25 ± 0.10	Smith *et al*. (2009)	6 months old (post weaning stage)
	0.42 ± 0.01	Mortimer *et al*. (2009)	15 - 16 month old ewes
	0.30 ± 0.10	Scholtz *et al*. (2010)	15 – 16 month unmulesed animals
BWS	0.36 ± 0.12	Smith *et al*. (2009)	6 months old (post-weaning stage)
	0.69 ± 0.05	Brown *et al*. (2010)	Late (yearling and hogget age)
	0.45 ± 0.13	Scholtz *et al*. (2010)	15 – 16 month unmulesed animals
	0.35 ± 0.06	Greeff *et al*. (2013)	Yearlings
	0.50 ± 0.08	Greeff *et al*. (2013)	Hoggets

The occurrence of flystrike depends on weather conditions, which are often transient and unpredictable. Selection gains under such conditions are often difficult to achieve, even without the added complication of ABS being evaluated on the binomial scale and the fact that adequate blowfly challenge impinges on the welfare of animals. Significant heritability estimates reported here and in the literature suggest that direct selection for ABS will be successful under adequate

challenge. Selection responses may, however, be slow under conditions of suboptimal challenge. The unpredictability of flystrike, the widespread use of prophylactic treatments such as jetting and crutching, as well as animal welfare concerns under adequate challenge conditions, adds to arguments for indirect selection instead of direct selection for ABS (Scholtz *et al*. 2010).

Genetic correlations of ABS with the indicator traits were all favourable and larger than twice the corresponding standard error, ranging from -0.47 for DS to -0.78 for NWS and BWS (Table 2). Wrinkly sheep are thus more susceptible to breech strike than their plainer contemporaries. Genetic correlations between breech strike and BWS amounted to 0.23 (Greeff and Karlsson 2009) as well as to 0.27 (yearling age) and 0.13 (hogget age) (Greeff *et al*. 2013). The difference in sign between the latter estimates and those in the present study stems from the ABS being analysed in the present study, whereas the latter authors studied the incidence of breech strike.

The genetic correlations among WS on different body parts were all higher than 0.88 and approached unity in some cases (Table 2), suggesting that WS is effectively the same trait irrespective of body location. These estimates are consistent with earlier genetic correlations exceeding 0.90 among WS on different body regions (Jackson and James 1970; Mortimer and Atkins 1993, Mortimer *et al*. 2009), confirming these traits to be genetically very similar.

CONCLUSIONS

It is evident that selection for a reduction in DS and wrinkle score on any part of the body will benefit ABS in this environment in South Africa. Further research on the incorporation of direct and indirect selection for ABS as part of an integrated blowfly management programme is needed, to ensure that the problem is dealt with in a sustainable manner while simultaneously ensuring that the welfare of animals is not compromised.

REFERENCES

Brown D.J., Swan A.A. and Gill J.S. (2010). *Anim. Prod. Sci.* **50**: 1060.

Cloete S.W.P., Olivier J.J. and Du Toit E. (2001) Proc. FLICS, flystrike and lice conference. Ed. S. Champion. University of Tasmania, pp. 395.

Dun R.B. and Hamilton B.A. (1965) *Aust. J. Exp. Agric. Anim. Husb.* **5**: 236.

Greeff J.C. and Karlsson L.J.E. (2009) *Proc. Assoc. Advmt. Anim. Breed. Genet.* **18**: 272.

Greeff J.C., Karlsson L.J.E. and Schlink A.C. (2013) *Anim. Prod. Sci.* Online: http://dx.doi.org/10.1071/AN12233

Jackson N. and James J.W. (1970) *Aust. J. Agric. Sci.* **21**: 837.

Misztal I. (2008) *J. Anim. Breed. Gen.* **125**: 363.

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.

NWGA (National Woolgrower's Association) (2009) Wool Farmer July 2009, p. 5.

Pickering, N.K., Dodds, K.G., Blair, H.T., Hickson, R.E., Johnson, P.L. and McEwan, J.C. (2010) Online: http://www.kongressband.de/wcgalp2010/assets/pdf/0294.pdf

Scholtz A.J., Cloete, S.W.P., Van Wyk J.B., Misztal I., Du Toit E. and Van der Linde T.C.de K. (2010) *Anim. Prod. Sci.* **50**: 210.

Scobie D.R., Hickey S.M., Maslen D.P. and the late Black G.M. (2011) *Proc. N.Z. Soc. Anim. Prod.* **71**: 251.

Smith J.L., Brewer H.G. and Dyall T. (2009) *Proc. Assoc. Advmt. Anim. Breed. Genet.* **18**: 334.

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:463-466

USING CROSS-VALIDATION IN A FAST EM ALGORITHM FOR GENOMIC SELECTION AND COMPLEX TRAIT PREDICTION

R.K. Shepherd[1], M.J. Drumm[1] and J. Yang[2]

[1]School of Engineering & Technology, Central Queensland University, Rockhampton QLD 4702
[2]The University of Queensland Diamantina Institute, Woolloongabba, QLD 4102

SUMMARY

This paper reports on changes to the EM algorithm emBayesB which estimates QTL effects using dense genome-wide SNP marker data. To overcome convergence issues, modifications were made to the original algorithm which included cross-validation for the estimation of model parameters. The modified algorithm called emBayesB_CV was used to analyse a trait simulated on real human genotypes consisting of 294,831 SNP measured on 3925 individuals. Three datasets were simulated for a trait determined by 10, 100 or 1000 additive QTL. The results showed that the modified algorithm emBayesB_CV was not only computationally fast, but also more accurate than GBLUP in predicting breeding value. However prediction accuracy declined as the size of QTL effects decreased due to the result that although emBayesB_CV could accurately locate the chromosomal location of large QTL effects, this was not the case for small QTL effects.

INTRODUCTION

Genomic prediction of breeding values is a new tool for selection in livestock breeding programs and for risk prediction with complex human diseases. In animal breeding genomic selection uses information from high-density genome-wide SNP markers to predict the breeding value of candidates for selection. Firstly SNP effects have to be estimated in the population of interest by analysing the relationship between phenotype and the SNP genotypes (called training). Then genomic estimated breeding values (GEBV) are calculated by summing the estimated SNP effects across the genome of each candidate. Usually the accuracy of GEBV is assessed in an independent dataset by calculating the correlation between GEBV and either True Breeding Value (TBV) or phenotype (called validation). Bayesian models can be used to include important prior beliefs about the QTL effects and are usually more accurate than BLUP methods using the realised relationship matrix (called GBLUP). But Bayesian prediction is computationally slow for large SNP panels, whereas GBLUP is much faster. emBayesB is an Expectation Maximisation (EM) algorithm which not only incorporates important prior information about QTL effects, but is also computationally fast like GBLUP. However convergence issues are known to occur with emBayesB unless arbitrary bounds are placed on the estimated parameter of the SNP effect distribution like in Shepherd *et al*. (2010). This paper investigates cross-validation for parameter estimation in addition to other modifications to the emBayesB algorithm.

MATERIALS AND METHODS

EM theory. Full details are in Shepherd *et al*. (2010). If we knew which SNP were in linkage disequilibrium (LD) with QTL, then the problem would be much easier. So we assume *a priori* that a fraction γ of the SNP are in LD with QTL and that SNP in LD with QTL have effects from a double exponential (DE) distribution with parameter λ. A linear data model $\mathbf{y} = \mathbf{Bg} + \mathbf{e}$ is assumed to relate phenotype y_i of individual *i* to the j^{th} SNP effect g_j where element b_{ij} of the $n \times m$ matrix $\mathbf{B}$ is the number (0, 1 or 2) of reference alleles (usually standardised) of SNP *j* for individual *i*. The errors are assumed normal and independent such that $\mathbf{y} \mid \mathbf{g} \sim N\left(\mathbf{Bg}, \mathbf{I}\sigma_e^2\right)$.

Using EM theory an iterative sequence of E and M-steps is developed which should converge to maximum *a posteriori* (MAP) parameter estimates. At iteration *k*, the E-step involves the calculation of p_j^k, the posterior probability that SNP *j* is in LD with QTL given the data and all current parameter estimates. This is done analytically and fast. Then given the data and the current values of p_j^k, the M-step calculates $\hat{g}_j = p_j^k DE_{j(\text{mode})}$, $\hat{\gamma} = \frac{1}{m}\mathbf{1}'\mathbf{p^k}$, $\hat{\lambda} = \mathbf{1}'\mathbf{p^k} \Big/ \left|\hat{\mathbf{g}}\right|' \mathbf{p^k}$ and $\hat{\sigma}_e^2 = \frac{1}{n}\left(\mathbf{y} - \mathbf{B}\hat{\mathbf{g}}\right)'\left(\mathbf{y} - \mathbf{B}\hat{\mathbf{g}}\right)$ where $\mathbf{p^k}$ is the vector of posterior probabilities at iteration k and $DE_{j(\text{mode})}$ is the posterior mode of g_j conditional on all current estimates using a DE prior only. Iterating between the E and M-steps the algorithm should converge quickly to produce MAP estimates of $\mathbf{g}$, posterior probabilities $\mathbf{p}$ and ML estimates of $\gamma, \lambda, \sigma_e^2$.

emBayesB_CV. Modifications were made to the original algorithm due to convergence issues. Firstly a complete E-step was performed before updating estimates of each SNP effect g_j. This was not done in the original algorithm. Also the total genetic variance in a dataset was estimated using GBLUP and the estimate of h^2 was fixed in emBayesB_CV. Then the parameter γ was estimated by *k*-fold cross-validation in the training data, with λ being determined by the equation $\lambda^2 = 2m\gamma \Big/ h^2\sigma_y^2$ using the fixed values of h^2 and γ. To speed up convergence, γ was updated each iteration using the M-step equation $\hat{\gamma} = \frac{1}{m}\mathbf{1}'\mathbf{p^k}$ with corresponding updates calculated for λ.

Data simulation. Genotypes of 3925 unrelated humans from GWA Studies were used as described in Yang *et al*. (2010). SNP were randomly selected (from the 294,831 SNP available) to be biallelic QTL and then removed as SNP in the analysis. Three datasets were simulated consisting of 10, 100 or 1000 additive QTL which meant the number of SNP used in each analysis was 294821, 294731 and 293831 respectively. QTL effects were independently simulated from a normal distribution and summed to produce the TBV of each individual. A trait with heritability 0.8 was produced by adding a normally distributed error term to the TBV of each individual. The number of QTL, which explain more than 0.1, 1, 5 and 10% of the total phenotypic variation, was 9, 8, 7 and 5 respectively in the 10 QTL dataset, whereas the number of QTL was 83, 29, 4 and 0 respectively in the 100 QTL dataset. For the 1000 QTL dataset, the number of QTL, which explain more than 0.1, 1 and 5% of the total phenotypic variation, was 297, 4 and 0 respectively.

Statistical analysis. The dataset for each of the three QTL scenarios, was initially split into a training set and a validation set consisting of 3500 and 425 records respectively. The value of γ was estimated using 5-fold cross-validation in the training set. For each γ value, the prediction equation $\mathbf{GEBV = B\hat{g}}$ was estimated using 4 folds (2800 individuals) and then used to calculate GEBV in the left out fold (700 individuals). This procedure was repeated 5 times, so that each fold was left out once, in order to produce GEBV for all 3500 individuals. The value of γ which maximised the correlation between GEBV and phenotype in the training data was chosen. Then this value of γ was used to estimate the SNP effects $\hat{\mathbf{g}}$ using all 3500 training records. To validate the estimated SNP effects $\hat{\mathbf{g}}$, the correlation between TBV and GEBV was calculated for the 425 validation records, as well as the linear regression of TBV on GEBV, which has a slope of 1 if the GEBV are unbiased. The SNP effects were also estimated by GBLUP in the training set

and then validated in the validation dataset. For GBLUP the estimated SNP effects were solutions to the training set equations $(\mathbf{B}'\mathbf{B} + \alpha\mathbf{I})\hat{\mathbf{g}} = \mathbf{B}'\mathbf{y}$ where $\alpha = \sigma_e^2 / \sigma_g^2 = m(1 - h^2) / h^2$.

RESULTS AND DISCUSSION

Using cross-validation to estimate γ, fixing h^2 and using separate E & M-steps solved the problem of convergence which can occur with emBayesB_CV. In Table 1 each emBayesB_CV run uses 5-fold cross-validation with an initial fixed γ and takes approximately 30 minutes on a single compute node of the High Performance Computing (HPC) facility at CQUniversity. Searching for the optimum γ usually took between 10 and 20 such runs which can be run simultaneously on a multiple node HPC facility. Hence for large SNP panels, emBayesB_CV will be significantly faster than a full Bayesian analysis as found by Shepherd *et al*. (2010).

Table 1. Correlation between GEBV and TBV ($\mathbf{r}_{TBV}$) for the validation data of GBLUP and emBayesB_CV using the 10, 100 or 1000 QTL datasets. b is the linear regression coefficient of TBV on GEBV while h^2 is the fixed heritability. The estimated number of SNP in LD with QTL ($m\hat{\gamma}$) is given as well as the estimated parameter $\hat{\lambda}$ of the SNP effect distribution.

No.	No.	Fixed	GBLUP	emBayesB_CV		
QTL	SNP (m)	h^2	$\mathbf{r}_{TBV}$ (b)	$\mathbf{r}_{TBV}$ (b)	$m\hat{\gamma}$	$\hat{\lambda}$
10	294821	0.8	0.15 (0.7)	0.88 (1.0)	93	10.5
100	294731	0.7	0.13 (0.6)	0.74 (1.0)	163	2.9
1000	293831	0.6	0.21 (0.9)	0.33 (0.8)	203	1.3

Table 1 shows the correlation between GEBV and TBV for the validation data. emBayesB_CV was significantly more accurate than GBLUP. The poor performance of GBLUP was due to the fact that unrelated individuals were chosen in the original dataset (Yang *et al*. 2010). Livestock populations have high levels of relatedness and so GBLUP would be expected to do much better in livestock populations. emBayesB_CV was most accurate when there were 10 simulated QTL ($r_{TBV} = 0.88$) with the accuracy decreasing to 0.74 and 0.33 for 100 and 1000 QTL respectively (Table 1). This decline in accuracy is not unexpected. As the number of QTL increase, the size of individual QTL effects decrease and thus it becomes more difficult to detect the location of the QTL as shown in Figure 1. Figure 1A shows chromosome 6 which has the two largest QTL effects in the 10 QTL dataset. It can be seen that the location of the 2 QTL is accurately determined by SNP with large posterior probabilities. However Figure 1B shows that the 69 small QTL effects on chromosome 4 in the 1000 QTL dataset, are not accurately located by SNP with large posterior probabilities. In fact there are only 4 SNP with posterior probabilities greater than 0.9 on chromosome 4 and none are located close to QTL.

Although emBayesB_CV was not able to accurately locate QTL with small effects, it was able to predict aggregate breeding value more accurately than GBLUP by detecting SNP in regions of trait variation. For example, in the 1000 QTL dataset, the chromosome explaining the largest percentage of the phenotypic variance (6.7%) was chromosome 4 which contained 69 QTL, and 13 detected SNP with posterior probabilities greater than 0.05 of being in LD with QTL (Figure 1B). On the other hand the chromosome explaining the smallest percentage of the phenotypic variance (0.9%) in the 1000 QTL dataset, was chromosome 21 which contained 13 QTL, and only 4 detected SNP with posterior probabilities greater than 0.05 of being in LD with QTL.

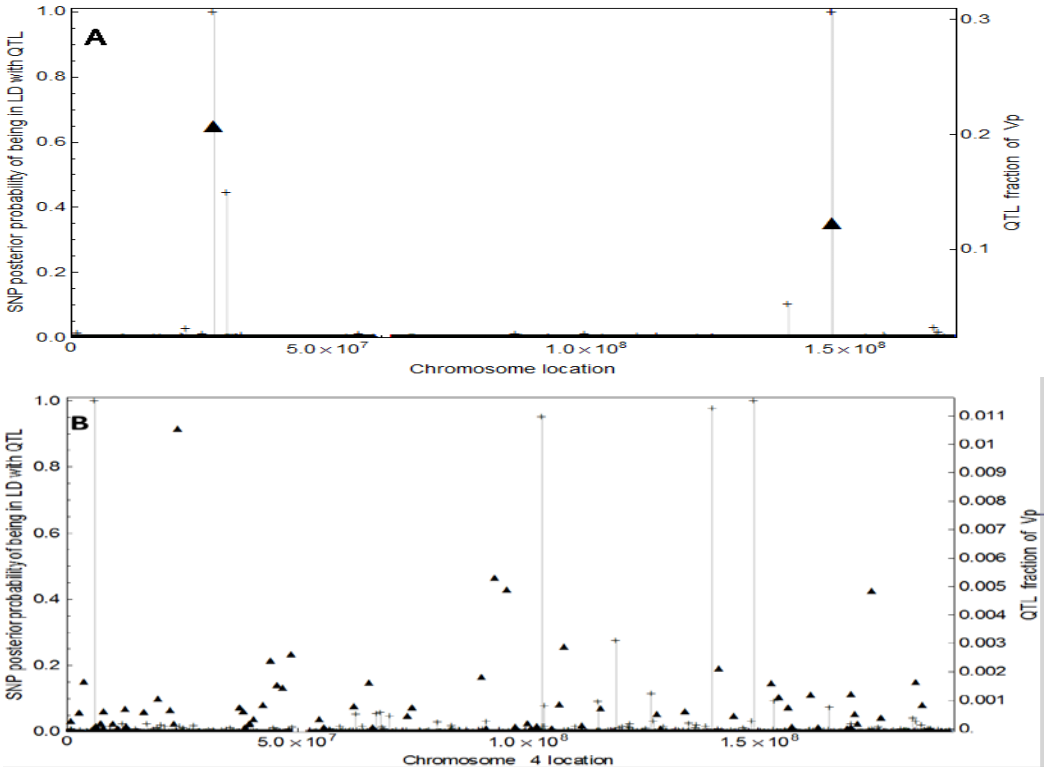


Figure 1. Posterior probability (+) of a SNP being in LD with QTL and the fraction of the total phenotypic variance (▲) explained by a QTL on chromosome 6 for the 10 QTL dataset (A) and on chromosome 4 for the 1000 QTL dataset (B).

The number of SNP with posterior probabilities greater than 0.5 of being in LD with QTL, was 51, 76 and 49 for the 10, 100 and 1000 QTL datasets. Also using the formula:

$No.\ SNP\ in\ LD = m\hat{\gamma}$ where $\hat{\gamma}$ estimates the proportion of SNP in LD with QTL,

Table 1 shows that emBayesB_CV predicts 93, 163 and 203 SNP are in LD with QTL for the 10, 100 and 1000 QTL datasets. This shows the difficulty emBayesB_CV has in detecting QTL as the effects get smaller. Increasing the number of genotyped individuals will help detect smaller QTL effects.

CONCLUSIONS

emBayesB_CV is a computationally fast method of predicting breeding value using dense genome-wide SNP marker data which was significantly more accurate than GBLUP for the scenarios investigated in this paper. emBayesB_CV overcame the convergence issue which often occurs with emBayesB. The chromosomal location of large QTL effects can be accurately located with emBayesB_CV. But this is not the case for QTL of small effect.

REFERENCES

Shepherd R. K., Meuwissen T.H.E. and Woolliams J.A. (2010) *BMC Bioinformatics* **11**: 529
Yang J., Benyamin B., McEvoy B. P., Gordon S., *et al.* (2010) *Nature Genetics* **42**: 565

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:467-470

POST-WEANING GROWTH IN BEEF AND DAIRY CROSSBRED STEERS

C. G. Vazquez[1], R.E. Hickson[2], S.T. Morris[2], N. Lopez-Villalobos[2], P.R. Kenyon[2] and J.G. García-Muñiz[3]

[1]Facultad de Medicina Veterinaria y Zootecnia, UNAM Ciudad Universitaria. 04510. México
[2]Massey University, Private Bag 11222, Palmerston North 4442, New Zealand.
[3]Departamento de Zootecnia, Universidad Autónoma Chapingo, Chapingo, Estado de México.

SUMMARY

A study was performed to evaluate post-weaning growth of 78 steers sired by straight-bred Hereford bulls (HH) over: straight-bred Angus dams (HHxAA, n=25), Angus-cross-Friesian dams (HHxAF, n=21); Angus-cross-Jersey dams (HHxAJ, n=21) and Angus-cross-Kiwicross (Friesian-Jersey) dams (HHxAK, n=11). The steers were divided into two groups for slaughter at 666 days of age (n=38) and at 763 days of age (n=40). Live weight (LWT) was measured monthly from weaning (168 days) to slaughter. Rib fat depth (RF) and body condition score (BCS; 1-5 scale) were measured every three months from weaning, and height at withers (HT) was measured at 12, 18 and 24 months of age. A longitudinal mixed model was used to compare breed group and slaughter season as main effects. At weaning, HHxAF steers were heaviest (237.4±1.4 kg), whilst HHxAK and HHxAJ were similar and intermediate to the other two breed groups (228.8±1.8 kg and 225.6±1.4 kg, respectively) and HHxAA steers were lightest (209.7±1.3 kg). At 21 months of age, HHxAF steers (555.6±7.5 kg) remained heavier than HHxAK and HHxAJ steers (515.8±5.4 and 520.7±4.1 kg, respectively), but HHxAA steers (532.6±6.6 kg) were similar (P>0.05) to all other breed groups. There were no differences (P>0.05) among breed groups for RF and BCS. Breed-group differences for HT existed at 12 and 18 months but not at 24 months of age. Measurements of LWT were strongly phenotypically correlated over time, and moderately phenotypically correlated with measurements of BCS and HT made at the same age.

INTRODUCTION

The use of beef-cross-dairy heifers as breeding cows in beef herds can increase the weight of calf weaned compared with straight-bred beef cows (Hickson *et al.* 2011). Weight of calf weaned is commonly used as an indicator of cow productivity, however, the ultimate product produced in the beef industry is a finished steer ready for slaughter rather than a weaned calf. Beef-cross-dairy breeding cows are likely to have inferior direct genetics for growth and finishing compared with beef straight-bred cows; the increased weaning weight of their calves is likely the result of the cow's increased milking ability rather than the calf's own growth potential. Therefore, the effect on the growth of the calf of being born to a beef-cross-dairy dam compared to a straight-bred beef dam should be examined right up to slaughter at around 2 years of age, rather than just in the pre-weaning period.

The objective of this study was to evaluate post-weaning performance of live weight, rib fat, body condition score and height at withers in four breed-groups of beef and dairy-crossbred steers using random regression models.

MATERIALS AND METHODS

Animals and measurements. Heifers of four breed groups: straight-bred Angus (AA), Angus cross Friesian (AF), Angus cross Jersey (AJ), and Angus cross Kiwicross (Friesian-Jersey and Jersey-Friesian; AK) were bred to straight-bred Hereford (HH) bulls at 16 months of age in December 2009. Male progeny were castrated at approximately 6 weeks of age, and were weaned from their dams at an average age of 168 days. The breed groups of the steers were: HHxAA (n =

25), HHxAF (n = 21), HHxAJ (n = 21) and HHxAK (n = 11). Prior to weaning, steers were grazed alongside their dams in four herds based on calving date of the dam and balanced for breed group. From weaning until 581 days of age, all steers were grazed in one herd under commercial management at Massey University's Tuapaka farm (15 km east of Palmerston North, New Zealand). At day 581, steers were allocated to either the 666-day or 763-day slaughter group, based on live weight (the heavier animals were allocated to the 666-day group) so that half of each breed group was included in each slaughter group. Steers slaughtered at 763 days of age (n=40) were moved to Massey University's Riverside farm (10 km north of Masterton, New Zealand) at day 581, where they grazed in one herd until slaughter. Steers slaughtered at 666 days of age (n=38) remained in one herd at Tuapaka until slaughter.

Live weight (LWT) was measured monthly from weaning (168 days average age) until slaughter. Rib fat depth (RF) was an ultrasound measurement of subcutaneous fat depth over the *M. longissimus* between the 12th and 13th ribs and was measured by the same commercial ultrasound technician each time. Body condition score (BCS) was assessed on a scale of 1-5 (1=emaciated, 5=obese) by one technician from weaning to 1 year of age, and another technician (trained by the first technician) from 15 months until slaughter. Rib fat depth and BCS were recorded every three months from weaning until slaughter. Height at withers (HT) was measured at 12 and 18 months of age for all steers and at 24 months of age for steers slaughtered at 763 days of age.

Statistical analyses. Data were analyzed using SAS version 9.3 (SAS Institute Inc., Cary, NC, USA). A repeated measures mixed model with unstructured and heterogeneous residual variance was used to assess the effects of rearing herd, breed group and slaughter group.

Random regression models using a Legendre orthogonal polynomial (Kirkpatrick *et al.* 1990) of 1st order $Y_t = \alpha_0 P_0(x)+\alpha_1 P_1(x)$ for RF and BCS, and 2nd order $Y_t = \alpha_0 P_0(x)+\alpha_1 P_1(x)+\alpha_2 P_2(x)$ for LWT and HT, were used to examine the fixed effects of breed and slaughter season, where: $P_0(x)=1$; $P_1(x)=x$; $P_2(x)=½(3x^2-1)$; and $x=[2(t-t_{min})/(t_{max}-t_{min})]-1$ represents the standardised unit of time from -1 to 1 from weaning to slaughter; Y_t is the measurement for each trait at age t, t_{min} is age at weaning and t_{max} is the greatest age recorded from weaning (Schaeffer 2004).

RESULTS

Variation in LWT increased as the animals became heavier. Steers from the HHxAF breed group were the heaviest until 18 months of age ($P<0.05$; Table 1), and were heavier than the HHxAJ and HHxAK steers at both slaughter ages. Steers from the HHxAA breed group were lighter ($P<0.05$) than all other breed groups at weaning and 9 months of age, after which they were similar ($P>0.05$) to the HHxAJ and HHxAK steers. By 21 months of age, HHxAA steers were also similar ($P>0.05$) to HHxAF steers. Furthermore, by 21 months of age, there was no difference for LWT between HHxAA and HHxAF steers. Steers from the HHxAF and HHxAA breed groups grew 5% faster ($P<0.05$) than HHxAJ and HHxAK steers, which had similar growth performance. Rib fat depth and BCS were similar for all breed groups throughout the experiment. Height of HHxAF steers was greater ($P<0.05$) than for HHxAA and HHxAJ steers at 12 and 18 but not 24 months of age.

Table 2 reports the phenotypic correlations of LWT with itself and the other traits across the different ages of measurement. All LWT measurements were strongly correlated with each other, although the correlation was less at 24 months of age than at other ages – probably the result of these correlations being based on only the 40 steers that were slaughtered at 24 months, and because the steers were considerably fatter at 24 months of age, indicating they had moved into a finishing phase of growth. Rib fat depth was moderately positively correlated with LWT up to 18 months of age, but it was not correlated after that. Body condition score was correlated with LWT at the same age throughout the experiment, and BCS at 12 and 15 months of age were correlated

Table 1. Least squares means ±SE for live weight, rib fat, body condition score and height for steers from Hereford bulls over: Angus dams (HHxAA), Angus-cross-Friesian dams (HHxAF), Angus-cross-Jersey dams (HHxAJ), and Angus-cross-Kiwicross dams (HHxAK).

Age (months)	HHxAA	HHxAF	HHxAJ	HHxAK
		Live weight (kg)		
weaning	209.7±1.3^{c}	237.4±1.4^{a}	225.6±1.4^{b}	228.8±1.8^{b}
9	264.9±3.7^{c}	297.9±3.9^{a}	280.2±3.7^{b}	285.9±4.3^{ab}
12	331.8±3.8^{b}	362.3±3.8^{a}	340.3±3.1^{b}	343.3±2.9^{b}
15	398.7±4.4^{b}	426.7±4.7^{a}	400.4±3.0^{b}	400.8±2.7^{b}
18	465.7±5.4^{b}	491.1±5.9^{a}	460.6±3.3^{b}	458.3±3.8^{b}
21	532.6±6.6^{ab}	555.6±7.5^{a}	520.7±4.08^{b}	515.8±5.4^{b}
24	599.6±7.9^{ab}	620.0±9.3^{a}	580.8±5.0^{b}	573.3±7.4^{b}
		Rib Fat depth (mm)		
9	2.49±0.13	2.58±0.15	2.71±0.16	2.82±0.24
12	3.12±0.11	3.20±0.15	3.37±0.15	3.28±0.25
15	3.75±0.11	3.82±0.15	4.03±0.17	3.75±0.29
18	4.38±0.12	4.43±0.18	4.69±0.19	4.22±0.33
21	5.01±0.15	5.05±0.21	5.35±0.23	4.68±0.38
24	5.64±0.18	5.67±0.24	6.01±0.26	5.14±0.42
		Body condition score (1-5 scale)		
9	2.64±0.07	2.81±0.08	2.59±0.08	2.59±0.11
12	2.91±0.05	3.02±0.06	2.83±0.06	2.81±0.08
15	3.17±0.04	3.23±0.05	3.06±0.05	3.03±0.07
18	3.40±0.04	3.45±0.06	3.29±0.05	3.24±0.08
21	3.69±0.06	3.66±0.07	3.52±0.06	3.46±0.10
24	3.95±0.08	3.88±0.09	3.76±0.09	3.68±0.13
		Height at withers (cm)		
12	114.3±0.60^{c}	118.3±0.58^{a}	115.7±0.57^{bc}	118.1±1.17^{ab}
18	125.1±0.48^{b}	128.3±0.62^{a}	125.4±0.48^{b}	127.7±1.37^{ab}
24	135.8±0.74	138.3±1.10	135.1±0.86	137.3±1.83

^{a,b,c} within row, least squares means with different letters are different (P<0.05)

with LWT during the previous 6 months. Height at withers was generally positively correlated with LWT throughout the experiment.

DISCUSSION

Steers from the three dairy-crossbred dams received similar, generous quantities of milk prior to weaning (Hickson *et al.* 2011), thus differences in their LWT at weaning likely reflected differences in their genetic potential for growth. In contrast, the lesser milk yield of straight-bred Angus dams could have restricted the growth to weaning of the HHxAA steers. Once in the post-weaning environment where all steers had the same feed availability, the growth of the HHxAA steers was greater than the other breed groups, perhaps reflecting compensatory gain that allowed them to reach LWT that was not different to any of the other breed groups by slaughter. The relative LWT of the maternal lines indicates that the genetic potential for LWT of the HHxAA and HHxAF steers would be similar (Hickson *et al.* 2011). There is limited literature detailing the post-weaning growth of calves born to beef-cross-dairy cows, so it is valuable to document that the advantage in LWT at weaning of calves from the dairy-type dams was lost by 21 months of age.

Beef-cross-dairy cows have been shown to have lesser body condition than straight-bred beef cows (Hickson *et al.* 2011), but there were no effects of breed group on RF or BCS of the steers in this study, presumably because the steers were of at least 75% beef breeds.

Table 2. Phenotypic correlation coefficients for live weight with rib fat, body condition score and height at various ages for steers from Hereford bulls over: Angus dams (HHxAA), Angus-cross-Friesian dams (HHxAF), Angus-cross-Jersey dams (HHxAJ), and Angus-cross-Kiwicross dams (HHxAK).

	Live weight at age (months)						
	Weaning	**9**	**12**	**15**	**18**	**21**	**24**
Live weight at age (months)							
9	0.93						
12	0.84	0.94					
15	0.54	0.66	0.75				
18	0.56	0.64	0.70	0.74			
21	0.59	0.70	0.78	0.80	0.88		
24	0.25	0.32	0.32	0.62	0.53	0.68	
Rib fat at age (months)							
9	0.33	0.32					
12	0.29	0.29	0.22				
15	0.33	0.38	0.38	0.15ns			
18	0.18ns	0.24	0.23	0.08ns	0.34		
21	0.01ns	0.12ns	0.14ns	0.00ns	0.21ns	0.11ns	
24	0.07ns	0.10ns	0.18ns	0.02ns	0.00ns	0.05ns	0.12ns
BCS at age (months)							
9	0.39	0.48					
12	0.27	0.35	0.41				
15	0.20ns	0.26	0.33	0.37			
18	0.05ns	0.11ns	0.14ns	0.26	0.37		
21	0.00ns	0.06ns	0.12ns	0.22ns	0.23	0.31	
24	-0.33ns	-0.24ns	-0.12ns	0.10ns	-0.13ns	0.05ns	0.33
Height at age (months)							
12	0.67	0.72					
18	0.39	0.42	0.41	0.47	0.47		
24	0.187ns	0.29ns	0.38	0.36	0.29ns	0.47	0.33

Correlations are significant at the P<0.05 level unless indicated with 'ns' (non-significant)

REFERENCES

Hickson R.E., Kenyon P.R., Lopez-Villalobos N., Laven R.L., Fraser K.R., Balcomb C.C.,and Morris S.T. (2011) *Proc. Soc. Sheep and Beef Cattle Veterinarians of the New Zealand Veterinary Association Annual Conf.* 4.22.1

Kirkpatrick M., Lofsvold D. and Bulmer, M. (1990) *Genetics* **124**: 979.

Schaeffer L.R. (2004) *Livestock Production Science* **86**: 35.

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:471-474

UPDATES TO THE NEW ZEALAND NATIONAL BREEDING OBJECTIVE FOR DAIRY CATTLE

P.R. Amer[1], B. Santos[1], T.J. Byrne[1], C. Ludemann[2], B Visser[1], B.L. Harris[3], and J. Bryant[4]

[1]AbacusBio Limited, PO Box 5585, Dunedin
[2]AbacusBio Pty Ltd, 331 Penshurst Rd, Byaduk, VIC, 3301
[3]Livestock Improvement Corporation Ltd, Private Bag 3016, Hamilton, New Zealand
[4]DairyNZ, Private Bag 3221, Hamilton 3240, New Zealand

SUMMARY

The National Breeding Objective for dairy cattle in New Zealand expressed as a genetic selection index called Breeding Worth (BW), assesses sire and cow genetic merit now and sets the direction for the New Zealand cow of the future. A major review of the calculations of the economic weightings that underpin the index has recently been undertaken. A modified approach to the costing of feed had only a modest impact on existing traits in the index, but opens up opportunities to calculate economic values for traits which shift feed requirements from one season to another. Such traits include autumn body condition score and lactation persistency. A further major change to the index related to assumptions about the farmer response to shifts in herd genetic merit for survival. Historically, lower survival was assumed to result in lower voluntary culling, whereas the new model assumes that lower survival will lead to an increased requirement to breed replacement heifers. As a consequence, the economic values for residual survival, fertility, and to a lesser extent somatic cell score have increased substantially. These changes have been generally well received in the industry and have led to noticeable impacts on rankings of AI sires.

INTRODUCTION

The New Zealand dairy industry is supported by a co-ordinated, integrated and comprehensive data recording and genetic evaluation system. New Zealand dairy farmers rely on independent evaluations encapsulated in the breeding worth (BW) of sires The National Breeding Objective which is to "identify animals whose progeny will be the most efficient converters of feed into farm profit" is expressed as Breeding Worth (BW) and is managed by New Zealand Animal Evaluation Limited (NZAEL) a subsidiary of DairyNZ. This paper describes the outcomes of a major review and development of the economic models to calculate economic values, as well as the practical implications for sire rankings and farmer selection decisions.

METHODS

A completely new set of models was developed for computing economic values used in the BW formulation. For milk yield traits, largely similar sets of assumptions and equations were used as in the previous model (Harris 1998). The cost of feed was modelled quite differently in the new approach (i.e. based on the models described by Chapman *et al.* (2012) for ranking forages), but despite this, new economic values for milk component yield and milk volume traits were very similar to those from the historic model. For as many inputs as practical, five year rolling averages of historic values were used as the model inputs. This was particularly relevant for the assumption about the relative value of protein versus fat, where this ratio has been highly variable over time, and the five year average avoided too much impact of very old and less relevant payments, while still evening out some short term fluctuations. In order to keep the economic values in strict concordance with the philosophy of improving the efficient conversion of feed into profit, a

rescaling methodology was applied to all of the economic values affecting per animal feed requirements. Any trait that increased the amount of feed required per cow on the home farm (commonly referred to as the milking platform) received a penalty under the assumption of a proportional reduction in stocking rate and therefore a loss of per cow profitability. There were three components to the economic value of milk volume which accounted for the volume charges associated with milk collection under the Fonterra payment system including a net effect of a peak season supply premium, the feed costs associated with milk lactose content which is very closely linked to milk volume, and an adjustment for the fact that high milk yielding cows have higher mastitis.

The economic value of live weight was computed using methodology that we had previously applied in breeding objective developments for many other livestock farming systems (e.g. Amer *et al.* 2001). Four independent components of the live weight economic value were calculated, each with separate discounted genetics expressions coefficients (e.g. Berry *et al.* 2006) to account for different timing and frequency of trait expression. The four live weight economic value components were bobby calf revenue, heifer rearing costs, annual cow maintenance feed requirements, and cull cow carcase value. Pricing schemes for bobby calf and cull cow values took account of both average per kg payment values, but also price premiums for heavier weight bands. Feed costs for heifer replacements used opportunity costs of feed on sheep and beef farms assuming contract heifer rearing costs off the home farm that would be directly proportional to the feed requirements of the animals. Similarly, dry cow feed costs assumed that all dry cows in the South Island would be fed on support blocks with lower opportunity costs of feed than occur on the milking platform.

The rationale for the economic value of Cow Survival was changed with a new assumption that lower survival would result in higher herd replacement heifer costs, and a higher proportion of younger cows in the herd which tend to be less profitable than mature cows due to lower milk yields. The previous rationale was that lower survival would result in less voluntary culling.

The somatic cell score economic value had three components, namely, a bulk tank penalty associated with milk processor charges when bulk tank average cell counts exceed thresholds which invoke price penalties, an account for the relationship between somatic cell score and cow survival, and a further relationship in the link between higher somatic cell counts in individual cows and their incidence of mastitis. Bulk tank penalties were modelled using aggregations of whole herd test results for somatic cell count to generate a distribution of bulk tank readings by region and farm. A certain proportion of farms capture price penalties which can be then translated into an average price penalty per litre of milk under this base level of somatic cell count. The same calculation was then undertaken to work out what the average price penalty per litre would be if all cows increased their somatic cell score by a single unit. The impact of somatic cell score on cow survival was quantified using a genetic regression coefficient derived from variance component estimates available from the New Zealand national genetic evaluation system. A corresponding genetic regression coefficient for clinical mastitis on somatic cell score was derived using a combination of values available in New Zealand, and values from the international literature, as no genetic evaluation currently exists for mastitis in the New Zealand system.

The three components which made up the final economic weight for fertility accounted for lost milk due to late calving, reduced survival due to culling on poor fertility, and lost premium value on heifer calves bred by AI. While late calving cows tend to have a truncated lactation curve when it is assumed that the whole herd is dried off on a constant date, the shape of the lactation curve is also influenced by calving date, with earlier calving cows tending to have a lower peak than late calving cows. The effect of poor fertility on reduced survival was quantified using a genetic regression coefficient derived from variance component estimates available from the New Zealand national genetic evaluation system. Cows that calve earlier also have a greater chance than their

later calving counterparts of producing high value replacements or high value beef calves that can then be sold. There is currently a market price differential between recorded, artificially bred (AB) heifer calves, and those that are unrecorded reflecting the superior genetic merit and scarcity value of recorded AB calves. This differential equates to a range between $350 to $400 per calf. Discounted genetic expressions coefficients were used to combine component economic values expressed in different animal classes (replacements, lactating cows vs cull cows) into a per lactating cow basis.

RESULTS

The results from the new model and corresponding assumptions led to a set of economic values for the Breeding Worth calculation as shown in Table 1. Milk protein has a much higher relative economic value than fat, reflecting both the relative price of fat in the market, and also the higher feed costs associated with the relatively energy dense milk fat component. While in absolute terms, the penalty for milk volume appears modest, this trait has a very high genetic standard deviation because of its units (litres), and in practice has a significant influence on bull and breed rankings. The feed cost associated with the lactose component of milk volume is a major contributor to the penalty, although volume charges applied by milk processors to cover trucking and processing costs of raw milk were also significant.

Both higher maintenance costs for the herd, and higher heifer rearing costs associated with increased live weight breeding value contributed in roughly equal proportions to the live weight breeding value. The higher revenues from bobby calf and cull cow sales only offset approximately 25% of the feed costs for larger cows and heifers.

The economic value of cow survival ($0.148 per day of average herd life in Table 1) represents a substantial increase on the economic value used previously ($0.048 per day). This further resulted in a modest increase in the economic value of somatic cell score, and a substantial increase in the economic value of fertility, as the impact of fertility on cow survival is a significant component of the overall economic value of fertility, whereas bulk tank penalties are a significant component of the economic value of somatic cell score.

In summary, the economic values for milk protein, milk volume and live weight have all changed by less than 3%. The economic value of milk fat has dropped by 7% while the economic values of survival, somatic cell score, and fertility have increased in magnitude by 200%, 20% and 135% respectively.

IMPLICATIONS

The correlation across all bulls with a minimum reliability for the index of 75% was 0.974 between the new index and the index used in 2012. The new index had moderate positive effects on breed averages (BW Reliability > 75%) for Jersey (+$9.20) and Kiwi Cross (+$5.00), but resulted in a lower average BW for Friesian (-$12.40), Aryshire (-$10.06) and Other (-$19.20) largely due to the increased emphasis on fertility. The new index also changed the breed representation in the top 100 bulls, 41 were Friesian (previously 45), 28 were Jersey (previously 23) and 31 were Kiwi Cross (previously 32). While the correlations between the new index and the old index appear high, significant shifts in rankings among the top AI bulls have been observed in practice. In particular, some bulls which are favourable for high production and live weight but weaker for fertility and survival have dropped substantially in their ranking. In general, there has been a high level of industry acceptance of the new index.

Table 1. Summary table detailing the calculation of economic weights for the new national breeding objective for the New Zealand dairy industry[1]

Trait (units)	Economic value ($/unit change)	Genetic regression	Discounted genetic expressions	Component economic weight ($/unit change)	Aggregated economic weight ($/unit change)	2012 BW
Milk Fat (kg)	**1.79**	**1.00**	**1.00**	**1.79**	**1.79**	**1.92**
Milk Protein (kg)	**8.63**	**1.00**	**1.00**	**8.63**	**8.63**	**8.69**
Milk Volume (litres)					**-0.091**	**-0.094**
Volume charge component	-0.038	1.00	1.00	-0.038		
Lactose feed cost component	-1.032	0.049	1.00	-0.051		
Mastitis component	-86.32	0.00002618	1.00	-0.002		
Live weight (kg)					**-1.52**	**-1.48**
Cow maintenance component	-1.16	1.00	1.00	-1.16		
Bobby calf value component	0.34	1.00	0.67	0.23		
Heifer replacement feed costs	-3.17	1.00	0.27	-0.86		
Cull cow carcase value component	1.51	1.00	0.18	0.27		
Cow Survival (days of average herd age)	**0.82**	**1.00**	**0.18**	**0.148**	**0.148**	**0.048**
Somatic cell score (log cells/ml)					**-38.57**	**-31.46**
Bulk tank penalty	-24.03	1.00	1.00	-24.03		
Survival component	0.82	-65.129	0.18	-9.62		
Mastitis component	-86.32	0.057	1.00	-4.92		
Fertility (% calving in first 42 days)					**7.35**	**3.12**
Lost milk component	1.84	1.00	1.00	1.84		
Survival component	0.82	27.847	0.18	4.11		
AB heifer calf premium component	1.41	1.00	1.00	1.41		

[1]Economic values give the change in dairy farm gross margins per industry average animal that expresses the trait. Bold font traits reflect traits for which economic weighting get applied in the index. Normal font traits are component traits for which economic values have been calculated for convenience, but their impact on the NBO comes either through aggregation of components or through their genetic relationship with other traits. Genetic regressions are of component traits on profit traits and explicitly account for the genetic relationships between the traits that capture the final weighting in the index and component traits. Discounted genetic expressions coefficients account for the fact that the expressions of some traits and their components comes with different timing and frequency of expression and this needs to be accounted for in the index formulation.

REFERENCES

Amer P. R., Simm G., Keane M. G., Diskin M. G., and Wickham B. W. (2001) *Livestock Production Science* **67**: 223.

Berry D. P., Madalena F. E. and Amer P.R. (2006) *Livestock Science* **99**: 159.

Chapman D.F., Bryant J.R., McMillan W.H., and Khaembah E.N. (2012) *New Zealand Grasslands Association*. **74**: 209.

Harris B. L. (1998) *Proceedings of the 6th World Congress on Genetics Applied to Animal Production* **25**: 383.

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:475-478

DERIVING ECONOMIC VALUES FOR REACTION NORMS OF GROWTH IN PIGS

S. Hermesch[1] and P.R. Amer[2]

Animal Genetics and Breeding Unit, a joint venture of NSW Department of Primary Industries and University of New England, University of New England, Armidale, NSW 2351;
AbacusBio Ltd, PO Box 5585, Dunedin, 9058, New Zealand

SUMMARY

Slopes of reaction norm models, also called reaction norms (RN), are alternative traits used in animal breeding for selection of genotypes that perform more consistently across a range of environments. Environmental sensitivity is of economic importance when the environment where selection takes place differs considerably from the commercial environment of slaughter pigs. The position on the environmental trajectory where intercept of reaction norm models is defined influences the economic values (EV) for slope and intercept. This position has to correspond to the trait definitions of intercept and slope of reaction norm models used to estimate variance components. The magnitude of EV for RN depends on the difference between the selection and production environments and the EV for the trait of interest. Economic values for RN may be negative or positive depending on whether the production environment is below or above the selection environment. Non-linear EV for growth across the environmental trajectory had minimal impact on the EV for RN of growth. Further genetic and economic analyses of extensive industry data are required to better quantify the economic importance of RN in pig breeding.

INTRODUCTION

Reaction norms quantify genotype by environment interactions by describing the response of genotypes to varying environmental conditions (Falconer and Mackay, 1996). As such, RN are alternative traits used in animal breeding for selection of genotypes that perform more consistently across a range of environments. For pig breeding, Knap (2005) derived the EV for the RN of days to reach market. It was assumed that pigs were selected in a superior environment typical for nucleus herds, while production was at an inferior environment representing the average customer farm with lower performance.

It was the aim of this paper to discuss economic implications of genetic differences in environmental sensitivity and to define EV for RN when selection is in the average environment using growth rate of pigs as an example trait.

METHODS

Selection in superior or inferior environments. Pigs are often selected in a superior nucleus environment and progeny of sires may have to perform in inferior environments prevalent on customer farms. International breeding companies, however, have nucleus herds in multiple countries with varying climatic and husbandry conditions. It is therefore feasible that sires may also be selected in an inferior environment with their commercial progeny raised in superior environments. Economic benefits of reduced environmental sensitivity of genotypes differ between these two scenarios. Low environmental sensitivity of genotypes is desirable when selection of sires occurs in the superior environment as it leads to superior performance of progeny in the inferior environments. In contrast, high environmental sensitivity of genotypes is economically beneficial when sires are selected in an inferior environment because progeny will be able to exhibit superior performance in better environments.

However, applying appropriate EV for RN when selection is in superior or inferior environments may not be the best approach because the intercept, which represents the traditional

trait, is defined for the selection environment and not the average environment of the environmental trajectory. Van Tienderen and Koelewijn (1994) outlined the dependency of (co)variances of intercepts and slopes on the position of the intercept on the environmental scale and suggested to define the intercept for the average environment of the environmental trajectory. This recommendation has generally been adopted in animal breeding applications (e.g. Kolmodin and Bijma, 2004; Su *et al*. 2006). In principle it is still possible to use reaction norm models for the situations outlined above. Intercept is then defined for the selection environment, which is situated above or below the average of the environmental trajectory, to ensure that trait definitions of intercept and slope of reaction norm models correspond to the EV for RN.

Selection in average environment of trajectory. Centering environments on the average environment in genetic analyses based on RN models (van Tienderen and Koelewijn, 1994) implies that the intercept corresponds to the estimated breeding value of the trait in the average (zero) environment. Genetic merit of genotypes across the environmental trajectory is defined as:

$$G_{gi}(E_{jk}) = G_{gi}(E_{j0}) + b_{(G_{gi}\cdot E_j)} * (E_{jk} - E_{j0})$$

where $G_{gi}(E_{jk})$ is genetic merit of genotype g in trait *i* for the k^{th} value of environmental variable *j*; $G_{gi}(E_{j0})$ is genetic merit of genotype *g* in trait *i* for the average value (*0*) of environmental variable *j*; $b_{(G_{gi}\cdot E_j)}$ is the RN quantifying the response *G* of genotype *g* in trait *i* per unit change in environmental variable *j* and $(E_{jk} - E_{j0})$ is the difference between the average (*0*) and k^{th} value of environmental variable *j*.

Knap (2005) defined the EV of RN for days to reach market weight as the EV for days to reach market weight times the difference in the environmental variable between the selection and production environments. This specific example can be extended to the generic case and EV for RN ($EV\_b_{(i\cdot E_{jk})}$) are then:

$$EV\_b_{(i\cdot E_{jk})} = (E_{jk} - E_{j0}) * EV\_i(E_{jk})$$

where $(E_{jk} - E_{j0})$ has been explained above and $EV\_i(E_{jk})$ is the EV of trait *i* for the k^{th} value of environmental variable *j*. In this way, there is only an EV associated with a RN when the average production environment of progeny (E_{jk}) differs from the average selection environment (E_{j0}). The magnitude of the EV for RN depends on the difference between the average environment for which the intercept is defined and the production environment of progeny of sires below or above the average environment.

Economic value of a trait varies across the environmental trajectory. For lifetime average daily gain (*ADG*) together with feed conversion ratio in the breeding objective, the EV is:

$$EV\_ADG = \left(\frac{Age_P}{Gr_P}\right) \times C_{NF}$$

where Age_P is the age of a finished pig at 90 kg live weight (130 days); Gr_P is the growth rate of a finished pig just prior to slaughter (900 g·day^{-1}) and C_{NF} is the daily non-feed costs per pig from weaning to slaughter ($AU 0.8 per day). The EV for ADG is $AU 0.116 per g·day^{-1} for an ADG of 692 g·day^{-1}, which was also used to derive the EV for RN of growth.

The EV for growth is affected by the level of performance in growth. It varies from $AU 0.139 to $AU 0.098 per g·day^{-1} for environments with a group average of ADG of 60 g·day^{-1} below or above a group average of ADG of 692 g·day^{-1}. This variation in the EV for growth across the environmental trajectory contributes to economic benefits of lower environmental sensitivity. A less environmentally sensitive genotype is economically advantageous as the economic losses of a reduced growth in high environments are lower than the economic benefits resulting from a higher growth in the low environments due to the non-linear relationship between growth and farm profit. This economic advantage is quantified by the proportion of pigs at each environmental level times

the relevant EV for growth at each environmental level and summed over all environmental levels. The economic advantage is larger for wider spread of progeny across the environmental trajectory.

RESULTS AND DISCUSSIONS

Economic values for RN are zero when the production environment of progeny equals the average environment. If progenies of a sire are raised in inferior or superior environments relative to the average environment, EV for RN of growth were $AU ±3.71 and $AU ±0.104 per (g·day^{-1}) per standard deviation of each environmental variable (Table 1). Please note, EV for RN are negative or positive (symbolized as +/-) depending on whether the production environment of progeny is below or above the average selection environment. Four distinct health environments were used by Schinckel *et al*. (1999) to evaluate line by environment interactions. Environments differed by about 80 g·day^{-1}, which corresponds to an EV for RN of growth of $AU +/- 9.28 per pig. Li and Hermesch (2012) found significant RN for growth for two environmental variables which were based on least squares means (LSM) for ADG and backfat (BF) of contemporary groups. The standard deviations of these two environmental variables were 32 g·day^{-1} and 0.9 mm. The range of RN estimates for growth is also shown for both environmental variables to illustrate genetic differences between sires. The standard deviations of sire solutions were 12.7 for the intercept and 0.025 and 1.079 for RN based on environmental variables of LSM for ADG and BF. Economic values per standard deviation of sire solutions are then $AU 1.47 for the intercept, $AU 0.093 for RN based on LSM for ADG and $AU 0.112 for RN based on LSM for BF.

Table 1 Standard deviations in environmental variables (SD EnVar, g·day^{-1} or mm), magnitude of economic value for reaction norm (RN) of average daily gain (ADG; $AU/pig per g·day^{-1} times SD EnVar) and range of RN for ADG (g·day^{-1} per EnVar))

EnVar[1]	SD EnVar	EV for RN of ADG[2]	Range of RN of ADG
LSM for ADG of CG g·day^{-1}	32	+/-3.71	-0.102 to 0.103
LSM for BF of CG mm	0.9	+/-0.104	-5.04 to 6.78

[1] LSM: least squares means, CG: contemporary group; [2] +/-: EV for RN may be positive or negative

Additional economic benefits resulting from lower environmental sensitivity depend on non-linearity of EV for growth across the environmental trajectory and the spread of progeny of sires across the environmental range below or above the average environment (Table 2). Economic values for growth are more variable across a lower environmental range, which lead to higher EV for RN for lower performance levels. Overall, the economic advantage of less environmentally sensitive genotypes is small because EV for growth is not sufficiently non-linear across a realistic environmental trajectory. However, this EV of RN ignores the benefits of more consistent performance across environments. For example, differences in environmental sensitivity of sires contribute to variability in performance of pigs within a batch. This variability within a batch may lead to non-linearity in profitability, resulting from lost revenues of light-weight pigs that do not reach target market weight. These under-weight pigs are sent to market in order to vacate housing facilities for the next batch. The EV for growth does not capture this loss in revenue as it assumes that all pigs reach target weight. Batch variability can be even more costly in production systems attempting to achieve a consistent market supply. This is because dips in growth create under-supplies of finished stock at certain times, and then over-supply subsequently.

The environmental variable is expected to be normally distributed for most situations as was found by Li and Hermesch (2012). Variation among contemporary groups may lead to skewness in the environmental variable as some values of the environmental variable may be more represented than others by individual contemporary groups. This may also lead to a skewed representation of

sires across the environmental trajectory. However, provided sires are expected to make equal contributions across individual contemporary groups, there will be no economic advantage for less environmentally sensitive genotypes. This is because the cumulative superior (inferior) genetic merit of a sire for environments below the average environment is matched by the cumulative inferior (superior) genetic merit of a sire for environments above the average environment.

Non-linear RN are likely to lead to non-linear profitability across the environmental trajectory, which contributes to the EV for environmental sensitivity. Deriving EV for multiple, higher-order RN parameters would be challenging with a higher-order polynomial parameterisation because of multi co-linearity with the other RN traits in the breeding objective. However, an economic rationale could be established to penalise genotypes that were predicted as being likely to deteriorate rapidly at an extreme end of the environmental continuum. An empirical approach would be required to integrate the economic rationale with the polynomial coefficients.

Table 2 Economic values (EV) for reaction norms of growth (ADG) due to changes in EV for ADG across environmental trajectories with mean performances of 500 to 800 g·day^{-1} and varying spread of progenies of sires across environmental trajectory (EnVar)

Spread of progeny of sires in standard deviation of EnVar	Mean growth performance (g·day^{-1})			
	500	**600**	**700**	**800**
40	1.361	0.78	0.488	0.326
20	0.365	0.211	0.133	0.089

CONCLUSIONS

Economic values of RN exist if the production environment of progeny differs from the average selection environment and when the EV of a trait varies considerably across the environmental trajectory. The magnitude of EV for RN depends predominantly on the difference between the selection environment and production environment of progeny as well as the EV for the trait of interest. Further genetic and economic analyses of extensive industry data are required to better quantify the economic importance of RN in pig breeding.

ACKNOWLEDGMENTS

This project was funded by the Pork CRC under Project 2B-102.

REFERENCES

Falconer D.S. and Mackay T.F.C (1996) 'Introduction to quantitative genetics' 4th ed. Longman Group, Essex.

Knap P.W. (2005) *Austr. J. Exp. Agric*. **7-8**: 763.

Kolmodin R. and Bijma P. (2004) *Genet. Sel. Evol.* **36**: 435.

Li L. and Hermesch S. (2012) AGBU Pig Genetics Workshop Notes. Available online at: http://agbu.une.edu.au/pig_genetics/pdf/2012/P9-Li-Hermesch-GxE.pdf.

Schinckel A.P., Richert B.T., Frank J.W. and Kendall D.C. (1999) Purdue University 1999 Swine Day Report. Available online at: http://www.ansc.purdue.edu/swine/swineday/sday99/13.pdf.

Su G., Madsen P., Lund M.S., Sorensen D., Korsgaard I.R. and Jensen J. (2006) *J. Anim. Sci*. **84**: 1651.

Van Tienderen P.H. and Koelewijn H.P. (1994) *Genet. Res.* **64**: 115.

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:479-482

THE ECONOMIC VALUE OF BODY CONDITION SCORE IN NEW ZEALAND SEASONAL DAIRYING SYSTEMS

T.J. Byrne[1], B. Santos[1], P.R. Amer[1], and J.R. Bryant[2]

[1]AbacusBio Limited, P O Box 5585, Dunedin 9058, New Zealand
[2]DairyNZ, Private Bag 3221, Hamilton 3240, New Zealand

SUMMARY

The body condition score (BCS) of a dairy cow will fluctuate according to her physiological state. Where pasture availability is variable through the seasons, these fluctuations will have an economic cost if body condition is lost when feed costs are low and replaced when feed costs are high. Differences in efficiency of BCS mobilisation and replenishment, seasonal differences in cost of feed, the value of additional milk solids production at the end of lactation (by milking longer to utilise body condition) or drying off with conservation of body condition were the basis of the calculations presented in the economic value (EV) of one unit of BCS in dairy cows.

INTRODUCTION

The body condition score (BCS) of an adult dairy cow will fluctuate according to her physiological state (Nicol and Brookes 2007). In farming systems where pasture availability varies throughout the year, these fluctuations have implications for the cost of feed and farm management. Body condition score at the end of spring (120 days from planned start of calving) is very important since it defines the feeding management required to return the cow to adequate pre-calving BCS targets which will support good production and fertility in the following lactation.

Currently, BCS is used as a correlated predictor in the genetic evaluation of dairy cow fertility in the New Zealand dairy industry. This paper describes the calculation of an EV for the trait 'autumn body condition score' in seasonal dairy production systems. The New Zealand dairy industry is currently considering a proposal to include 'autumn body condition score' as a trait with a direct economic weight in its national breeding objective.

RATIONALE

The economic impact of genetic differences in ability to maintain body condition can be assessed by balancing the cost of extra feeding to maintain milk production in cows that retain body condition in spring rather than converting it to milk against the costs of three alternative management strategies as follows:

1) Excess body condition present in late lactation can be converted directly into higher milk solids revenue by milking slightly longer without providing additional feed;

2) More body condition at dry off can lead to savings in autumn/winter feed costs as less BCS gain is needed to meet pre-calving BCS targets for the following season;

3) Cows with high BCS could have an extended lactation with additional feeding to support milk production because low BCS cows need to be dried off early.

The approach used followed the standard practice of treating one unit of BCS change as equivalent to 6.58% of live weight when considering energy requirements (Anonymous, 2010) since this allowed the associated energy requirements to be scaled for breed differences in mature live weight. Calculations were also based on the assumption that cows would have recovered their pre-calving body condition scores by the end of the lactation i.e. that there would be no future cost associated with reduced production or effects on subsequent fertility. Further assumptions were that a gain of 1 kg of live weight in late lactation requires 50 MJME, a loss of 1 kg live weight saves 37 MJME for milk production in lactating cows, and 72 MJME is required per 1 kg of live

weight gain in dry cows (Anonymous 2010). These assumptions, along with seasonal differences in feed costs (Chapman *et al.* 2012), and the value of extra milk solids production late in the lactation season whether through fat mobilisation or later dry off, formed the basis of the calculation of the EV per unit of BCS in dairy cows.

METHODOLOGY

Quantifying the extra feed costs on the milking platform from retaining BCS in spring. It was assumed that a cow that is one BCS unit higher than her herd-mates at the end of the spring period has achieved this by mobilising less body condition during early lactation. On a herd basis, it was also assumed that 30% of this extra body condition came from condition retained in early spring, while 70% came from condition retained in late spring. Early spring is a period of relatively high feed costs where supplementation is often required, whereas late spring is a time of pasture surplus and relatively low feed costs. Because of the typical timing and spread of calving, the majority of overall herd body condition mobilisation is expected to occur in late spring.

Cows with genetic propensity to retain body condition mobilise less energy, and so need more feed (37 MJME) to produce an equivalent amount of milk than cows with a genetic propensity to lose body condition. The following equation predicts an aggregated feed cost over the lactation to supply energy for BCS retention during lactation (i.e. BCS retention in early and late spring associated with a one unit increase in BCS in lactating cows, FC_b^{ABCS}).

$$FC_b^{BCS} = \sum_{s=1}^{2}\left[LW_b \times 0.0658 \times EC \times P_s \times Pf_s\right], \qquad (1)$$

where for cow breed b, LW is the cow mature live weight, EC is the energy change associated with the BCS retention (37 MJME/BCS unit/kg live weight) and for season s (s = early spring or late spring), P is the proportion of the condition score gain captured through associated BCS retention (30% in early spring and 70% in late spring), and Pf is the price of feed per MJME (Chapman *et al.*, 2012).

Quantifying the gain in milk solids revenue resulting from higher BCS through lactation leading to later drying off without additional feed inputs. The benefits of higher BCS throughout lactation can be captured as the value of feed used directly for milk production instead of BCS gain. The energy value saved through not having to gain body condition score (50 MJME per kg of live weight gain in lactating cows) was assumed to be converted to extra milk production in later lactation. The following calculation predicts the milk solid equivalent of one unit of BCS captured through later dry off in lactating cows, without any extra feed being supplied to the lactating cow.

$$MSV_b^{BCS} = \left(\frac{LW_b \times 0.0658 \times ES}{MS_b}\right) \times Pm, \qquad (2)$$

where for cow breed b, LW is the cow mature live weight, ES is the energy saved per kg live weight from not having to gain BCS, which can be used for milk production (50 MJME), MS is the MJME required per kg MS production, and Pm is the milk solids price.

Quantifying the savings in autumn or winter (May) feed costs through less need for supplementary feeding to gain BCS in late lactation. A cow that is one BCS unit lower than her herd-mates at the end of spring was assumed to have to recover that condition by additional feeding in late lactation. It was also decided that this additional feeding would be undertaken in the month of May, irrespective of the region of New Zealand. This assumption is not completely correct since in the North Island, May is counted as a winter feed cost, whereas in the South Island it is counted as an autumn feed cost. The decision was necessary however to provide some

standardisation. The following calculation describes the savings in May feed costs associated with carrying forward an additional BCS unit.

$$WFC_b^{BCS} = LW_b \times 0.0658 \times EC \times Pf_{May}, \qquad (3)$$

where for cow breed b, LW is the cow mature live weight, and for season s (s = winter), EC is the energy saved from not having to achieve BCS gain (50 MJME per kg of live weight gain in late lactation cows) and Pf_{May} is the price of feed per MJME for the region in May.

Quantifying the profit resulting from prolonging lactation and providing additional feed to support milk production. Some farms capture the benefits of less mobilisation of condition during early lactation by milking cows longer while continuing to provide additional feed inputs. This alternative rationale reflects the opportunity to utilise retained BCS to ensure a longer lactation. This later drying off aspect of high BCS animals is not captured by breeding values for milk production traits.

The ability to milk cows with more condition for longer depends on the rate of decline in BCS that occurs in late lactation. The farm model currently used to derive the National Breeding Objective, assumes a lactation length of 270 days and a winter period of 61 days (Amer 2013). This leaves the balance of 34 days available. Under the assumption that cows have returned to pre-calving condition before the winter period (meaning they only need energy for maintenance and foetal growth during winter) a cow going into late lactation with one BCS unit more than her herd-mates could be milked for an additional 34 days provided additional feed was available on the milking platform to support her total feed requirements over and above the feed costs of a dried off cow that needs to recover body condition score prior to the beginning of winter. End of lactation daily milk production in kg MS, after accounting for the proportions of early-, mid-, and late-season calving cows, was incorporated for each breed. The following calculation describes the revenue component (MSP) associated with prolonged lactation, with additional feeding to prolong milk production.

$$MSP_b^{BCS} = 34 \times MP_b \times Pm, \quad (4)$$

where for cow breed b, over 34 days MP is the daily milk production in kg MS at the end of lactation and Pm is the milk solids price ($/ kg MS).

In order to incorporate the marginal cost of extended lactation, feed costs to support milking are calculated for a cow over and above those required for herd-mates which are one BCS unit lower, have been dried off, and are being fed to recover condition. The following calculation describes feed energy costs to support that milk production ($MSFC$) associated with prolonged lactation, with additional feeding to prolong milk production.

$$MSFC_b^{BCS} = 34 \times MP_b \times MS_b \times Pf_a - LW_b \times 0.0658 \times ER \times Pf_a \qquad (5)$$

where for cow breed b, over 34 days, MP is the daily milk production in kg of milk solids at the end of lactation, MS is the MJME required per kg MS production, Pf_a is the price of feed per MJME in autumn, LW is the cow mature live weight, and ER is the energy change per kg live weight from gaining BCS in dry cows (72 MJME). Feed costs for maintenance are the same for the high and low BCS cows and so have been omitted from equation (5).

The following calculation describes variable costs (electricity and labour) to support the additional milk production ($MSVC$) associated with prolonged lactation.

$$MSVC_r^{BCS} = \sum_{d=1}^{34} \left[\frac{(VCL_r + VCE_r)}{LL_r} \right] \qquad (6)$$

where for region r, VCL is variable cost associated with labour per lactation, VCE is variable cost associated with electricity per lactation, and LL is the lactation length.

Economic value calculation. The EV of BCS in dairy cows can be assessed by balancing the extra feed costs on the milking platform from retaining BCS in spring against the three alternative strategies for farmers to extract value from higher BCS going in to autumn as follows.

- Extra feed costs on the milking platform from retaining BCS in spring (equation 1) against extra milk production in late lactation (equation 2):
$$EV_b^{BCS\_MS} = MSV_b^{BCS} - FC_b^{BCS}$$
- Extra feed costs on the milking platform from retaining BCS in spring (equation 1) against feed cost savings in May (equation 3):
$$EV_b^{BCS\_WF} = WFC_b^{BCS} - FC_b^{BCS}$$
- Extra feed costs on the milking platform from retaining BCS in spring (equation 1) against extra milk production in prolonged lactation (equation 4) after accounting for the extra cost of feed energy to support milk production over and above the feed required to recover body condition in a dry low BCS cow (equation 5), and the additional variable costs (electricity and labour) of prolonged lactation milk production (equation 6):
$$EV_b^{BCS\_PL} = MSP_b^{BCS} - MSFC_b^{BCS} - MSVC_r^{BCS} - FC_b^{BCS}$$

Assuming a fixed supply of feed on the milking platform, changes in the amount of feed required per cow with a trait change resulted in corresponding changes in carrying capacity. This had consequences for farm profitability on a per cow basis after all costs, including feed, had been accounted for (Amer, 2013). After accounting for the proportion of cows by breed in each region of New Zealand, and constraining feed supply on the milking platform, the EVs, per body condition score unit, were \$83.23 (extra milk production in late lactation), \$23.16 (feed cost savings in May), and \$128.60 (extra milk production in prolonged lactation).

CONCLUSION

The incorporation of BCS into the genetic improvement programme for the New Zealand dairy industry represents an economic opportunity for New Zealand dairy farmers. Differences in efficiency of body condition mobilisation, seasonal differences in cost of feed, and the value of extra milk solids production (through fat mobilisation or later dry off) form the basis of calculations of the economic value of one unit of BCS in dairy cows. The final EV for BCS will need to also take into account the proportion of dairy farms that capture the value of increases in BCS by saving feed rather than additional milk production. The incorporation of these proportions and development of the final EV calculation is on-going.

REFERENCES

Amer P. R., Byrne T. J., Santos B. F. S., Bryant J., 2013 *Proc. Assoc. Advmt. Anim. Genet.* (this journal)

Anonymous, 2010. Facts and Figures for New Zealand Dairy Farmers. DairyNZ.

Chapman, D.F., Bryant, J.R., McMillan, W.H., Khaembah, E.N., 2012. New Zealand Grasslands Association.74 209.

Nicol, A.M., Brookes, I.M., 2007. New Zealand Society of Animal Production, pp. 151-172

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:483-486

ECONOMIC WEIGHTS FOR MATERNAL PIG TRAITS IN AUSTRALIA MOTIVATE GENETIC IMPROVEMENT FOR ROBUSTNESS

C.I. Ludemann[1], P.R. Amer[2] and S. Hermesch[3]

[1]AbacusBio Pty Ltd, 331 Penshurst Rd, Byaduk, VIC, 3301
[2]AbacusBio Ltd, 1st Floor Public Trust Building, 442 Moray Place, P O Box 5585, Dunedin 9058
[3]Animal Genetics and Breeding Unit, a joint venture of NSW Department of Primary Industries and University of New England, University of New England, Armidale, NSW 2351

SUMMARY

A new model was developed to provide Australian pig breeders with an ability to estimate economic values and economic weights of key traits in a way that was more flexible and relevant to pig producers. Economic weights were converted using a genetic standard deviation scalar so that the relative contributions of each trait to the overall maternal breeding objective could be made. Number of piglets born alive had the greatest contribution (30.9%) to the maternal index followed by daily gain (maternal) (20.5%) and sow mature weight (13.6%). Other traits considered in the maternal breeding objective were pre-weaning survival (13.2%), sow longevity (11.3%), gilt age at puberty (7.9%), and piglet survival at birth (2.6%). The emphasis on growth rates in pigs has led to heavier sow mature weights and associated economic and animal welfare costs. Inclusion of the mature weight trait into the maternal index will allow farmers to assess the trade-off between their desired rates of progress in pig growth traits and that of sow mature weight.

INTRODUCTION

The application of crossbreeding is fundamental to modern pig production systems and has important implications for breeding programs (Harris 1998). This includes the need for separate development of maternal and terminal breeding lines of pigs. For genetic improvement of maternal lines, both maternal traits such as litter size and sow longevity and also terminal traits such as growth rate, feed efficiency and carcass attributes are highly relevant. Terminal traits tend to have higher heritabilities, and are recorded earlier in the lives of selection candidates, making them easier to improve than maternal traits. Greater emphasis has been placed on terminal traits in pigs in Australia, resulting in the potential genetic progress in maternal traits being underutilised.

The number of traits in genetic evaluations of Australian pigs has increased over time. In regard to maternal traits, only litter size was considered in breeding objectives initially based on the model developed by Stewart *et al*. (1990). The bio-economic models developed by De Vries (1989) was used by Cameron and Crump (2001) to derive economic weights for litter size based on production and market parameters relevant for Australian conditions at the time. However, breeders require greater flexibility to set up company-specific breeding objectives for a wider range of traits. This paper provides a general overview of the Pig Economic Value (PigEV) model and lists economic weights for maternal traits based on Australian pig industry data. We also compare the relative contribution each trait's economic weight has to the maternal breeding objective.

MATERIALS AND METHODS

PigEV model. The PigEV model includes independent sub-models to derive economic values for maternal and terminal traits. However, only the maternal trait sub-model is described in this paper. Inputs are divided into those required to customise the breeding objective for a certain situation or operation and those which are not expected to change over time, or across farms. Input

parameters include production and price assumptions for growing pigs, replacements, and sows, as well as operational costs including the facility costs, depreciation and discount rates.

Partial economic values from the PigEV model are a quantification of the change in profit for a unit-change in a trait, expressed independently to each other. Maternal economic values included: the longevity of the sow; piglet survival at birth, piglet survival pre-weaning; number of piglets born alive; gilt age at puberty; mature weight of sow; and average daily gain (maternal) of piglets. Equations to estimate the economic value of each of these traits were described in more detail by Hermesch *et al*. (2012).

Trait units. The economic value of longevity (LONG) was calculated based on the marginal economic benefit of a sow achieving an extra parity over her lifetime. Survival at birth (SB) is a trait of the sow defined as the number of live born piglets divided by the total number of piglets born including still births. The economic value for survival at birth accounts for the gestation cost of the sow associated with the stillborn piglet and disposal costs. The pre-weaning survival (SW) trait was defined as the number of piglets surviving divided by the number of piglets born alive. Two alternative scenarios were considered to estimate the economic value of number of piglets born alive based on the pig operation being limited by a fixed number of piglets (NBAp) or a fixed number of sows (NBAs). Both of these traits cannot be applied in the same breeding objective at the same time and will depend on the pig operation. Larger pig operations may have less flexibility to sell more pigs into their existing supply contracts, without price reductions. Therefore, these producers may opt for the method based on fixed number of piglets. In contrast, smaller pig operators tend to sell opportunistically into larger markets. Hence, the approach based on a fixed number of sows is more appropriate for smaller producers. The gilt age at puberty (AP) trait was based on a one-day increase in the number of days a gilt required to achieve a weight suitable for mating. A one-gram-per-day increase in average daily gain of piglets as influenced by maternal genes (ADGm) was used for the average daily gain (maternal) trait. The mature weight (MW) trait (measured in kg of live weight at maturity) was a combination of more than one 'component' trait. Component traits represent different economic aspects of a change in a trait which contribute to its overall relative economic weighting. The economic value for mature weight for example accounted for the economic impact of a change in: energy requirements for replacement gilts to achieve final mature weight (MWg); sow maintenance energy requirements (MWm); sow capital costs (MWk); and sow cull value (MWc), for a one-kilogram increase in sow mature weight.

Discounted genetic expressions. Economic values do not take into account the timing and contribution a selection candidate's genes make to a trait over an extended period of time. Values were discounted back to the time of birth of a gilt replacement. The traits SB, SW, NBA, MWm, MWk and ADGm were traits expressed once per parity. LONG and MWc were expressed at the end of the sows life, while AP and MWg were assumed to be expressed at the time of first farrowing. Economic values were multiplied by the discounted genetic expression (DGE) coefficients, which account for the timing and frequency of expression of selection candidate's genes over an extended period of time, to produce an economic weight for each trait.

Trait genetic standard deviations. Absolute economic weights ($|EW|$) with different trait units can be multiplied by their genetic standard deviation (σ_G) to facilitate a comparison of the relative importance of traits to the breeding objective. The fixed-number-of-piglets variant of the number of piglets born alive economic weight was used for the comparison. Percent importance of traits to the breeding objective was calculated as the values of $|EW| \times \sigma_G$ divided by the sum of these values across all traits. The relative importance of each trait was computed either within maternal traits only (corresponding to a maternal sub index) or across a broader maternal role index (maternal line) that also included economic values for finishing pig (terminal) traits.

RESULTS AND DISCUSSION

The values of the three DGE coefficients required were 3.68 for traits expressed once per farrowing, 0.88 for traits expressed at the end of the sow's life and 0.96 for traits expressed at the time of first farrowing. Table 1 summarises the economic weights of the maternal traits and their relative contributions within the maternal sub index (M%) and also considering a more complete maternal line index which also includes terminal traits (I%).

The trait with the greatest overall contribution to the maternal pig breeding objective was number of piglets born alive. This was followed by average daily gain (maternal) and mature weight (overall). The economic value for number of piglets born alive for a fixed number of piglets (AU$31.4) estimated here was similar to that estimated by Cameron and Crump (2001) (AU$31.7). However, the economic value with a fixed number of sows was more than double that estimated for a fixed number of piglets. This demonstrates the importance of defining the specific limiting factor of each commercial production system. To our knowledge this effect of the production system on the economic value for litter size has not previously been considered.

Table 1. Genetic standard deviations (σ_G), economic values (EV, $AU), economic weights (EW=economic value × discounted genetic expression) and the relative contribution of maternal traits within the maternal sub index (M%) and the contribution of maternal and terminal traits to a more complete maternal line index (I%) in the Australian PigEV model

Trait	Units	σ_G	EV	EW	M%	I%
Longevity	Parities	0.4	99.0	86.9	11.3	6.1
Piglet survival at birth	Proportion	0.08	27.0	99.7	2.6	1.5
Pre-weaning survival	piglets·farrow^{-1}	0.03	404.4	1354	13.2	7.1
Number of piglets born alive	piglets·farrow^{-1}					
Fixed number of piglets		0.82	31.4	115.7	30.9	16.5
Fixed number of sows		0.82	68.6	252.89		
Gilt age at puberty	Days	10	-2.51	-2.4	7.9	4.2
Mature weight overall	kg live weight	10		-4.2	13.6	7.3
Gilt energy			-0.40	-0.39		
Sow maintenance			-0.37	-1.35		
Sow capital costs			-1.29	-4.79		
Sow cull value			2.66	2.33		
Average daily gain (maternal)	grams·day^{-1}	20	0.85	3.14	20.5	11.0
Other terminal traits						46.3

Knap (2005) defined robustness traits as pre-weaning survival, growing pig survival, and the number of litters a sow has over a lifetime. In that study, the robustness traits were shown to contribute significantly to overall pig production profitability (31%) in relation to conventional production traits such as carcass lean content (17%), days to slaughter (21%), average daily feed intake (19%), and litter size at farrowing (11%). It can be argued that the broader suite of maternal traits included in this study would further contribute to improvements in robustness. Selection pressure to slow increases in mature weight and maintain age at puberty will reduce the rate of genetic gain in growth rate. Maternal weaning weight also reflects the ability of the sow to support piglet production. Thus at 37.2%, the relative importance of the new maternal traits (excluding NBA) will have a significant impact on an overall maternal line index.

As the population average for traits change, so too can the optimal weighting for each trait. A Canadian study (Quinton *et al.* 2006) suggested additional emphasis needed to be placed on piglet

perinatal survival for Canadian herds with litter sizes over 10 piglets. In our model, as litter size increases, the number of discounted genetic expressions of piglet survival traits increases modestly. In addition, we would expect increasingly unfavourable genetic correlations for NBA with SB and SW.

Ongoing selection pressure on growth rates in Australian pigs is increasingly leading to concerns that mature weight may require additional emphasis in selection goals. Economic progress brought about by selection for growth traits will be tempered by the positive correlation between growth rates in pigs and mature weight of sows. Lewis and Bunter (2013) for example estimated a genetic correlation of 0.32 between weight of pigs at 20 weeks and weight of sows at mating for the fifth parity. Furthermore, Hermesch *et al.* (2010) suggests there is a 3kg increase in sow mature weight for every 10 gram per day increase in ADG. In terms of the overall effect of mature weight in a pig operation, our study shows that the small benefits from higher sow cull values will be outweighed by greater feed requirements for sow maintenance and replacement gilts, as well as higher capital costs for housing facilities.

CONCLUSIONS

The PigEV model provides greater flexibility for pig breeders to create breeding objectives for their own breeding programs. The relative contributions of these new traits to the breeding objective suggest that both recording effort within breeding programmes and infrastructure development to include genetic evaluation capability for these traits within PIGBLUP is warranted. In addition to improving the profitability of maternal line pigs selected on this expanded index we would expect the resulting sows to be more robust with more modest mature size and better survival.

ACKNOWLEDGEMENTS

This project was supported by the Australian Pork CRC under project 2B-102.

REFERENCES

Cameron N.D. and Crump R.E. (2001) *Manipulating Pig Production VIII Proc. Aust. Pig. Sci. Assoc.* **8**: 146

De Vries A.G. (1989) *Livestock Production Science* **21**:49.

Harris, D. L. (1998) *Journal of Animal Science* **76**: 2294.

Hermesch S., Jones R.M., Bunter K.L. and Gilbert H. (2010) *Proc. of the 9th World Congress on Genetics Applied to Livestock Production*

Hermesch S., Ludemann C.I. and Amer P.R. (2012) *AGBU Pig Genetics Workshop* October 2012. Armidale: University of New England available at: http://agbu.une.edu.au/pig_genetics/pdf/2012/P8-Hermesch-Ludemann-Amer-PigEV-model.pdf

Knap P.W. (2005) *Australian Journal of Experimental Agriculture* **45**:763.

Lewis C.R., and Bunter K.L. (2013) *Journal of Animal Science accepted.*

Quinton V.M., Wilton J.M., Robinson J.A. and Mathur P.K. (2006) *Livestock Science* **99**:69.

Stewart T.S., Bache D.H., Harris D.L., Einstein M.E., Lofgren D.L. and Schinckel A.P. (1990) *Animal Breeding and Genetics* **107**:340.

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:487-490

ECONOMIC IMPACT OF CHANGES TO THE BREEDING OBJECTIVES USED WITHIN THE NEW ZEALAND BEEF BREEDING INDUSTRY

J.A. Sise[1], T.J. Byrne[1], M.J Young[2] and P.R. Amer[1]

[1]AbacusBio Limited, PO Box 5585, Dunedin
[2]Beef + Lamb New Zealand Ltd., PO Box 39-085, Harewood, Christchurch, New Zealand

SUMMARY

The economic impact of potential changes to the breeding objectives and selection approaches used within the New Zealand (NZ) beef industry were quantified. A selection index model was used to derive economic weightings applied to commonly recorded traits, according to current genetic trends for NZ Angus and Hereford beef breeding herds. Breeders were assumed to record a range of liveweights on young bull selection candidates, as well as days to calving on their mothers. These traits were used to predict economic response to traits within the breeding objective (growth, days to calving and calving ease). Adding additional selection criteria, focussing on gestation length and body condition score, showed a relatively modest 3% increase in the annual response to selection within relatively intensive beef farming enterprises. Use of muscle and fat scanning and recording cow mature weights, which is already quite common in the industry, demonstrated a larger increase in the annual response to selection (approximately 10%). The impact of genotype-by-environment interaction was also examined, whereby the genetic parameters were modified to reflect situations where bulls bred on lowland country for a lowland breeding objective were used on hill country. Modifications to the breeding objectives and improved selection approaches to accommodate the specific needs of hill country farmers has the potential to provide substantial improvements in potential rates of genetic progress, which could lead to increased productivity within the commercial farming environment.

INTRODUCTION

Many bull breeders in New Zealand manage and select their cattle on easier farming country than the nutritionally challenging farming conditions that the majority of commercial beef cows endure. This regularly raises the question as to whether beef cattle breeding schemes could be modified to better suit the extensive hill country environment occupied by the majority of beef cows in New Zealand. The research required to characterise this problem is likely to be both time consuming and expensive, so there is a need to first quantify the magnitude of industry benefits that might be achieved through a better understanding of the extent and impact of a genotype by environment interaction, and the role of new selection criteria in improving rates of progress. This paper reports on the impact of the current indexes on genetic change, with economic weightings applied to commonly recorded traits estimated according to typical phenotypic recording practices and rates of genetic progress observed by the NZ Angus and Hereford populations. We then report on the impact of potential changes to the economic weightings, and discuss the application of the updated indexes in both lowland and extensive hill country herds.

SELECTION INDEX MODELLING

A selection index model was used to predict the response to selection in beef breeding herds, where herd bulls aged 2 at first use are selected from the top 30% of candidates, whilst artificial insemination (AI) bulls are selected from the top 10% at 8 years of age. Twenty percent of cows were assumed to be AI mated, and 80% mated to herd bulls, with an average cow age of 4.9 years (AbacusBio, commercial data), resulting in a generation interval of 4.1 years.

Objectives

Four different phenotypic recording strategies were evaluated. In the base model (strategy 1), recorded phenotypes for birth weight (BW) and 200 and 400 days weights (w200, w400) were assumed to be available on all young bull candidates, and 90 paternal siblings. Maternal 200 day weights (w200m) and days-to-calving (DC) records were assumed available on the candidates' mothers and 30 paternal half siblings. The AI bulls were assumed to have in excess of 200 progeny with BW, w200, w400 and 600 day weight (w600), as well as 30 daughters with w200m and DC records. Strategy 2 includes recording of mature cow weight (MCW) on the candidates' mother plus 30 paternal half siblings for herd bulls, and 30 daughters for the AI bulls. Strategy 3 builds in ultrasound records for rib fat depth (FDrib) and eye muscle area (EMA), whilst strategy 4, adds in direct and maternal gestation length (GL, GLmat) and cow body condition scores (BCS).

The genetic parameters applied within the selection index model were sourced from Archer *et al*. (2004), Bourden *et al.* (1982) and Meyer *et al*. (1995). Implied economic weightings were estimated as the values that would be required to achieve levels of genetic progress derived by visual interpolation of published genetic trends in NZ herds (Angus NZ 2010; Hereford NZ trends 2010). Table 1 shows the genetic trends (in trait units) reported, relative to expected rates of progress for the recording model used by strategy 1 breeders. Differences between Angus and Hereford data were small, so data were pooled to derive economic weightings.

Table 1: Genetic trends achieved over the last 10 years by NZ Angus and Hereford herds, compared to expected rates of progress for typical NZ breeders

Trait (unit)	Angus	Hereford	Expected progress	Estimated economic weight (cents)
BWT (kg)	0.05	0.03	0.11	-69
w200 (kg)	1	1	1.33	23
w200m (kg)	0.4	0.5	0.07	50
w400 (kg)	2	1.6	2.79	101
w600 (kg)	2.8	2.1	2.73	-
MCW (kg)	2	1.9	2.03	-
CE (%)	0.02	0.03	-0.01	399
CEm (%)	0.05	0.13	0.03	211
EMA (cm^2)	0.05	0.07	0.25	241
DC (days)	-0.2	-0.16	-0.28	-200

ECONOMIC WEIGHT REVISIONS

The recording models for strategies 2, 3 and 4 were used to include additional recorded trait information, and an alternate index was devised according to assumptions around the relative importance of key traits to the NZ beef industry. The alternate index included body condition score (which is highly correlated with both BW and other weights) and a negative weighing on MCW (to account for the economic impact of heavier cows). Gestation length was also included as a profit trait, reflecting breeder's desires to shorten GL to allow heifers and cow additional time to gain condition between calving and mating.

Table 2 shows the trait unit response to selection using the current and alternate index weightings. To enable a clear comparison between the existing and alternate indexes, total economic response has been estimated using the revised index weightings only, where the trait responses (generated using each set of index weightings) were multiplied and summed over all

traits within the objective. As more traits are recorded (within strategy models 2, 3 & 4) the overall economic response to selection increases. However, progress in some traits decreased, as the new information diverted selection pressure away from initially recorded traits onto those traits with new information. A good example of this can be observed in the results for strategy model 3, where, using the current index, ultrasound scan information for EMA results in an additional 28 cents of progress in EMA but only 14 cents in overall index.

Selection using the alternative index results in an increase of 50-60 cents (20%) per annum for all models. Inclusion of recorded trait information for BCS and GL had a minor impact on overall response, however the strong genetic correlation between BCS and MCW within strategy 4 resulted in larger animals, with unfavourable shifts in the genetic trends for MCW and CE.

Table 2: Trait response to selection using current and alternate index weightings, with the economic response to selection estimated using the alternate index weightings

	Current index Trait response to recording strategy					Alternate index Trait response to recording strategy				
Trait (unit)	Weight	1	2	3	4	Weight	1	2	3	4
BWT (kg)	-69	0.11	0.11	0.11	0.10	0	0.22	0.22	0.21	0.20
w200 (kg)	23	1.33	1.32	1.22	1.27	0	1.46	1.43	1.31	1.42
w200m (kg)	50	0.07	0.13	0.12	0.07	111	0.12	0.18	0.18	0.11
w400 (kg)	101	2.79	2.77	2.64	2.71	0	2.96	2.90	2.70	2.89
w600 (kg)	0	2.73	2.72	2.68	2.74	98	3.59	3.55	3.44	3.53
MCW(kg)	0	2.03	1.49	1.53	1.82	-36	3.40	2.68	2.66	2.99
CE (%)	399	-0.01	0.02	-0.01	-0.03	162	-0.15	-0.11	-0.13	-0.17
CEm (%)	211	0.03	0.01	0.03	0.04	72	0.02	0.01	0.03	0.03
FD P8 (mm)	0	0.01	0.01	0.02	0.02	0	0.01	0.01	0.04	0.03
FDRib (mm)	0	0.01	0.01	0.01	0.01	-116	0.01	0.01	0.04	0.03
IMF (cm^2)	312	-0.01	-0.01	-0.01	-0.01	0	-0.01	-0.01	-0.01	-0.01
EMA (cm^2)	241	0.25	0.25	0.48	0.48	122	0.28	0.29	0.45	0.46
DC (days)	-200	-0.28	-0.28	-0.26	-0.26	-96	-0.14	-0.14	-0.13	-0.12
BCS (units)	0	-0.01	-0.01	-0.01	0.00	3670	0.01	0.01	0.01	0.01
GL (days)	0	-0.07	-0.08	-0.07	-0.08	-274	-0.06	-0.06	-0.05	-0.07
GLm (days)	0	0.00	0.00	0.00	0.00	-128	0.00	0.00	0.00	0.00
Economic response (cents)		260	285	295	304		310	336	355	361

EXTENSIVE HILL COUNTRY GRAZING

To assess the impact of farm system type, where animals are grazed on extensive "hill" country versus easy "lowland" terrain, an alternative breeding objective was used. Two new traits were introduced to model hill country beef herds; i) Body condition score (BCS Hill); and ii) Days to calving: DC Hill). These traits were introduced to reflect a herd in which cows primarily perform a pasture clean up role on a mixed sheep and beef farm. The genetic correlation between BCS and BCS Hill was set to 0.5, the strong correlations between BCS with W600 and MCW were removed, and the correlation between DC and BCS Hill was set to -0.3. Similarly, the genetic

correlation between DC Hill and DC was set to 0.5, as was the genetic correlation between DC Hill and BCS Hill. However positive (unfavourable) correlations were introduced between DC Hill and growth traits, reflecting the unsuitability of fast growing genotypes for hill country herds.

Table 3 shows the economic responses to selection for recording strategy models 3 and 4, and the breeding program and recording objectives are shifted from the lowland to the hill. When BCS and DC are recorded on the hill, similar levels of economic progress can be achieved as would be observed using the alternate index on lowland. However, selection pressure is diverted away from growth, which results in smaller cows with improved maternal weaning weights and easier calving. If BCS and DC are recorded on lowland country, and the bulls are then used to generate replacements on the hill, growth rates increase, resulting in bigger cows, longer gestation times, and loss of calving ease. This results in a gross reduction in the benefits of genetic selection. Under recording strategy 3, BCS and GL are not recorded, therefore the genotype by environment interaction effect is less severe. With strategy model 4, selection emphasis applied to BCS on lowland, results in larger cows which are less suitable for hill country farming environments.

Table 3: Impact of land type on response to selection using recording strategies 3 and 4

Recording strategy model	3: GL & BCS not recorded			4: GL & BCS recorded		
Commercial Land type	Lowland	Hill	Hill	Lowland	Hill	Hill
Breeding program land type[1]	Lowland	Lowland	Hill	Lowland	Lowland	Hill
Annual response (cents)	355	343	373	361	222	391

[1] The focus of the breeding objective used for selection is assumed to match the breeding program land type, but annual response value is based on the economic values corresponding to the commercial land type.

POTENTIAL BENEFITS TO NEW ZEALAND

Using national herd size and discounted genetic expression methodologies, it is possible to estimate the value of genetic improvements over a10 year period to the commercial beef farming tier. Over 10 years of cumulating genetic progress, 1 unit of genetic trend cumulates to 1+2+3+....+10=55. If we count a further 5 years of progress (i.e. we consider the benefits of 10 years of progress over 15 years) then the benefits increase to 105. However, if we discount the benefits to account for the time lag using a discount rate of 7% (the farm mortgage rate minus the rate of inflation), this reduces to 59.4 units of progress in present value terms.

Assuming the current rate of progress (valued using the alternate index) is $2.60 per cow mated per year, the estimated value of 10 years of genetic progress (valued over a 15 year period) is $61.8 M or $154 per cow mated. Improvements in selection criteria and /or indexes could result in increases in the rate of genetic progress and economic returns to the beef industry. A 30% increase in genetic progress, combined with an industry penetration rate of 40% (i.e. 400,000 beef cows mated per year), would result in national benefits of up to $18.5M.

To achieve significant productivity increases within the beef farming sector, further input is required from both breeders and commercial farmers as to the relative importance of traits currently used within the index. Industry consultation is needed to determine farmer attitudes for change, and to establish whether opportunities exist to optimise both index weightings and phenotypic recording practices applied within New Zealand.

REFERENCES

Angus NZ (2010) http://angusnz.com/genetics_trends_2010.php
Archer J.A. *et al*. (2004) *Australian Journal of Experimental Agriculture* **44**: 393.
Bourdon R.M. and Brinks J.S., (1982) *Journal of Animal Science* **55**: 543.
Hereford NZ trends, (2010) http://www.herefords.co.nz/breedplan_sire_list.php
Meyer K., (1995) *Livestock Production Science* **44**: 125.

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:491-494

HERITABILITY OF TRACK CONDITION AFFINITY IN THE AUSTRALIAN THOROUGHBRED RACING POPULATION

B.D. Velie, N.A. Hamilton and C. M. Wade

University of Sydney, Faculty of Veterinary Science, Camperdown, Sydney, NSW, Australia 2006

SUMMARY

Performance records were obtained for 31,441 Thoroughbreds racing in Australia. Three traits, best race track condition, best average finish position track condition, and best earnings per start track condition, were calculated for each horse. Heritability of each trait was estimated using an animal model incorporating sex as a fixed effect and trainer as a random effect. Heritabilities of best race track condition, best average finish position track condition, and best earnings per start track condition were 0.03 ± 0.007, 0.04 ± 0.008, and 0.03 ± 0.007, respectively.

INTRODUCTION

The ability of a race horse to consistently perform under varied environmental conditions undoubtedly plays a significant role in the success it will have on the racecourse. This is arguably most on display when the skies open and it starts to rain. The ability of certain horses to perform well in wet conditions has given rise to labels such as “mudder” and “swimmer” with some horses branded as “mudders” and “swimmers” based solely on their breeding. Although multiple studies have demonstrated significant relationships between racing success and track surface (Silveira & Ferreira 2008; Cheetham *et al.* 2010), few studies have investigated the idea that the ability of a horse to perform well under wet conditions is heritable. In this study we explore this notion in the Australian Thoroughbred racing population and provide estimates of heritability for 3 traits associated with a horse’s performance under specific track conditions.

MATERIALS AND METHODS

Population. Performance records were made available by Racing Information Services Australia (RISA) for all Thoroughbreds entered in a race or official barrier trial in Australia from 1 August 2000 until 22 February 2011. The data were filtered to include only horses that had raced on the turf and that were under the supervision of a single trainer or training partnership during the entire time frame of the study. The filtered sample included a total of 31,441 horses representing offspring from 2,269 sires and 22,716 dams. The sample consisted of 1,743 (5.5%) intact males, 14,244 (45.3%) females, 15,444 (49.1%) geldings, 8 (<0.1%) cryptorchids and 2 (<0.1%) horses where the sex was not listed in the raw data provided by RISA.

Best Race Track Condition. Horse races in Australia are classed according to the previous performance of the horse and consist of ‘restricted’ and ‘open’ classes. Restricted races place conditions on horses eligible to race and can be restricted by age, sex, and/or number of previous wins (maidens, class 1 to 6 in increasing order of performance). Open class races have fewer restrictions based on previous performance (although they may be limited by sex and age) and are thus of a higher class than restricted races. The highest class of horse races are black type races, consisting of Listed, Group 3, Group 2 and Group 1 races (in increasing order of difficulty and prestige). Turf tracks in Australia are rated based on 5 categories (fast, good, dead, slow, heavy) and on a scale of 1 to 10 (Table 1).

Table 1: Australian turf track ratings

Category	Scale	Description
Fast	1	A dry hard track
Good	2	A firm track
Good	3	Ideal track with some give
Dead	4	Track with give, better side of dead
Dead	5	Significant amount of give, worse side of dead
Slow	6	A mildly rain effected track, better side of slow
Slow	7	Rain affected, worse side of slow
Heavy	8	Soft track, just into heavy range
Heavy	9	Very soft, genuine heavy
Heavy	10	Very soft, wet and muddy, heaviest category

Best race track condition (BRTC) was recorded as the category rating of the track in which a horse won its highest class of race in Australia. For horses with multiple wins at the same class of racing over a variety of track ratings, the category rating of the track for the race with the largest amount of prize money was chosen.

Best Average Finish Position Track Condition. The average position in which a horse finished under each category track rating was calculated for each horse. Best average finish position track condition (BAFPTC) was recorded as the category rating of the track in which a horse had its best average finish position.

Best Earnings Per Start Track Condition. Earnings per start under each category track rating were calculated for each horse. Best earnings per start track condition (BEPSTC) was recorded as the category rating of the track in which a horse had the highest earnings per start.

Heritability. Analyses were carried out using a single trait animal model in ASReml-R (R Development Core Team 2011). Sex and colour were included as fixed effects while trainer and horse were included as random effects. Cryptorchids and horses with no documented sex were excluded from the analysis (n = 10). Only fixed effects and covariates with a Wald-test $P<0.05$ were retained in the final model.

RESULTS

Descriptive statistics, stratified by sex, of BRTC, BAFPTC, and BEPSTC are shown in Tables 2, 3, and 4. Analysis of BRTC, BAFPTC, and BEPSTC yielded heritability estimates of 0.03 ± 0.007, 0.04 ± 0.008, and 0.03 ± 0.007, respectively. Sex and trainer were significant ($P < 0.01$) for all traits and were retained in the final models. Colour was not significant ($P = 0.06 - 0.40$) for any trait and was therefore dropped from the final models.

Table 2: Descriptive statistics, stratified by sex, of best race track condition for a sample of Thoroughbreds racing in Australia between 2000 and 2010

Best race track condition	Sex				
	Intact males	Females	Geldings	Cryptorchids	Not listed
Fast	67	565	663	0	0
Good	1016	8147	8912	5	1
Dead	370	2839	2869	1	0
Slow	183	1557	1682	2	0
Heavy	107	1136	1318	0	1
Total	1743	14244	15444	8	2

Table 3: Descriptive statistics, stratified by sex, of best average finish position track condition for a sample of Thoroughbreds racing in Australia between 2000 and 2010

Best average finish position track condition	Sex				
	Intact males	Females	Geldings	Cryptorchids	Not listed
Fast	102	819	1118	0	0
Good	555	4313	4496	2	1
Dead	443	3507	3441	3	0
Slow	394	3145	3514	2	0
Heavy	249	2460	2875	1	1
Total	1743	14244	15444	8	2

Table 4: Descriptive statistics, stratified by sex, of best earnings per start track condition for a sample of Thoroughbreds racing in Australia between 2000 and 2010

Best earnings per start track condition	Sex				
	Intact males	Females	Geldings	Cryptorchids	Not listed
Fast	95	685	923	0	0
Good	764	6047	6356	4	1
Dead	458	3453	3607	1	0
Slow	254	2361	2499	3	0
Heavy	172	1698	2059	0	1
Total	1743	14244	15444	8	2

DISCUSSION

On a wet track, mud and dirt are often kicked back into the faces of the horses that sit back in the field, potentially resulting in a dry track preference for horses that are put off by flying debris. While this debris may unfavourably affect these horses, it is just as likely to play a favourable role for horses that are unfazed by the flying mud, and is just one example of why a horse may finish in a higher position on a wet track compared to a dry track. In the current study BRTC, BAFPTC, and BEPSTC were used to assess each horse for its "preferred" track condition and to estimate the heritability of this preference. Heritabilities for BRTC (0.03 ± 0.007), BAFPTC (0.04 ± 0.008) and BEPSTC (0.03 ± 0.007) were estimated to be very low; however, it is interesting to note that a similar trait used to evaluate a horse's ideal race distance, has been shown to be significantly heritable ($h^2 = 0.61$-0.98) (Williamson and Beilharz 1996; Velie *et al*. [Under Review]). Because BRTC assumes that winning a race of lower class is always better than placing in a race of higher class, it was thought that the true "preference" of a horse may not be accurately assessed using BRTC. With this in mind, BAFPTC and BEPSTC were also analysed as these traits were able to account for a superior finish position without a horse having won the race. Unfortunately, although arguably a better assessment of a horse's track "preference", both BAFPTC and BEPSTC yielded heritability estimates less than 0.05, providing evidence to refute the racing of progeny under similar track conditions to that of their parents based solely on the track condition "preference" of the parents.

Multiple theories have been put forward as to why certain horses are able to perform well on wet tracks and others show a distinct "dislike" for wet track conditions. There is no doubt that the genetic composition of a horse's dam and sire significantly contributes to how well it performs on the racetrack (Ekiz *et al*. 2005; Ekiz and Kocak 2007; Bakhtiari and Kashan 2009; Binns *et al*. 2010; Hill *et al*. 2010). Our results suggest that the sire and dam contributions reflect attributes of

the horse that are separate from its affinity for a specific track condition (Cust *et al.* 2012). Regardless, additional research exploring the genetic contribution to a horse's "preferred" track condition would undoubtedly provide more insight into the reasons behind common observations that horses express an affinity for particular track conditions.

ACKNOWLEDGEMENTS

We thank the Australian Stud Book for their support and for providing the pedigrees for our analyses. Raw copyright racing material data supplied for analysis by Racing Information Services Australia. © 2012 RISA (and other parties working with it).

REFERENCES

Bakhtiaria J. and Kashan N.E.J. (2009) *Livestock Sci.* **120**: 1512.

Binns M.M., Boehler D.A. and Lambert D.H. (2010) *Anim. Genet.* **41**: 154.

Cheetham J., Riordan A.S., Mohammed H.O., McIlwraith C.W., and Fortier L.A. (2010) *Equine vet J.* **42**: 346.

Cust A.R.E., Anderson G.A., Whitton R.C. and Davies H.M.S. (2012) *Aust vet J.* DOI: 10.1111/avj.12012

Ekiz B., Kocak O. and Yilmaz A. (2005) *Arch Tierz Dummerstorf* **48**: 121.

Ekiz B and Kocak O. (2007) *Turkish J. Vet. Anim. Sci.* **31**: 1.

Hill E.W., Gu J., Eivers S.S., Fonseca R.G., McGivney B.A., Govindarajan P., Orr N., Katz L.M., and MacHugh D. (2010) *Plos One* **5**: e8645.

R Development Core Team (2011) 'R: A Language and Environment for Statistical Computing', R Foundation for Statistical Computing, Vienna Austria.

Silveira M.D.D.M. and Ferreira M.D.G. (2008) *UDO Agricultural Science J.* **8**: 127.

Velie B.D., James J.W., Hamilton N.A., and Wade C.M. (Under Review) *Anim. Genet.*

Williamson S.A. and Beilharz R.G. (1996) *J. Anim. Breed. Genet.* **113**: 505.

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:495-498

GENOME WIDE ASSOCIATION STUDY USING THE OVINE SNP50 BEADCHIP AND LAMBS SELECTED FOR EXTREMES FOR CARCASS LEAN MEAT YIELD

P.L. Johnson, T.C. Van Stijn, H. Henry, N.J. McLean and M. Lee

AgResearch, Invermay Research Centre, PB 50034, Mosgiel 9053, New Zealand

SUMMARY

A phenotype resource of extremes for carcass lean meat yield in lambs was established in an attempt to identify regions of the genome associated with carcass lean meat yield. Data were available on 1150 lambs genotyped using the Ovine SNP50 BeadChip. Only two SNPs reached nominal significance (P-values of the order 10^{-8}), with both on Chromosome 2 in the region of GDF8, the gene which contains the already known GDF8 c.1232 *G>A* mutation. The c.1232 *G>A* SNP is not on the Ovine SNP50 BeadChip, with the closest SNPs 10 to 30kbp away, however, these proximal SNPs do not appear to be in LD with c.1232 *G>A* and were not identified in the analysis. The most significant SNPs identified actually lie 2 to 4Mbp away from GDF8, but are in higher LD with c.1232 *G>A*. Models were developed to test whether or not the significance of the SNPs was due to their LD with c.1232 *G>A*. The c.1232 *G>A* genotypes, fitted in models, explained a large proportion of the difference, however, there was still significant residual carcass yield variation explained by these SNPs. This suggests the presence of other SNPs within GDF8 or in neighbouring genes affecting carcass yield, a hypothesis supported by work in cattle and other sheep resources. Using the resource developed no other genomic regions containing significant QTL were identified. The ability to detect other smaller QTL that account for less of the genetic variation may require an increased sample size and/or information from higher density SNP chips.

INTRODUCTION

Data generated from the Ovine SNP50 BeadChip (www.sheephapmap.org) can be used in several ways, including the development of Molecular Breeding Values, or the identification of genomic regions which explain either all of the variation in monogenic traits or a large amount of variation in polygenic traits. For the former, there are already a number of publications that report the identification of the gene causing various monogenic disorders such as Arthrogryposis, Achondroplasia and Progressive Muscular Dystrophy through the selective genotyping of case and control animals using the Ovine SNP50 BeadChip. Such studies often require only small numbers of animals exhibiting the disorder to be genotyped. There are fewer published reports on the ability of the Ovine SNP50 BeadChip to detect regions that explain variation in polygenic traits; one such report is the identification of SNPs on OAR4 associated with average daily gain, staple length, wool grade, and fleece weight in Rambouillet sheep (Hadfield *et al.* 2012). The genotyping of different breeds using the Ovine SNP50 BeadChip has also been used to look at breed differentiation and signature sweeps, and one often reported is the signature of Texel sheep in the region of Myostatin (GDF8) on OAR2 (Kijas *et al.* 2012).

An industry-derived data set was collected for animals containing phenotypic extremes for meat yield with the aim of identifying genomic regions that have a large effect on the trait.

MATERIALS AND METHODS

A description of the data set has previously been provided by Johnson *et al.* (2011). Briefly, data were collected in 2008 and 2009 at Alliance Group Ltd meat plants. Lambs were selected from large mobs of greater than 200 lambs, with carcass weights between 15.5 and 19kg. One to three of the most extreme yielding pairs of animals (high and low, matched for carcass weight) were identified from a total of 344 mobs. No information about breed, age or origin was available

on the lambs. Measurements recorded on the whole carcass are described by Johnson *et al.* (2011) and included carcass weight, carcass length, buttock circumference (BC) and VIAscan® carcass measurements of the lean meat yield of the leg, loin, and shoulder expressed as a percentage of the carcass weight, together with their sum total.

The lambs were genotyped using the Illumina OvineSNP50 BeadChip (www.sheephapmap.org), and for the GDF8 c.1232 *G>A* variant with data analyses were performed using the R package GenABEL (Aulchenko *et al.*, 2007). Illumina OvineSNP50 BeadChip genotype and phenotype data were available on 1,150 lambs made up of pairs of high and low yielding extremes representing 344 mobs. The data for each trait was adjusted for sex, year of birth, and mob together with the first six principle components fitted to the autosomal SNPs from each animal's 50K SNP data. A polygenic model using a kinship matrix was calculated, by the 'ibs()' function of GenABEL using the weight = 'freq' option, from all autosomal makers as identity by state (Aulchenko *et al.* 2007).

The residual from this mixed model was tested for association with each SNP independently (Chen and Abecasis, 2007). Principal component analysis (PCA; prcomp function in R) or classical multidimensional scaling (CMD) was used to calculate principal components or co-ordinates to check for further population structure and outliers.

Further investigation of SNPs identified as significant was conducted using SAS (SAS, 2004), using the General Linear Model procedure based on models described by Johnson *et al.* (2011), with the SNP of interest fitted as a fixed effect. An additional model for each SNP involved firstly fitting the GDF8 c.1232 *G>A* genotype of the animal as a fixed effect.

RESULTS AND DISCUSSION

Of the traits assessed the most significant results were observed for BC with evidence of a large peak on Chromosome 2 (Figure 1). Two SNPs (s02728 and OAR2_128772350) contributing to this peak had nominal P-values of the order 10^{-8}, a level of significance of interest in GWAS studies (Barsh *et al.* 2012). There were also peaks on Chromosome 2 for other lean meat yield related traits. For example, total yield estimated by VIASCAN® had similar results to BC, with the most significant SNPs, which were generally the same SNPs as for BC, having nominal P-values in the order of 10^{-7}. Interestingly, no associations of this order of significance were discovered outside of the region of the GDF8 locus for the traits analysed.

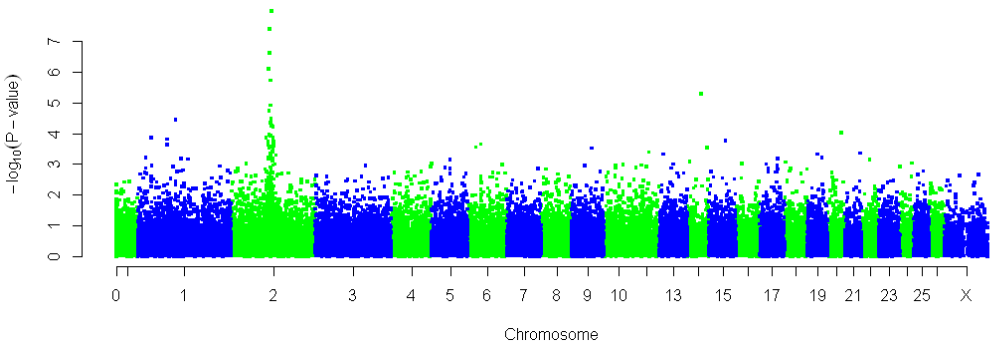


Figure 1. Manhattan Plot for Buttock Circumference in a Lean Meat Yield Extremes Industry Derived Sheep Resource

The region defined by the most significant peaks contains GDF8 (Figure 2, Top), with the SNPs 4Mbp and 2Mbp either side of GDF8. The known GDF8 c.1232 *G>A* mutation derived from

Texel sheep are not on the Ovine SNP50 BeadChip, with the closest SNPs (OAR2_125305996 and OAR2_126354465) 10 and 30kbp either side of GDF8 (Figure 2, Bottom). The effect of c.1232 *G>A* on carcass traits in this resource was reported by Johnson *et al.* (2011). Whether or not the SNPs identified are simply acting as markers for c.1232 *G>A* was investigated through looking at the level of LD between the SNPs (results not presented). This analysis showed that there was actually a higher level of LD (but not perfect) between the significant SNPs identified in the analysis and c.1232 *G*>A than the most proximal SNPs and c.1232 *G>A.*

The results from further analysis of the most significant SNPs are in Table 1 and Table 2 where the SNPs are fitted separately, and then with c.1232 *G>A* genotype fitted in the model prior to the inclusion of the SNP. Table 1 shows that most of the variation explained by the SNPs can be attributed to their LD (even though not perfect) with c.1232 *G>A*, but that there is also evidence of other mutations within this region affecting lean meat yield, via the presence of a residual significant effect. The presence of other mutations related to lean meat yield in the region is supported by evidence from cattle for multiple mutations within GDF8 and evidence from Johnson *et al.* (2005) for a second QTL in the region of GDF8 in Texel sheep.

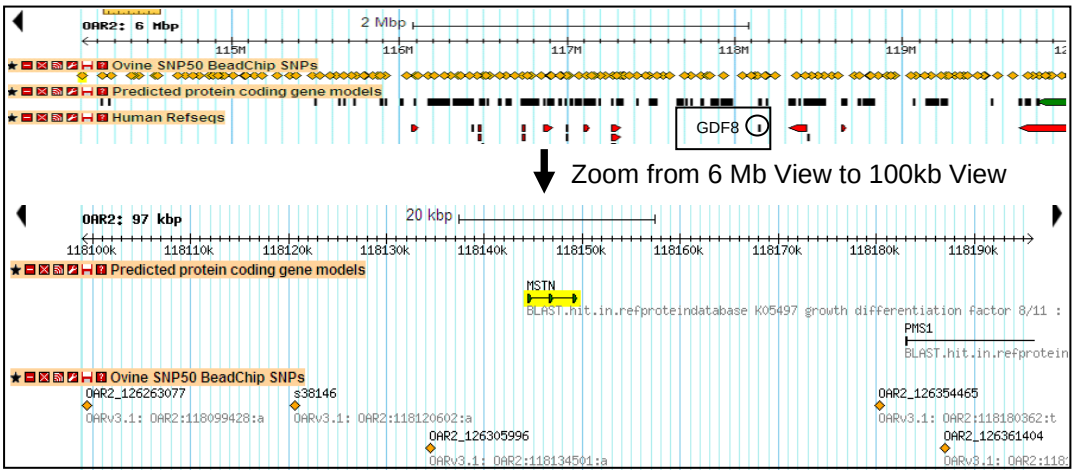


Figure 2. Illustration of Ovine Chromosome 2: Top, region that is bounded by the two most significant SNPs (s02728 and OAR2_128772350); Bottom, an enhanced view of the region immediately around GDF8 (MSTN)

Table 1: Significance of associations between Buttock Circumference (cm) and SNPs identified using the OvineSNP50 BeadChip, before and after inclusion of GDF8 c.1232 *G>A* genotype in the model

SNP	GDF8 c.1232 G>A Fitted First in Model	Partial R2	P Value	LSMeans For Genotype[1]		
				1	2	3
GDF8 c.1232 G>A		0.11	P<0.0001	64.4	63.4	62.4
s02728		0.11	P<0.0001	64.3	63.3	62.3
OAR2_128772350		0.13	P<0.0001	64.4	63.5	62.3
s02728	✓	0.02	P<0.0001	63.9	63.5	62.8
OAR2_128772350	✓	0.03	P<0.0001	64.0	63.6	62.7

[1] 1 and 3 represent the two homozygous genotypes, and 2 represents the heterozygous genotype

Table 2: Significance of associations between Total ViaSCAN® (lean?%) and significant SNPs identified using the OvineSNP50 BeadChip before and after inclusion of GDF8 c.1232 *G>A* genotypes in the model

SNP	GDF8 c.1232 G>A Fitted First in Model	Partial R2	P Value	LSMeans For Genotype[1]		
				1	2	3
GDF8 c.1232 G>A		0.14	P<0.0001	58.0	55.3	52.9
s02728		0.14	P<0.0001	57.9	55.1	52.9
OAR2_128772350		0.14	P<0.0001	57.7	55.4	52.8
s02728	✓	0.02	P<0.0001	56.9	55.4	54.2
OAR2_128772350	✓	0.02	P<0.0001	56.4	55.9	54.3

[1] 1 and 3 represent the two homozygous genotypes, and 2 represents the heterozygous genotype

CONCLUSIONS

Genotyping of carcass lean meat yield extreme lambs using the Ovine SNP50 BeadChip confirmed the region of GDF8 Ovine Chromosome 2 as influencing lean meat yield. In addition this study provides evidence for other mutations in the region that affect carcass lean meat yield, in addition to GDF8 c.1232 *G>A*, albeit they have a smaller effect.

That no other genomic regions were identified suggests there are unlikely to be QTL of a similar magnitude, at a reasonable allele frequency, segregating in the population studied. Increasing the sample size and/or the use of information from higher density SNP chips may be valuable to examine other regions in the genome that have QTL that account for less of the genetic variance in the traits studied.

ACKNOWLEDGMENTS

This work was funded by Ovita Ltd. Nessa O'Sullivan, Allison Knowler and Ryan Tecofsky carried out the data collection. The authors would like to acknowledge the staff at Alliance plants, Mataura, and Lorneville, for their help in data collection, especially Karen Thompson for providing the download for the VIAscan® data.

REFERENCES

Aulchenko Y. S., Ripke S., Isaacs A. and van Duijn C.M. (2007) *Bioinformatics* **23**: 1294.

Barsh G. S., Copenhaver G.P., Gibson G. and Williams S. M. (2012) *PLoS Genet* **8**: e1002812.

Chen W. M. and Abecasis G.R. (2007) *Am J Hum Genet* **81**: 913.

Hadfield T. *et al.* (2012). *PAG* 2012. p P0578.

Johnson P. L., McEwan J.C., Dodds K.G., Purchas R.W. and Blair H.T.. (2005) *J Ani Sci* **83**: 1988.

Johnson P. L., Van Stijn T.C., Henry H. and McLean N.J. (2011) *Proc. Assoc. Advmt. Anim. Genet.* **19**: 219.

Kijas J. W. *et al.* (2012) PLoS Biol **10**: e1001258. SAS. (2004). Sas/stat 9.1 user's guide. SAS Publishing.

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:499-502

GENOMIC REGIONS ASSOCIATED WITH DIFFERENCES IN FAT PERCENTAGE IN MILK BETWEEN HOLSTEIN AND JERSERY CATTLE

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

[1]Department of Agriculture and Food Systems, Melbourne School and Land and Environment, University of Melbourne, Parkville, VIC, 3010; [2]Department of Primary Industries, AgriBio, Bundoora, VIC, 3083

SUMMARY

Holstein and Jersey cattle exhibit large phenotypic differences in milk traits such as percentage of fat in milk (fat%). However, the genetic basis for this differentiation is unknown. Past strategies have attempted to identify selection in the genome using distortions to neutral loci and then locating candidate genes in these regions. In this paper we use the predicted difference between Holstein and Jersey breeds for fat% in milk to identify genomic regions and then examine these regions for evidence of selection. We localise a small predicted breed difference in fat% to regions on chromosome 14 and 5 but find little evidence for selection in these regions.

INTRODUCTION

Long-term selection is expected to increase the frequency of favourable alleles over time and leave a selection 'signature' in the surrounding genome. Under the classic 'hitchhiker' model of Maynard-Smith and Haigh (1974), selected loci are swept rapidly to fixation and this causes a reduction in heterozygosity at neutral loci surrounding the selected mutation. This type of selection signature is found for mutations with large effects, such as the *IGF1* mutation affecting stature in dogs (Stutter *et al.* 2007). However, evidence supporting this model for polygenic traits is limited (Pritchard, Pickrell and Coop 2010). This is because polygenic traits are influenced by hundreds or thousands of loci, each with relatively small phenotypic effect. Under these conditions, selection may cause only a small increase in the frequency of favourable alleles at many loci and leave little evidence for a selection signature.

Holstein and Jersey cattle differ markedly in fat% in milk, presumably due to differences in past selection. Different selection histories should leave evidence of selection in surrounding neutral loci. In this paper we introduce a novel method for identifying regions of the genome subject to past selection. We use predictions of the effect of single nucleotide polymorphisms (SNPs) on fat% to find regions of the genome where Holsteins and Jerseys are predicted to differ in genetic value for fat%. We then examine these regions for two traditional signatures of selection – large between breed allele frequency differences at neutral markers (i.e. high F_{ST}) and reduced SNP heterozygosity in either breed.

MATERIALS & METHODS

Data. Phenotypes and genotypes for 616,350 single nucleotide polymorphisms (SNP) were available for 2,767 Holstein and 825 Jersey bulls. All genotypes were quality checked, imputed from 50K to high density (as required) and phased with BEAGLE (Browning and Browning 2007) following Erbe *et al.* (2012). Phenotypes were daughter-trait-deviations from the Australian Dairy Herd Improvement Scheme for milk and fat yield from which daughter deviations in fat% were calculated. Holstein cattle have, on average, 1 % lower fat in milk compared to Jersey.

Estimating SNP effects for fat%. The effect of each SNP on fat% (b-hat) was estimated using BayesR, fitting the mean, SNP effects and a (residual) polygenic variance following Erbe *et al.* (2012). We analysed Holstein and Jersey bulls together in an analysis with 30,000 iterations and 20,000 discarded as burn in. Fat% was standardised within breed [i.e. $(x_i - \mu)/\sigma$] to have a

mean of zero and a phenotypic standard deviation of 1 prior to estimating the SNP effects. Therefore, we estimated within breed effects for the SNPs. SNP effects were the posterior mean of 5 replicate chains.

Identifying genomic regions predicting between breed differences in fat%. The autosomes were divided into sliding windows of 250 kb, with adjacent windows separated by 50 kb. The between breed difference in fat% was calculated as:

$$\sum_i p_{(Hol)i}\,\hat{b}_i - \sum_i p_{(Jer)i}\,\hat{b}_i \qquad [1]$$

where i is the i^{th} SNP in the 250 kb window, $p_{(Hol)i}$ and $p_{(Jer)i}$ are the allele frequency of SNP i in Holsteins or Jerseys, and $\hat{b}_i$ is the estimated effect of the SNP on fat%. Thus the sign of the between breed differences predicts if Holstein (positive values) or Jersey (negative values) have a higher fat%. The top 1% of windows were selected for further investigation. Windows from the top 1% were merged into regions when windows were separated by ≤ 250 kb.

Testing genomic regions for evidence of selection. For each of the regions identified above, we added ±125 kb of flanking sequence and calculated the mean of Wright's F_{ST} and the mean haplotype heterozygosity. Wright's F_{ST} measures the degree of allele frequency change between two populations and it was calculated per SNP following Weir and Cockerham (1984) as:

$$F_{ST} = \frac{\left(\overline{p_j^2} - \bar{p}^2\right)}{\bar{p}(1-\bar{p})} \qquad [2]$$

where p_j is the allele frequency in either Holstein or Jersey, $\bar{p}$ is the average allele frequency from Holsteins and Jerseys at the locus and $\overline{p_j^2}$ is the mean of the squared frequencies. The haplotype heterozygosity was calculated by dividing phased genotypes from BEAGLE into non-overlapping 30 SNP haplotypes and calculating 1-freq(homozygotes) in either Holstein or Jersey. P-values were determined by sampling 1000 random regions (of equal size as the observation), calculating F_{ST} and heterozygosity for these regions and determining the proportion of the sampled regions less than (or equal to) the observed F_{ST} or greater than (or equal to) the observed heterozygosity. Hence, $P < 0.05$ when the observed value was in the top (F_{ST}) or bottom (heterozygosity) 5 % of sampled regions. Finally, we re-calculated the between breed effect for fat% with [1] for the merged regions to avoid double counting of SNP from overlapping windows.

RESULTS

Between breed differences for fat%. For all SNP, the predicted difference in fat% was - 0.04 SD (i.e. - 0.002 %), implying a lower fat% for Holstein compared to Jersey cattle. This predicted between breed difference is much smaller than the observed phenotypic difference, probably because the SNP effects (b-hat) were estimated within breed. However, the direction of the between breed effect was consistent with phenotypic observations. Therefore, across all loci, Holstein cattle have a slightly higher frequency of alleles with negative effect on fat% than Jersey.

There were 510 windows identified from the top 1% of windows contributing to the between breed differences in fat%. The effects (per window) from the top 1 % had effects of between 0.003 and 0.36 SD. The 510 windows were consolidated into 110 genomic regions of up to 21 windows, from 250 kb to 1.4 Mbp.

Most (6/8) regions with between breed effects > 0.01 SD predict a slightly higher fat% in Holsteins than Jerseys (Table 1) but the largest effects on BTA5 and BTA14 predict a lower fat% in Holstein. These two regions potentially cause the lower fat% in Holstein, relative to Jersey.

For the measures of selection across all locations, the mean F_{ST} between Holstein and Jersey was 0.07 and haplotype heterozygosity was higher for Holsteins (mean heterozygosity = 0.84) compared to Jersey (mean heterozygosity = 0.75). However, the regions identified with breed differences in fat% showed little evidence for selection in the form of high F_{ST} or low heterozygosity (Table 1). In particular, the regions identified as contributing a large relative

increase in fat% for Jersey on BTA5 and 14 show do not show genomic evidence of selection for either measures of selection (Figure 1).

Table 1. Regions with large breed differences in fat% of milk for Holstein and Jersey cattle. Reported is the location, average F_{ST} and heterozygosity for each region

BTA	Region location (Mbp)	F_{ST}		Het. Holstein		Het. Jersey		Avg. effect (SD) (Hol-Jer)
3	15.25-15.7	0.146	*	0.726	*	0.43	*	0.037
3	16.55-17.2	0.046		0.817		0.78		0.026
5	93.35-94.35	0.099		0.885		0.78		-0.179
13	46.05-47.45	0.098		0.737	*	0.782		0.025
14	1.6-2.5	0.054		0.814		0.822		-0.358
14	2.6-3.15	0.058		0.815		0.725		0.020
19	42.6-43.2	0.076		0.852		0.747		0.020
20	33.9-34.8	0.095		0.735	*	0.754		0.063

$^{\#}P \leq 0.1$; $^{*}P \leq 0.05$

Several regions have previously been identified within Holsteins as associated with fat% in milk. Notably, the two regions with the extreme between breed differences for fat% on BTA14 and 5 contain the well-known *DGAT1* mutation (~1.8 Mbp; Grishart *et al.* 2004) and a region previously associated with fat% by Cole *et al.* (2011).

One of the largest regions associated with between breed differences in fat% was on BTA20 (30.9 - 32.3 Mbp), surrounding the growth hormone receptor gene (*GHR*, ~32 Mbp). This gene has been previously identified as associated with milk yield and composition (Blott *et al.* 2003). However, there was almost no predicted difference in fat% between the breeds over this region because it contained windows that predicted a higher fat% in Holstein and other windows predicting a high fat% in Jerseys and when summed across the whole region the effects tended to cancel out.

DISCUSSION

Large predicted between breed differences occur when a difference in allele frequency between Holstein and Jersey coincided with a large estimated SNP effect for fat% (i.e. see eq. [1]). The regions with the largest between breed differences in fat% were on BTA5 and 14, where previous studies have also identified genetic markers associated with fat%. However, we did not observe evidence of selection through increased F_{ST} or reduced heterozygosity in these regions. This could be because selection has not caused a big enough change in allele frequency between Holstein and Jersey or because the causative mutation is very old. If the favoured mutation is old, the linkage disequilibrium on the selected haplotype may have broken down (through mutation and recombination) or it could have existed on multiple haplotypes prior to selection.

This is a preliminary study which aimed to investigate if selection for a polygenic trait could be associated with regions of the genome. Our approach first identified regions where within breed QTL segregate with different SNP allele frequencies in Holstein and Jersey cattle, and then tested these regions for evidence of selection. This approach is similar to humans studies, where height-associated SNP were found at different frequencies in European populations and the allele frequency differences were attributed to selection (Turchin *et al.* 2012). However, although we identified some regions which could contribute to between breed differences, we found no evidence for selection surrounding these loci. Our approach could be improved by using different measure of selection (such as extended haplotype heterozygosity) or a different method to identify QTL responsible for between breed differences. For example, our analysis may be weak when the

linkage disequilibrium between the QTL and SNP varies between Holstein and Jersey. It is also likely that selection and drift has driven alternative alleles underlying between breed differences to fixation (or extreme frequencies) in our two populations. Such regions could be identified from studies with crossbred cattle.

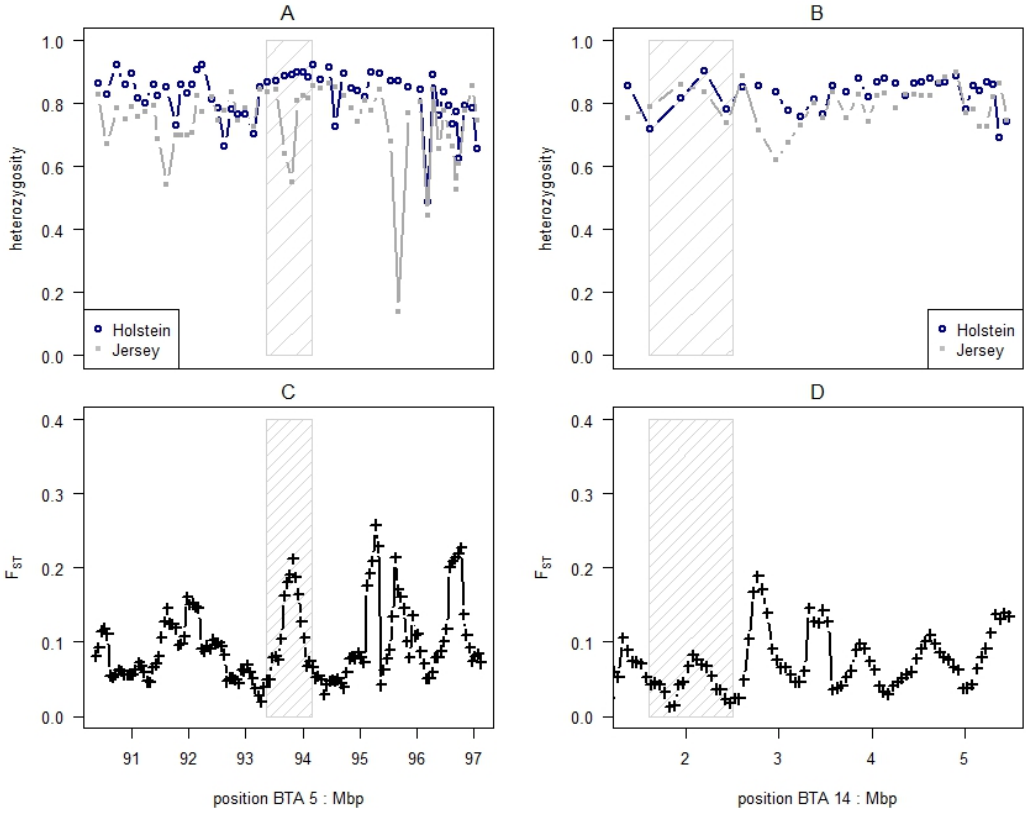


Figure 1. Heterozygosity (A, B) and F_{ST} (C, D) for regions with large between breed differences for fat% on BTA 5 (A, C) and 14 (B, D). F_{ST} is averaged over 250 kb windows

REFERENCES

Browning S.R. and Browning B.L. (2007) *Am. J. Hum. Genet.* **81:** 1084.
Cole J.B., Wiggans G.R. Ma L., Sonstegard T.S., *et al.* (2011) *BMC Genomics* **12:** 408.
Blott S., Kim J.J., Moisio S., Schmidt-Kuntzel A., Cornet A. *et al.* (2003) *Genetics* **163:** 253.
Erbe M., Hayes B.J., Matukmalli L.K., Goswami S., *et al.* (2012) *J. Dairy Sci.* **95:** 4114.
Grisart B., Farnir F., Karim L., Cambisano N., Kim J.J., *et al.* (2004) *Nat. Acad. Sci.* **101:** 2398.
Prichard J.K. Pickrell J.K. and Coop G. (2010) *Curr. Biol.* **20:** R208.
Stutter N.B., Bustamante C.D., Chase K., Gray M.M., Zhao L. *et al.* (2007) *Science* **316:** 112.
Maynard-Smith J.M. and Haigh J. (1974) *Genet Res.* **23**: 23.
Turchin, M.C., Chiang C.W.K., Palmer C.D. *et al.* (2012) *Nat. Genetics* **44:** 1015.
Weir B.S. and Cockerham C.C. (1984) *Evolution* **38:** 1358.

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:503-506

SEARCHING FOR SNPS THAT AFFECT SHEEP ROBUSTNESS: *CYP17* SNP AFFECTS BEHAVIOURAL RESPONSES TO PSYCHOLOGICAL STRESS

D. Hough[1], P. Swart[1], J.J.E. Cloete[2,3] and S.W.P. Cloete[2,4]

[1]Department of Biochemistry, Stellenbosch University, Private Bag X1, Matieland, 7602
[2]Department of Animal Sciences, University of Stellenbosch, Private Bag X1, Matieland 7602
[3]Elsenburg Agricultural Training Institute, Private Bag X1, Elsenburg 7607, South Africa
[3]Directorate Animal Sciences: Elsenburg, Private Bag X1, Elsenburg 7607

SUMMARY

The ability of animals to adapt to stress is not only an animal health and welfare concern, but also influences reproduction potential and robustness. An important pathway involved in the stress response is the hypothalamic-pituitary-adrenal axis (HPAA) that results in the release of cortisol from the adrenal gland. In this study the cortisol responses of South African Merinos were measured to assess HPAA responsiveness to stress and relate it to behavioural stress responses to flock-isolation. The experiment was structured according to a 2×2 statistical design, with *CYP17* genotype (*WT1/WT1* vs. *WT1/WT2*) and selection line (H-line vs. L-line) as factors. Selection line criteria was based on divergent selection for (H-line) or against (L-line) maternal multiple rearing ability, where the H-line generally outperformed the L-line in terms of reproduction, animal welfare and resistance to certain pathogens. The *CYP17* genotype is involved in the biosynthesis pathway of cortisol. In the present study the *CYP17* genotype showed a significant influence on behavioural stress responses, where three parameters of the flock-isolation test were affected ($P<0.05$), namely the number of bleats uttered, the urinating frequency and the average distance from a human operator. It is suggested that the *CYP17* genotype affects behavioural responses via its effects on cortisol production, and that the SNPs located within the *CYP17* genotype may have application in marker-assisted selection of sheep.

INTRODUCTION

The ever increasing global population continues to place pressure on improving the 'efficiency' of animal production to meet local and global demands. It is, however, a difficult task to improve animal production in a commercial practice by means of genetic progress if the environment in which the animals are raised does not support the full expression of their genetic potential (Mormède *et al.* 2011). It is thus important to include robustness-related traits in breeding objectives, since 'robustness' is described as the ability to combine a high production potential with resilience to stressors, which allows for the unproblematic expression of a high production potential in a wide variety of environments (Beilharz 1998; Knap and Rauw 2009). In this respect, the hypothalamic-pituitary-adrenal axis (HPAA) plays an important role in adaptation to stress, via the release of cortisol from the adrenal cortex.

The Directorate Animal Sciences of the Western Cape Department of Agriculture at Elsenburg research farm in South Africa embarked on a strategic breeding program in 1986, where selection responses to divergent selection for maternal multiple rearing ability was assessed. The assumption was that selection for this trait would include characteristics for both fitness (increased lamb survival) and 'efficiency' of animal production (number of lambs reared per joining) as suggested by Snowder and Fogarty (2009). Two distinct Merino lines (upward selection: H-line vs. downward selection: L-line) were established that showed a marked divergent response in overall reproduction and animal welfare (Cloete and Scholtz 1998; Cloete *et al.* 2004; 2005a; 2005b; 2009; Scholtz *et al.* 2010; 2011). These lines differed in their behavioural responses to flock-isolation during an arena test (Cloete *et al.* 2005a; 2010), as well as in their cortisol responses to

insulin-induced hypoglycemic stress (Van der Walt *et al.* 2009; Hough 2012), where the H-line generally displayed a superior ability to adapt to stressful situations than the L-line.

The present study investigates the contribution of the *CYP17* genotype towards the cortisol response to physiological stress and its implications for behavioural responses to flock-isolation stress (psychological stress). The *CYP17* genotype is considered, since it encodes for an enzyme, namely cytochrome P450 17α-hydroxylase/17,20-lyase (CYP17), that plays a key role in the cortisol biosynthesis pathway (Miller and Auchus 2011). This paper studies the role of selection line and *CYP17* genotype in responses of sheep to induced hypoglycemia and flock isolation.

MATERIALS AND METHODS

Animals and breeding program. All animals belonged to a South African breeding program that commenced in 1986, where Merino sheep have been divergently selected for maternal multiple rearing ability (Cloete *et al.* 2004, 2009). Records of progeny from the 2001-2008 birth years were used, which included their behavioural performances in the arena test (Cloete *et al.* 2005a; 2010) and *CYP17* genotypes (for complete protocol of SNP genotyping, refer to Hough *et al*. 2013; Hough 2012). Records were grouped in a 2 X 2 factorial design, according to selection line (H- vs. L-line) and *CYP17* genotype (homozygous *WT1/WT1* vs. heterozygous *WT1/WT2*), and assigned the following abbreviations: H_E for H-line heterozygous *WT1/WT2*; H_O for H-line homozygous *WT1/WT1*; L_E for L-line heterozygous *WT1/WT2*; and L_O for L-line homozygous *WT1/WT1*. Arena test records were available for 260 H_E-, 74 H_O-, 53 L_E-, and 13 L_O-grouped sheep.

Hypoglycemic stress test. Merino sheep in the above mentioned breeding program (11 H_E-, 6 H_O-, 15 L_E-, and 6 L_O-grouped rams, 2-6 years of age) were subjected to intravenous administration of human insulin at a dose of 0.1 IU/kg body mass, after which blood samples were collected at times: 0, 15, 30, 60 and 120 min post challenge. Blood plasma glucose and free cortisol was determined with radioimmunoassay. Since it is not only the magnitude of the cortisol response that is important, but also the duration of the cortisol output the cortisol responses to hypoglycemic stress was expressed as the area under the curve (AUC) from measurements over 2 hours. Ethics approval was obtained from the Departmental Ethics Committee for Research on Animals (DECRA ref: R08/21) of the Western Cape Department of Agriculture.

Arena stress test. In this flock-isolation stress test, sheep entered a 10.6 m X 4.0 m arena (marked out in 18 squares) one-by-one, as described by Cloete *et al.* (2005a). The arena was surrounded by wooden panels to prevent escape, but still allowing visual contact with six contemporary group sheep on the opposite side of the arena, behind a split-pole fence, where a human operator was situated on a chair. The operator remained motionless, while the behaviour of the sheep was assessed for 3 minutes according to the following parameters: number of bleats, number of defecation events, number of urinating events, average distance from human operator (meters); and movement based upon the number of boundaries between squares that were crossed (crosses).

Statistical analysis. GraphPad Prism (version 4) software (GraphPad Software, San Diego, California) was used for all statistical analysis. The interaction effects of *CYP17* genotype and selection line was tested with a two-way analysis of variance and a Bonferroni's post-test for each recorded item in the arena test (average distance between the sheep and the human operator, movement in arena depicted by number of squares crossed, number of bleats, number of urinating events and number of defecating events). The selection line × *CYP17* genotype interaction was investigated with a two-way ANOVA of the AUC for the cortisol responses (normalized with glucose concentrations as measurement of the degree of hypoglycemic stress) of each subgroup.

RESULTS AND DISCUSSION

Hypoglycemic stress test. Comparison of the cortisol responses of the *CYP17* genotype × selection line groups with a two-way ANOVA indicated that the interaction between the selection line and the *CYP17* genotype was significant (P=0.0226). Differences in cortisol responses between the *CYP17* genotypes were only found within the L-line (Bonferroni post-test: P<0.05; 2751.5±57.5 AUC for the L_O-group vs. 1765.0±179.0 AUC for the L_E-group). In contrast, cortisol output was independent from *CYP17* genotype in the H-line (Bonferroni post-test: P>0.05; 2528.5±225.5 AUC for the H_O-group vs. 2610.5±37.5 AUC for the H_E-group). These results indicated that the effect of the *CYP17* genotype is dependent on the genetic background of the animal, since it has been shown that the H-line sheep have a superior HPAA function that allows them to adapt to stressful situations more effectively than L-line sheep (Hough 2012).

Arena stress test. The arena test performance of sheep in the H_E (n = 260), H_O (n = 74), L_E (n = 53) and L_O (n = 13) groups were compared. As seen by the sample size, the L-line sheep were poorly represented compared to the H-line, due to the effects of downward selection on the birth rate and survival of L-line animals. These statistics need to be improved in future studies. The behavioural response to stress was tested before one year of age (prior to exposure to various handling procedures) in lambs born from 2001-2008 of which the *CYP17* genotypes were known.

The effect of the selection line and *CYP17* genotype, as well as their interaction, was assessed with a two-way ANOVA, followed by a Bonferroni's post-test, for each arena test item (Table 1). It was found that the *CYP17* genotype (P<0.05), but not the selection line (P>0.05) or its interaction with the *CYP17* genotype (P>0.05), had a significant effect on three out of five of the arena test parameters, namely the number of bleats (P=0.0038, heterozygous *WT1/WT2*: 18.58±0.69 bleats vs. homozygous *WT1/WT1*: 13.76±1.15 bleats); number of urinating events (P=0.0083; heterozygous *WT1/WT2*: 1.45±0.15 events vs. homozygous *WT1/WT1*: 3.45±0.71 events); and the average distance of the sheep from the human operator (P=0.0192; heterozygous *WT1/WT2*: 4.50±0.17 meters vs. homozygous *WT1/WT1*: 4.26±0.28 meters). The animals of the L_E-group on average kept a longer distance (Bonferroni post-test: P<0.05; L_E: 4.21±0.28 meters) from the human operator (signal of stress) compared to the L_O-group (3.05±0.55 meters) that coincided with a higher cortisol response to hypoglycemia (superior stress response). The H_E-group uttered more bleats (17.50±0.76 bleats), but urinated less frequently (1.79±0.23 events) during the arena test than the H_O group (P<0.05; H_O: 12.82±1.23 bleats, 4.01±0.78 urinating events). Although the psychological stress responses of these two H-line groups seemed to be different, their responses to physiological stress (insulin-induced hypoglycemia) were the same. The line differences reported by Cloete *et al.* (2005a) were not evident in the present study. However, selection line tended to interact with *CYP17* genotype for the average distance from the human operator (P=0.06) and the number of defecating events (P=0.10). Future research on more animals is needed to elucidate the separate and combined effects of line and *CYP17* genotype.

It is known that cortisol can affect behaviour via its effects on the brain (Pryce *et al.* 1988; Da Costa *et al.* 2004; Dwyer *et al.* 2004). Results from the present study indicate that a higher cortisol response from the adrenal cortex is related to less stressed behaviours during flock-isolation, namely smaller average distances maintained from humans and fewer vocalizations of protest. The higher urinating frequency associated with the higher cortisol response might be ascribed to alterations in steroid hormone synthesis, which directs steroid biosynthesis towards cortisol production and away from aldosterone production. Subsequently there is an increase in hormonal signals (via the renin-angiotensin regulating mechanism) to increase urination. The remaining parameters in the arena test might be related to other complex traits, but would seem to not be related to the *CYP17* genotype.

Table 1. Summary of the behavioural responses of sheep during the arena test. Values depict means±SEM and *P*-values from two-way ANOVA with *CYP17* genotype (CG) and selection line (SL) as factors. Traits considered were: average distance from human operator (ADIS), number of crosses (NCROSS), number of bleats uttered (NBL), number of urinating events (NUR) and number of defecating events (NDEF).

	H-line		L-line		*P*-values		
Trait	***WT1/WT1***	***WT1/WT2***	***WT1/WT1***	***WT1/WT2***	**SL**	**CG**	**SL × CG**
ADIS	3.65±0.16	3.79±0.99	3.05±0.55	4.21±0.28	0.7301	0.0192*	0.0646
NCROSS	18.70±1.33	19.40±0.75	15.46±2.54	17.32±1.30	0.1709	0.5096	0.7650
NBL	2.82±1.23	17.50±0.76	11.69±3.11	18.42±1.43	0.9574	0.0038**	0.6017
NUR	4.01±0.78	1.79±0.23	4.54±2.45	2.49±0.68	0.4466	0.0083**	0.9116
NDEF	1.16±0.14	1.06±0.07	0.77±0.26	1.26±0.13	0.6093	0.2852	0.1018

CONCLUSIONS

The present results suggest that the *CYP17* genotype affects cortisol production and behavioural responses to psychological stress, where the presence of *WT1* seems to be more beneficial for adaptation to stress compared to the presence of *WT2*. The effect of the *CYP17* genotype, however, also depends on the genetic background of the animal to cope with stressors. More research is needed understand the interaction between selection line and *CYP17* genotype. The two SNPs within the ovine *CYP17* gene may have application via marker-assisted selection to improve the ability of sheep to cope with stress and to adapt to their environment more effectively.

REFERENCES

Beilharz R.G. (1998) *J. Anim. Breed. Genet.* **115**: 433.

Cloete J.J.E., Cloete S.W.P. and Hoffman L.C. (2005a) *Small Rumin.Res.* **60**: 227.

Cloete, J.J.E., Cloete, S.W.P. and Hoffman, L.C. (2010) Proc. 9th World Cong. Gen. Appl. Livest. Prod., 1-6 Aug 2010, Leipzig. http://www.kongressband.de/wcgalp2010/assets/pdf/0703.pdf.

Cloete S.W.P. and Scholtz A.J. (1998) *Aus. J. Exp. Agric.* **38**: 801.

Cloete S.W.P., Gilmour A.R., Olivier J.J. and van Wyk J.B. (2004) *Aus. J. Exp. Agric.* **44**: 745.

Cloete S.W.P., Scholtz A.J., Cloete J.J.E. and van Wyk J.B. (2005b) *Aus. J. Exp. Agric.* **45**: 1131.

Cloete S.W.P., Misztal I. and Olivier J.J. (2009) *J. Anim. Sci.* **87**: 2196.

Da Costa A.P., Leigh A.E., Man M. and Kendrick K.M. (2004) *Proc. Royal Soc. Lon. B* **271**: 2077.

Dwyer C.M., Gilbert C.L. and Lawrence A.B. (2004) *Horm. Behav.* **46**: 529.

Hough D. (2012) PhD Thesis, Stellenbosch University.

Hough D., Cloete S.W.P., Storbeck K., Swart A.C. and Swart P. (2013) *J. Anim. Sci.* **91**, 1193.

Knap P.W. and Rauw W.M. (2009) In 'Resource allocation theory applied to farm animal production', pp. 210-229, editor W.M. Rauw, CABI Publishing, Wallingford, UK.

Miller W.L. and Auchus R.J. (2011) *Endocr. Res.* **32**(1): 81.

Mormède P., Foury A., Terenina E. and Knap P.W. (2011) *Animal* **5**(5): 651.

Pryce C.R., Abbott A.H., Hodges J.K. and Martin R.D. (1988) *Phys. Behav.* **44**: 717.

Scholtz A.J., Cloete S.W.P., Van Wyk J.B., Kruger A.C.M. and Van der Linde T.C.deK. (2010) *Anim. Prod. Sci.* **50**: 203.

Scholtz, A.J., Cloete, S.W.P., Cloete, J.J.E., Kruger, A.C.M., van Wyk, J.B. and Van der Linde, T.C.de K. (2011) *Proc. Assoc. Advmt. Anim. Breed. Genet.* 19: 175.

Snowder, G.D. and Fogarty, N.M. (2009) *Anim. Prod. Sci.* **49:** 9.

Van der Walt D., Cloete S.W.P., Storbeck K. and Swart P. (2009) *Proc. Assoc. Advmt. Anim. Genet.* **18**: 100.

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:507-510

A SNP ASSOCIATION STUDY VERIFIES A MAJOR LOCUS FOR FIBRE DIAMETER RELATED TRAITS ON CHROMOSOME 25 IN SHEEP

H.W. Raadsma[1], E. Jonas[1,2] and P.C. Thomson[1]

[1] ReproGen, Faculty of Veterinary Science, The University of Sydney, Camden, NSW, 2570
[2] Swedish University of Agricultural Sciences – Department of Animal Breeding and Genetics, 750 07 Uppsala, Sweden

SUMMARY

The present paper reports on the fine mapping of a previously identified quantitative trait locus (QTL) for fibre diameter and fibre diameter variability on chromosome 25 in sheep. A medium-density SNP panel was used to narrow down a major linked region using 319 animals from an Awassi-Merino backcross population. We could narrow down the QTL region of interest to 5000 kbp (3000 to 6000 kbp) on chromosome 25. Strong pleiotrophic effects were previously seen for this QTL as linkage had been identified for mean fibre diameter, fibre diameter variability, proportion of coarse fibres and comfort factor. Histological examination of animals with extreme fibre diameter characteristics showed strong effects for mean diameter of primary follicles, and a much higher ratio of secondary to primary follicles (S:P ratio) for animals inheriting the fineness QTL allele. A possible mode of action of the QTL on secondary follicle branching has been proposed. A strong positional candidate gene has been suggested in a companion paper (Jonas *et al.* 2013-these proceedings), however, further investigations are needed for a better understanding of the underlying causal mutations.

INTRODUCTION

Fibre production is typical of sheep, which shows marked diversity across and within breeds. Across breeds a range of fleece types can be seen including coarse from short-hair types, medulated carpet-wool types, through to ultrafine apparel wool types. In some cases selection for desired fleece characteristics has come under intense selection as for example in specialised wool breeds such as Merino. The development of efficient breeding programs relies on the identification of appropriate traits for improvement; a precise knowledge of the genetic parameters and evaluating selection strategies. Despite our broad knowledge that fleece characteristics will respond strongly to directional selection (Atkins 1997), relatively little is known about underlying genetic architecture of genes contributing to such variation within and between breeds for all major fleece characteristics. Recent developments in molecular genetics have broadened our understanding of the genetic architecture of polygenic complex-traits under selection. A better understanding of the effects and magnitudes of allelic differences that influence these traits may significantly enhance overall response of selection by improved management of antagonistic and pleiotrophic effects. Fibre diameter is one such trait of major interest, yet limited QTL detection studies have been conducted for this trait (Purvis and Franklin 2004). The advent of high-density genotyping platforms for Single Nucleotide Polymorphisms (SNP) has opened the possibility to undertake high resolution mapping approaches exploiting variation between and within breeds.

In this paper, we report a high-density SNP marker association analysis using a paternal half-sib design within an Awassi × Merino resource population. Animals were genotyped using the ovine 50k SNP array to fine-map a QTL region with impact on fibre diameter reported previously (Raadsma et al unpublished).

MATERIALS AND METHODS

Animals and Phenotypes. A resource population derived from crosses between fat-tail Awassi (A) and small-framed Merino (M) sheep was established to exploit the extreme differences between these two types in a range of production characteristics (Raadsma *et al.* 2009). In the association study reported here, data from 319 Merino backcross ((Awassi x Merino) x Merino) progeny of the first F_1 sire were analysed in detail. All lambs were shorn at 16 month of age and wool quantity was measured from mid-side samples at week 75 by the Riverina Wool Testers in Wagga Wagga, Australia. Among many other traits, mean fibre diameter, variability in fibre diameter, percentage coarse fibres, prickle, and follicle curvature were recorded from animals of

this population. For the study presented here, only fibre diameter data was used for the association analysis. QTL transmission probabilities were calculated for all animals inheriting either the paternal copy of the "coarse fibre allele-Q" or the "fine fibre allele-q". From each population, 20 animals with the most extreme fibre diameter (highest mean fibre diameter in case of Q and lowest mean fibre diameter in case of q) were selected for histological examination. Histological measurements included mean fibre diameter and fibre diameter variability of primary (P) and secondary (S) follicles, total follicle density, P and S follicle density, and S:P ratio.

Genotyping. All 319 backcross animals and the F_1 sire were genotyped using the ovine 50kb SNP array (http://www.sheephapmap.org). The predicted map positions of each SNP were used to select a subset of 757 SNP on chromosome 25. Genotypes with a minor allele frequency < 0.05 and a call rate < 0.95 were excluded from the final analysis. PLINK (Purcell *et al.* 2007) was used to check the gender and pedigree information and estimate the similarities between sire and offspring. Inheritance of the SNPs was checked according to pedigree expectation and corrected using code written in 'R' (R Core Team 2012).

Association analysis. Two slightly different models were applied to the data for the association analysis. The 'identical-by-state' matrix (IBS) was obtained in PLINK (Purcell *et al.* 2007) and the factor of similarity between sire and each offspring from the IBS was used for further analysis in 'R'. The following model was first fitted to the data of chromosome 25:

$$FT_j = \beta_0 + \beta_1 SSim_j + SNP_i + \varepsilon_{ij} \qquad [1]$$

where FT_j = Fibre trait of offspring *j*; $SSim_j$ = Similarity between offspring *j* and the sire derived from the IBS matrix; SNP_i = *i*-th SNP (i = 1 to 757 SNPs used in the study); and ε_{ij} = residual random error term.

Additionally a model was used following the approach previously applied for QTL mapping. Similar to the QTL model in QTL-MLE used to identify linkage regions (Raadsma *et al.* 2009), the probability of allele '1' derived from the dam was deducted from the dataset and used as a fixed effect for the whole-genome association analysis using an alternative linear model in 'R'.

$$FT_j = \beta_0 + \beta_1 PDam_i + \varepsilon_i \qquad [2]$$

where FT_j = Fibre diameter of offspring *j*; $PDam_i$ = probability of allele 1 at SNP *i* transmitted from Dam; and ε_i = residual random error term.

Using these two slightly different models aimed to exploit some background information provided through the sire and or dam side. Firstly using additional information from the sire (Awassi x Merino) which aimed to utilize allelic information on the paternal side transmitted through the F1 sire. And secondly by using additional information on inherited maternal Merino alleles (1 or 2 alleles) more explicitly as fitted in model 2.

RESULTS AND DISCUSSION

Phenotypic data. The average fibre diameter was 24 μm, which equates to medium-strong wool, with values between 18 μm (equivalent to super-fine Merino) and 30 μm (equivalent to strong-wool Merino or crossbred wool type) (Atkins 1997).

Association analysis. Using the same resource flock, we have previously identified a number of QTL for various fleece quality and quantity traits using a genome wide linkage analysis with microsatellite markers as detailed by Raadsma et al (2009), among which a region on chromosome 25 stood out revealing one or more highly significant linkage regions (Raadsma et al unpublished). Details of initial QTL probability using microsatellite linkage analyses have been shown in Figure 1 for reference. However as no candidate gene had been described previously within the identified region in sheep and/ or comparative chromosomes in other species, further studies had been suggested to narrow down the region of interest. The study here aimed therefore to fine-map and to verify the previously identified QTL on chromosome 25. Models considering either Awassi or Merino influence on offspring genotypes showed significant association for fibre diameter in the region around 5000 kbp (3000 to 6000 kbp) on chromosome 25. Results of association analysis of with fibre diameter are shown in Figure 1 for each SNP positioned along the chromosome. The

results of both models were in good agreement. Results also verified the previous linkage region and could narrow down the significant associated region.

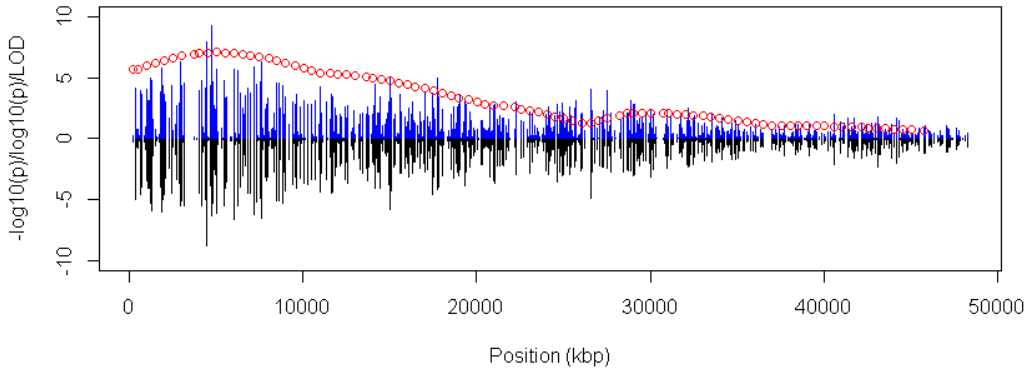


Figure 1. Results of the association analysis for fibre diameter using model [1] (above x-axis) and model [2] (under x-axis). Indicated with red dots are the LOD scores from the QTL study for fibre diameter from an unpublished linkage analysis using the same resource flock and data set

As reviewed by Purvis and Franklin (2004) only a limited number of QTL and GWAS studies have been conducted on fibre traits. QTL were identified for fibre diameter on chromosome 25 in a in a synthetic breed Merino based INRA401 (Ponz *et al.* 2001), a backcross Sarda x Lacaune sheep resource population (Allain et al. 2006), fine and superfine Merino sheep (Bidinost *et al.* 2008). This study confirms the importance of a major gene for fibre diameter characteristics and fleece quality. Within the positional candidate region, a positional candidate gene could be identified, results are shown elsewhere (Jonas *et al.* 2013).

Table 2. Contrasts of skin follicle characteristics in animals inheriting the paternal coarse-fibre (Q n=11) vs fine fibre (q n=9) QTL on OAR 25 from an Awassi-Merino to Merino backcross QTL mapping population

trait	Paternal coarse wool allele-Q		Paternal fine wool allele-q		% change
	mean	sd	mean	sd	(Qvs q)/q
Follicle density (n/mm**2)	54	8.4	73	13	+36
Density primary follicle(n/mm**2)	3.9	0.64	4.2	0.92	+7
Density secondary follicle(n/mm**2)	50	8.4	69	13	+38
Ratio S:P follicle	13	3.2	17	3.6	+30
Mean FD all follicles (um)	28	3.0	23	1.8	-19
SD FD all follicles (um)	8.0	0.7	4.0	0.8	-22
Mean FD primary follicles (um)	49	2.8	26	3.4	-46
SD FD primary follicles (um)	10	1.5	4.4	0.7	-56
Mean FD secondary follicles (um)	27	3.0	23	1.8	-15
SD FD secondary follicles (um)	5.1	0.72	4.0	0.81	-22

The mode of action of the QTL is likely to be through regulation of secondary follicle density during embryonic development given the large difference in secondary follicle density and ratio between secondary and primary skin follicles in animals with contrasting (“coarse fibre allele-Q” versus “fine fibre allele-q”) QTL alleles. Follicle branching is often thought to be a major characteristics of fine wool Merino sheep resulting in high secondary follicle populations in the skin. The main influence for differences of fibre diameter in the sheep used here was through a much higher S:P ratio in the animals inheriting the allele for fineness, suggesting a gene linked to

follicle development and possibly the branching of secondary follicles (Table 2).

CONCLUSION

The application of high density SNP assays to genotype animals of an ovine resource population showed high utility to provide high resolution mapping information for fine-mapping approaches. The results shown here did verify previous identified highly significant linkage or association on chromosome 25. In future studies we will implement population data and genetic similarity among offspring into the analysis. Results will also be tested using more families of the same sheep population.

ACKNOWLEDGEMENT

We acknowledge the support received from all current and previous staff involved in different stages of the project. We are grateful for the technical support provided by Ms. Gina Attard, Ms Marilyn Jones, and Mr David Palmer throughout this experiment, and for Ms Mary Lam for conducting initial QTL analyses.

REFERENCES

Allain D., Schibler L., Mura L., Barillet F., Sechi T., Rupp R., Casu S., Cribiu E., Carta A. (2006) *Proc. 8th World Congr. Genet. Appl. Livest. Prod., Belo Horizonte, Minas Gerais, Brazil*

Atkins K.D. (1997) In: *The Genetics of Sheep.* CABI, 471.

Bidinost F., Roldan D.L, Dodero A.B., Cano E.M., Taddeo H.R., Mueller J.P., Poli M.A. (2008) *Small Rum. Res.* **74:** 113.

Jonas E., Thomson P.C., Raadsma H.W (2013) *Proc. Assoc. Advmt. Anim. Breed. Genet.* **20**

Ponz R., Moreno C., Allain D., Elsen J.M., Lantier F., Lantier I., Brunel J.C., Perez-Enciso M. (2001) *Mamm. Genome* **12:** 569.

Purcell S., Neale B., Todd-Brown K., Thomas L., Ferreira M.A.R., Bender D., Maller J., Sklar P., de Bakker P.I.W., Daly M.J. and Sham P.C. (2007) *Am. J. Hum. Genet.* **81**:

Purvis I.W. and Franklin I.R. (2005) *Genet. Sel. Evol.* **37:** S97.

R Core Team (2012) R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org/.

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) *Genet. Select. Evol.* **41:** 34.

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:511-514

A BINARY CLASSIFIER USING SNP DATA FOR PREDICTION OF PHENOTYPIC OUTCOMES IN HANWOO (KOREAN) CATTLE

D.C. Detterer[1,2], S.H. Lee[3], P. Kwan[1] and C. Gondro[2]

[1]School of Science and Technology, University of New England, Armidale NSW 2351, Australia
[2]The Centre for Genetic Analysis and Applications, University of New England, Armidale NSW 2351, Australia
[3]Hanwoo Experiment Station, National Institute of Animal Science, RDA, Pyeongchang, Korea

SUMMARY

Korean *Hanwoo* cattle are prized for their high marbling ability and meat quality. Classically, these cattle possess a homogeneous yellow coat colouring, with farmers believing that *Hanwoo* with white spotted coats are crossbred and therefore unacceptable for breeding purposes. In this study we first attempted to determine if the coat spots were due to a non-*Hanwoo* genetic background or, alternatively, if the trait is intrinsic to the breed. By genotyping 232 (136 spotted) animals from half-sib families on the Illumina Bovine 50K SNP array, we compared the genotyped *Hanwoo* to other unrelated Hanwoo and European taurine breeds using principal component analysis. Results showed no evidence of crossbreeding in the spotted animals. A differential evolution algorithm was then used to evolve a classifier for the trait which selected 12 SNP with an accuracy of ~82% in separating individuals; further investigation using only haplotypes inherited from the sires resulted in a marked improvement to ~92% accuracy for these 12 SNP. This research highlights the potential for using these SNP as genetic markers to either entirely remove the trait from the population in the long term or manage matings so that the trait is not expressed in the offspring.

INTRODUCTION

The shift towards breed analysis via large scale genomic data has provided greater accuracy in prediction and opened avenues for a more complete understanding of the biology underpinning phenotypic traits useful for selection (Hayes *et al*. 2009; Habier *et al*. 2010).

The most common markers found throughout the genome are single nucleotide polymorphisms (SNP), single points of differentiation between individuals within a population. Through linkage disequilibrium, marker SNP associated with a disease or quantitative trait can suggest areas in the genome that demand further investigation (Carlson *et al*. 2004).

In the present research, after quality control filtering, we used 37,065 SNP genotyped on the Illumina 50K array for 232 Korean *Hanwoo* cattle derived from 28 sires. 136 of the 232 cases exhibited spots in their colouring—a trait deemed undesirable by breeders (see Brown & Lawrence 2010 for a study linking coat colouring to beef carcass grading). Note that none of the sires exhibited coat spots. An additional 229 animals from 5 European taurine breeds and other unrelated *Hanwoo* were also used in the analyses.

Currently, spotted *Hanwoo* cattle are simply culled due to a belief that these animals are not purebred. Although this is effective as a *brute force* method for removing undesired animals from the population, genomic technologies offer more efficient means of removing undesired traits by informing breeding choices. Ultimately, the objective is to breed-out the undesired trait rather than waste resources breeding cattle only to cull undesired offspring.

METHODS

A three-step process was used in this study. Firstly, principal component analysis (PCA) was conducted on the genomic data in order to ascertain the relationship between the spotted *Hanwoo*

cattle, the homogeneous *Hanwoo* population and other taurine breeds. Secondly, a differential evolution (DE) algorithm was used to search for a set of SNP within the genomic data which would classify the animals in the *Hanwoo* data set according to their status as either spotted or unspotted. Thirdly, the *Hanwoo* data set was phased into sire and dam haplotypes, with the SNP-based classifier derived in the previous step applied to the sire haplotypes.

In the first step, a Genomic Relationship Matrix (GRM) was used to characterise the relationship between individuals and the principal components of the matrix were then calculated.

In the next step of this investigation, *k*-means clustering was used to drive a DE algorithm as a strategy for stochastically selecting SNP that could separate between the spotted and unspotted animals.

DE is a heuristic in the family of evolutionary algorithms which also includes genetic algorithms, evolution strategies and evolutionary programming (Price *et al*. 2005; Fogel 2000). DE is a relatively straightforward algorithm to implement; it evolves real-valued vectors of parameters against an objective function. The purpose of the objective function is, in turn, to evaluate the "fitness" of a given vector in relation to how successfully its parameters solve a given problem.

In the present research, the objective function was based on *k*-means clustering, testing a set of SNP as to their ability to effectively separate the cases into spotted and unspotted cattle. Each real-valued vector in the DE algorithm represents a set of SNP using relative position indexing (Onwubolu & Davendra 2009); over the course of a run, SNP with greater predictive value in the clustering are given greater weight so that by the end of a run, the SNP selected collectively perform better as predictors. 100 runs of the DE/*k*-means algorithm were carried out in order to discover the most commonly selected SNP across the genome. From SNP selected 3 or more times, further exploratory runs of the algorithm were conducted to experimentally find a set of SNP that offered the greatest separation between spotted and unspotted cases.

As a final step of this analysis, an attempt was made to apply the classifier based on the selected SNP to the haplotypes each individual inherited from its sire.

The SNP data from the 232 *Hanwoo* cattle were phased into haplotypes inherited from the sire and dam respectively for each offspring. After removing any animals having unphased alleles on any of the selected SNP on the sire-inherited haplotypes, 60 animals remained, 33 of which were spotted.

RESULTS

As the initial hypothesis was that the spotted cattle were not purebred *Hanwoo* cattle but instead crossbreds, we ran a PCA of the genomic data. The purpose of this step in the overall analysis was to ascertain the relationship between the spotted *Hanwoo* cattle, the homogeneous *Hanwoo* population and other taurine breeds.

Applying PCA to the data resulted in a clear clustering separating the *Hanwoo* cattle from other breeds. Furthermore, spotted and unspotted Hanwoo form the same tight cluster—this suggests that the spotted *Hanwoo* are purebreds and not the result of crossbreeding. It is thus reasonable to assume that the genetics determining spotted and unspotted cattle are intrinsic to the breed itself. In Figure 1, the spotted and unspotted *Hanwoo* cattle clustered indistinguishably together on the left, with European beef and milk breeds forming clusters on the right.

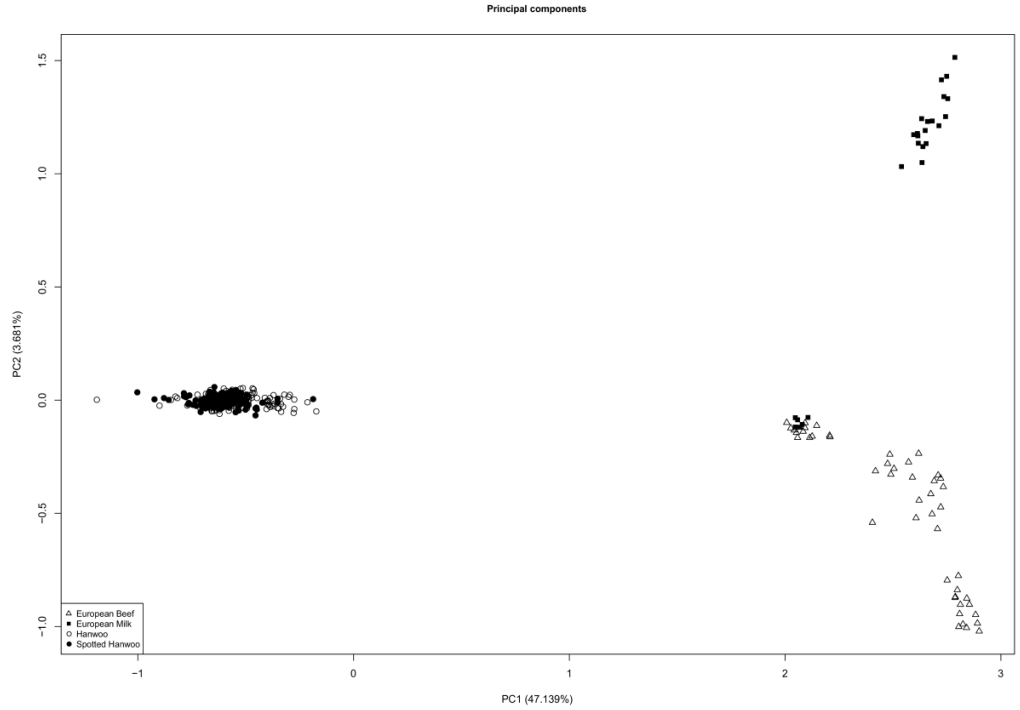


Figure 1. Distribution of animals according to the first two Principal Components of the SNP data set. Values in parenthesis are the variance explained by each component.

After applying the DE/k-means algorithm to the data set, 12 SNP were selected, giving a classification accuracy of ~82%; selecting more SNP did not increase the classification accuracy. In contrast, 10,000 random sets of 12 SNP selected from the 37,065 initial SNP resulted in classification accuracy with a mean of ~53% and a standard deviation of ~1.5%.

Table 1 lists the 12 selected SNP, the chromosome where each SNP is located and the position of the SNP on its respective chromosome.

Table 1. Selected SNP used in *Hanwoo* cattle classifier

SNSeP	Chromosome	Position on Chromosome
6149	3	120898833
7282	4	69056689
7827	4	106325551
7874	4	110856142
8800	5	58579368
13603	8	37461264
19453	12	901956
21220	13	42667485
27491	18	55895369
29257	20	44031834
34170	26	1632525
36876	29	41247528

Finally, performing *k*-means clustering on the phased sire haplotype data of the remaining 60 animals using the 12 SNP yielded a classification accuracy of ~92%. Given that the sires' coats were unspotted, this suggests the classifier may be effective at predicting the potential for spotted offspring given unspotted sires.

CONCLUSION

In this paper, an evolutionary algorithm was used to build a binary classifier for a phenotypic trait relevant to the Korean *Hanwoo* industry. Principal component analysis demonstrated that coat spotting in *Hanwoo* cattle is not due to crossing with other breeds, indicating that the trait is intrinsic to the breed. Using a differential evolution algorithm, 12 SNP were selected which were able to classify the cattle with an accuracy of ~82% via *k*-means clustering. Furthermore, classification on sire-inherited haplotypes gave an increase in accuracy to ~92%, suggesting that selective breeding based on SNP data is a viable path for removing the spotted trait from the population.

This is a promising start to a larger investigation into the utilisation of genomic markers to remove spotting from the population. However, the small sample size combined with a possible over-parameterization of the data means that independent validation is needed before the markers are adopted by industry.

Future work will involve validation of these results, including the use of cross-validation methods and further data gathering, which is currently underway. In addition, an investigation into the biology underpinning the relationship between the selected SNP and the spotted phenotypic trait will be undertaken.

Ultimately, the aim is to provide industry with a marker set to enable breeding decisions in the *Hanwoo* cattle industry.

ACKNOWLEDGEMENTS

This study was funded by grants from the Next-Generation BioGreen 21 Program (No. PJ008196), Rural Development Administration, Republic of Korea and the Australian Research Council project DP130100542.

REFERENCES

Brown, T.R. and Lawrence, T.E., (2010) *Professional Animal Scientist* **26**: 611.
Carlson C.S. et al., 2004. *The American Journal of Human Genetics* **74:** 106.
Fogel D.B. (2000) *IEEE Spectrum*, **37:**26.
Habier D. et al. (2010) *Genetics Selection Evolution* **42:**5.
Hayes B.J. et al. (2009) **41:** 51.
Onwubolu G. and Davendra D. (2009) *Differential Evolution: A Handbook for Global Permutation-Based Combinatorial Optimization*, pp.13–34.
Price K.V., Storn R.M. and Lampinen J.A. (2005) *Differential evolution*, Springer Berlin, Germany.

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:515-518

MULTI-TRAIT QTL MAPPING IN BEEF CATTLE

S. Bolormaa[1], J.E. Pryce[1], A. Reverter [3], Y.D. Zhang[4], W. Barendse[3] and M.E. Goddard[1,2]

[1]- Department of Primary Industries, Victoria; [2]- University of Melbourne, Victoria; [3]- CSIRO Animal, Food and Health Sciences, Queensland; [4]- Animal Genetics and Breeding Unit , NSW

SUMMARY

We report on the preliminary results of a multi-trait QTL mapping experiment using a genome-wide association study (GWAS) for 32 growth, feed efficiency, carcass, meat quality and reproduction traits of beef cattle. The GWAS were performed on 10,181 animals using the 800K Illumina SNP chip. The multi-trait analyses increased power to detect and map QTL. Each QTL appeared to have a pattern of pleiotropic effects across traits that was unique.

INTRODUCTION

Polymorphisms that affect complex traits or quantitative trait loci (QTL) often affect multi-traits, yet genome wide association studies (GWAS) are usually performed one trait at a time. When correlated traits are analysed independently the sampling errors tend to be correlated and this makes the interpretation of the results difficult. Also some account needs to be taken of the multi-trait testing that arises from performing many significance tests. Multi-trait analysis of linkage experiments has been reported to increase the power to detect QTL (Knott and Haley 2000; Korol *et al*. 2001). This paper investigates whether additional power can be extracted from a GWAS by analyzing traits together rather than one at a time.

Even if a QTL is detected for more than one trait in GWAS that is performed on single traits, it is possible that the most likely position for the QTL varies from trait to trait. Therefore we also consider whether multi-trait analysis can provide an increase in the precision of mapping QTL.

The obvious solution to the deficiencies of testing one trait at a time is a multi-trait analysis. However, typically not all animals have been measured for all traits and the individual animal data may not even be available. Therefore an approximate meta-analysis was used the estimates of SNP effects from individual trait GWAS.

The objectives of this study were to test a simple multivariate method to detect SNPs affecting beef traits, to understand the patterns of pleiotropic effects of genes that affect feed efficiency, growth, carcass, meat quality and fertility traits and to examine the ability of multi-trait analysis to increase the precision with which QTL are mapped.

MATERIALS AND METHODS

SNP data. In total, 729,068 SNP were genotyped. The SNP genotype data used in this study was a subset of Beef CRC genomic dataset. Details on genotyping, editing and imputation of the Beef CRC genomic data set has been described by Bolormaa *et al* (2013). A total of 10,181 animals with full genotypes and measured for at least one trait were used in this study.

Animals and population structure. The cattle were sourced from 9 different populations of 3 breed types. They include 4 different *Bos taurus* (Bt) breeds (Angus, Murray Grey, Shorthorn, Hereford), 1 *Bos indicus* (Bi) breed (Brahman cattle), 3 composite (Bt×Bi) breeds (Belmont Red, Santa Gertrudis, Tropical composites), and 1 recent Brahman cross population (F_1 crosses of Brahman with Limousin, Charolais, Angus, Shorthorn, and Hereford) (Bolormaa *et al.* 2013).

Traits. Phenotypes for 32 different traits were collated from 5 different sources including growth, feedlot, carcass, meat quality and reproduction. The trait definitions, number of records for each trait and heritability estimate and mean and its SD of each trait were reported by Bolormaa *et al.* (2013) and Zhang *et al.* (2013).

Statistical analysis. The association between each SNP and each of the traits was assessed by a regression analysis using the ASReml software (Gilmour 2009) and the following mixed model: trait ~ mean + fixed effects + SNP_i + animal + error; with animal and error fitted as random effect. Model details are given in Bolormaa *et al.* (2013) and Zhang *et al.* (2013). The effects of 729,068 SNPs were divided by their corresponding standard errors to calculate signed t values.

A multi-trait test of the effect of SNP i was conducted by storing the signed t-values for the 32 traits for SNP i in the vector $\mathbf{t_i}$. Then $\mathbf{t_i'V^{-1}t_i}$, where **V** is the correlation matrix among the SNP effects, is distributed as a chi-squared with 32 degrees of freedom under the null hypothesis that the SNP does not affect any of the traits. The correlation matrix **V** was approximated by the correlations among the estimated SNP effects across 729,068 SNPs. To avoid identifying a large number of closely linked SNPs whose association with traits is due to the same QTL, only the most significant SNP was retained from each 1Mbp interval. The most significant SNPs from the 2,523 1-Mbp-intervals were retained if it was significant at $P<10^{-4}$ and these SNPs were used to construct a new **V** matrix for use in clustering the SNPs into groups that have a similar pattern of effects on the 32 traits.

RESULTS AND DISCUSSION

In the multi-trait analysis 2,028 SNPs were significant ($P<5\times10^{-7}$), corresponding to a false discovery rate of 0.17%, and this was better than for any individual trait. When traits were analysed individually, for 29 out 32 traits the FDR was less than 2.5%. Therefore the multi-trait test did have greater power to detect QTL than the individual trait analyses. The multi-trait analysis was particularly successful in detecting QTL whose pattern of effects across traits was unusual.

Many highly significant SNPs from the multi-trait analyses were found within narrow regions on *Bos taurus* autosomal chromosomes (BTA) 3, 5, 6, 7, 14, 20 and 29 (Figure 1A). Many of the significant SNPs in both single trait and multi-trait analyses were linked and could be associated with the same QTL. When only the most significant SNPs in each Mb interval were retained, 418 SNPs were significant at $P<10^{-4}$.

A cluster analysis was performed on these 418 SNPs resulting in 12 clusters. Most clusters contained closely linked SNPs indicating that they were associated with the same QTL. Thus the long range LD that exists in cattle caused association between SNP and QTL separated by some Mb. The clustering of all SNPs in a region indicates that they all have the same pattern of effects across traits and therefore all detect the same QTL rather than multiple QTL each affecting an individual trait. However, the cluster analysis did separate the SNPs on BTA 7 into a group near 98Mb and a group near 93 Mb. The group at 98 Mb had a large effect on shear force whereas the group at 93 Mb had effects on muscling and fatness. Thus the analysis points to two separate QTL in this region of BTA 7.

The pattern of pleiotropic effects might be an important clue to the nature of the causative mutation and the function of the gene in which it occurred. Genes that operate in the same pathway might be expected to show the same pattern of pleiotropic effects. Therefore the patterns between QTL were compared to see if they fall into groups that might correspond to pathways. SNPs associated with different QTL seldom clustered together indicating that QTL seldom shared the same pattern of pleiotropic effects. However, there were some consistent patterns. For instance, SNP alleles that decreased shear force nearly always increased marbling. There was also a tendency for SNP alleles that increased hip height to increase weight and decrease fatness. However, this pattern was not consistent across all QTL.

Table 1 shows the effects of some of the significant SNPs that identify different QTL. One might describe these QTL as belonging to 3 groups. The first two QTL had a large effect on shear force and mapped to the positions of known genes affecting this trait (Calpastatin and Calpain 1).

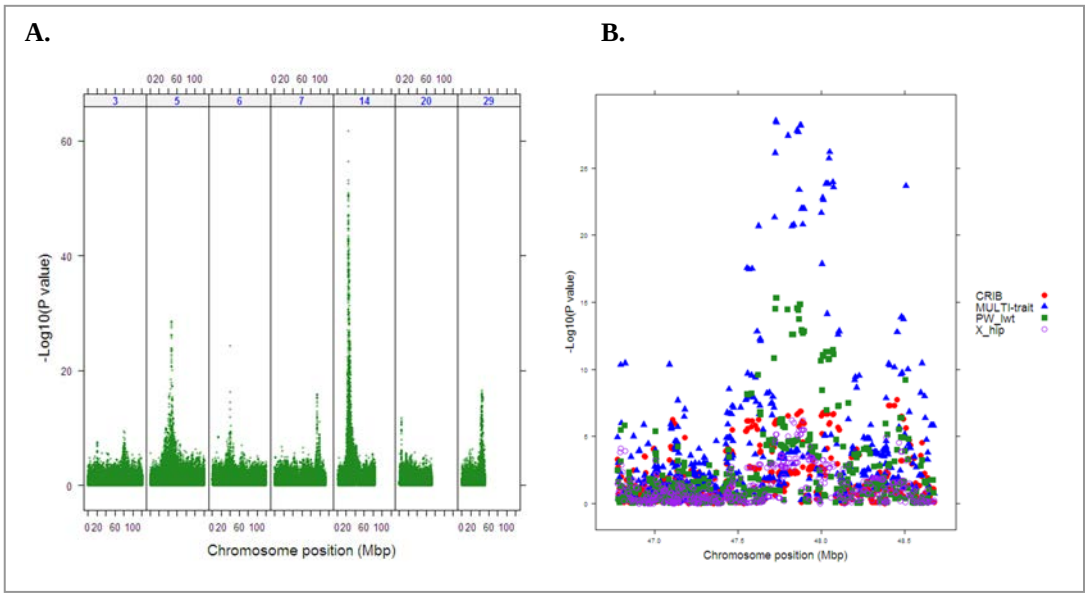


Figure 1. A. The –log$_{10}$(*P*-values) of multi-trait test based on GWAS for 32 traits on chromosomes 3, 5, 6, 7, 14, 20 and 29; B. The –log$_{10}$(*P*-values) of single SNP regressions for 3 traits and the multi-trait test on a region of chromosome 5.

The third SNP appeared to mark an unknown gene affecting shear force on BTA 6. The next 3 QTL could be affecting mature size. The allele of these 3 QTL which increased height also increased weight and decreased fatness. The last 3 SNP primarily affected fatness.

However, even QTL that have a similar pattern of pleiotropic effects, show differences in the detail of this pattern. For example, of the 3 'mature size' QTL, the one on BTA 14, which is presumably PLAG1 (Fortes *et al*. 2013), was the only one of the 3 mature size QTL that decreased shear force. It also had more marked effects on blood IGF concentration, fatness and reproduction than the other two SNPs. On the other hand, the QTL on BTA 5 had an unusual pattern of effects in that it redistributes fat from the P8 site to the rib and intramuscular depots. This QTL map was close to the gene HMGA2, which contains polymorphisms affecting growth, fatness and fat distribution in humans, mice and pigs (Anand and Chada 2000; Kim *et al*. 2006;Voight *et al*. 2010).

Table 1 also shows 3 SNPs associated with effects on marbling or intramuscular fat. There was a tendency for alleles that increase marbling to increase subcutaneous fat depth but this was not consistent. The QTL on BTA 7 had little effect on subcutaneous fat depth but a large effect on retail beef yield; the QTL on BTA 3 increased weight as well as fat; and the QTL on BTA 10 decreased shear force.

Based on these limited results, it would appear that each QTL has its own pattern of effects. Thus we have failed to discover groups of QTL that belong to the same pathway except for calpain and calpastatin. This could be explained if genes exist in a network rather than in pathways. Then each gene has a unique position in the network and therefore a unique pattern of effects.

The precision with which a QTL can be positioned on the genome in a GWAS is limited by two sources of errors. Firstly, the LD between SNP markers and the QTL is highly variable and therefore the nearest SNP is not necessarily the one in greatest LD with the QTL. Secondly, the

LD with the QTL is not observed directly but only via the effect of the QTL on a phenotypic trait.

Table 1. **Effect of some of significant multi-trait SNPs in the individual traits (signed t-values >1 are shown)**

chr	position	shear force	p8 fat depth	rib fat depth	intra-muscular fat	retail beef yield	weaning wt	IGF at weaning	weaning hip height	age at puberty in BB*	age at puberty in TC*
7	98540675	-8.6	1.1	1.4	1.5		1.7	2.1			
29	45778237	-10.5	2.9	2.5	4.1		-1.9				
6	68101121	-6.6	1.7	2.9	2.9	1.4			1.3	2.2	
5	47727773	1.9		-5	-4.4	-2	8.1	-1.9	9.6	1.2	3.3
6	40093712	1.7	-1.9	-2.6	-2.5		8.1	-2.6	9.5	2	1.1
14	25015640	-2.3	-7	-4.1	-1.6		9.8	-7.6	10.9	6.3	3.5
7	93007435	-2.5	-1.4		-3.2	6.3	2	-1.7	-2.4		2.9
3	80105316		1.1	3	2.5	-2	1.6				
10	89027305	-5.8			3.9		1.3	1.3			

* = Age at first detected corpus luteum in BB and TC

Because the QTL typically only explains a small amount of the variance of the trait , the effect of a SNP on the trait is estimated with error and this can also cause a SNP that is not the nearest to the QTL, to have the largest effect. By using more than one independent trait to map the QTL, the second source of error can be reduced but not the first source. Figure 1B shows the significance of SNPs from the multi-trait analysis and for 3 single trait GWAS in a region of BTA 5. The 3 separate traits map the QTL to slightly different positions and the multi-trait analysis may represent a good compromise.

ACKNOWLEDGEMENTS

We gratefully acknowledge the funding from the former Cooperative Research Centre for Beef Genetic Technologies for this study and ARC project DPI093502.

REFERENCES

Anand A. and Chada K. (2000). *Nat. Genet*.**24**:377.

Bolormaa S., Pryce J.E., Kemper K., Savin K., Hayes B.J., Barendse W., Zhang Y., Reich C.M., Mason B.A., Bunch R.J., Harrison B.E., Reverter A., Herd R.M., Tier B., Graser H.-U. and Goddard M.E. (2013) (submitted to *J. Anim. Sc.*).

Fortes M.R.S., Kemper K., Sasazaki S., Reverter A., Pryce J.E., Barendse W., Bunch R., McCulloch R., Harrison B., Bolormaa S., Zhang Y.D., Hawken R.J., Goddard M.E., and Lehnert S.A. (2013) *Anim. Genet.* (accepted)

Gilmour, A. R., B. J. Gogel, B. R. Cullis, and R. Thompson. (2009) ASReml User Guide Release 3.0 VSN International Ltd, Hemel Hempstead. HP1 1ES, UK.

Kim K.S., Lee J.J., Shin H.Y., Choi B.H., Lee C.K., Kim J.J., Cho B.W. and Kim T.H. (2006) *Anim. Genet.* **37**:419.

Knott S. A. and Haley C. S. (2000) *Genetics*. **156**:899.

Korol A.B., Ronin Y.I., Itskovich A.M., Peng J. and Nevo E. (2001). *Genetics*. **157**:1789.

Zhang Y.D., Johnston D.J., Bolormaa S., Hawken R.J. and Tier B. (2013) (submitted to *Anim. Prod. Sc.*).

Voight B.F., Scot L.J., *et al*. and McCarthy M.I. (2010) *Nat. Genet*.**42**:579.

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:519-522

HANDLING A SUBSET OF A LARGE DAIRY INDUSTRY DATASET FOR QUANTITATIVE GENETIC ANALYSES OF EXTENDED LACTATION TRAITS IN AUSTRALIAN DAIRY CATTLE

M. Abdelsayed[1], H.W. Raadsma[1], and P.C. Thomson[1]

[1] ReproGen Animal Bioscience Group, Faculty of Veterinary Science,
The University of Sydney, 425 Werombi Road, Camden NSW 2570, Australia

SUMMARY

Handling large-scale industry data is challenging both in the preliminary screening of data for validation, and also in the subsequent statistical analyses to obtain estimates of fixed effects, genetic parameters and estimated breeding values (EBVs). This paper summarises how a small subset from large industry-derived dairy data records have been explored and edited to be ready for genetic analysis prior to the analysis of a much larger data set, with a focus on lactation persistence and extended lactation traits in Australian dairy cattle. To cope with the large data volume (>158 million test-day records) test-day data were randomly divided into data-subsets, each of which could be effectively managed through a semi-automated quality control procedure for removal of extraneous outlier records. The methods as applied to one such data subset are reported here. The goals of this research were an investigation of extended lactation, but it was apparent that there were many instances of the calving date not being recorded, and hence initially being flagged as an extended lactation. Methods were derived for detecting potential "double lactations", based on comparing single- and double-lactation curve models. These techniques are illustrated using test-day records from a sample of ~1 million cows recorded between 1985 and 2010 obtained from the Australian Dairy Herd Improvement Scheme (ADHIS).

INTRODUCTION

In many areas of scientific study, and in society more generally, there has been a huge growth in the amount of data collected, the so-called "big data" phenomenon (Howe *et al*. 2008), particularly evident in the life sciences. In livestock industries, automatic recording of data has resulted in the availability of extraordinarily large industry data sets. They offer the promise of providing measures of new complex traits, as well as extremely accurate estimates of genetic parameters and breeding value predictions. However to attain these goals, some substantial difficulties need to be overcome. Firstly, automated methods of data cleaning need to be developed. Secondly, computationally efficient methods of data analysis need to be developed, as routinely used statistical methods may not "scale up" to deal with the massively increased data volume (Jacobs 2009). Also, before conducting any comprehensive analysis it is essential to explore the data to understand the overall trends, but more importantly the type of errors that can be encountered. This paper describes some approaches developed using a subset from large dairy industry data records for the analysis of test-day records from Australian dairy cattle milk yield data. A subset was used initially to trial different methods, for ease of handling and to obtain set steps which then can be applied automatically on the larger data file.

MATERIALS AND METHODS

For this study, dairy industry data were obtained from the Australian Dairy Herd Improvement Scheme (ADHIS) consisting of ~158 million test day records ranging from 1985 up to 2010. To handle the large volume of data, the dataset was randomly split into eight subsets, each subset consisting of ~1 million cows with ~20 million test day records, giving eight separate subsets. Splitting the data in this way ensured that all test-day records from the one cow were kept in the

same subset, so that all lactations from the one cow were in the same data file. Hence, the editing and analysis were then conducted on each of the eight data subsets separately.

For this paper and preliminary analysis, a randomly selected sample of 10,000 cows' data was selected; these were further filtered to only include Holstein Friesians (6,018 cows, 29,882 lactations). Analysis of this small data set would allow decisions of modelling techniques and automated data screening to be developed, and then applied to each of the eight large data subsets.

Graphical and numerical summaries from the sample data set revealed the necessity of log-transformation of some of these traits (e.g. SCC and lactose). Extreme outliers were filtered out (on transformed data where necessary) using a criterion linked to the number of records, specifically if it is more than k standard deviations away from the trait mean, where $k = |\Phi^{-1}(1/n)|$, n is the number of data records, and $\Phi^{-1}(\cdot)$ is the inverse cumulative distribution function of a standard normal distribution. For example for n = 30,287 test-day records as in the current example, k = 3.99, so four standard deviations from the mean would be an appropriate cut-off.

The ADHIS database records the calving date, and from this, days in milk (DIM) can be calculated for a particular recorded test-day. As a primary reason for undertaking this study was to explore variation in lactation curve shape, particularly those related to lactation persistence and extended lactation (Abdelsayed *et al*. 2013), cow-lactation data sets were removed from the analysis when fewer than three test-days records were available. Also, any test-day record beyond 750 DIM was excluded, as was any test-day record at birth (0 DIM).

An exploratory plot of the average milk yield over days in milk revealed a number of apparently extended lactations were in fact new lactations for which the calving date had not been recorded. This was important to ascertain, as this could substantially affect genetic estimates of lactation persistence and extended lactation traits. To identify this, a hypothesis testing method was used fitting one- and two-lactation Wood (1967) models to the test-day data. The Wood model as implemented here has the form $W(t; k, b, c) = \exp(k + b\log_e t - ct)$, where t is DIM, $W(\cdot)$ is the model-based mean yield, and k, b, and c are parameters that describe the lactation curve. Hence an observed milk yield for cow-lactation i on DIM j can be modelled as $y_{ij} = W(t_{ij}; k_i, b_i, c_i) + \varepsilon_{ij}$, where ε_{ij} represents a random error associated with the test-day record.

Single vs double lactation screening was conducted as follows, for each cow-lactation:

1. Fit a single-lactation Wood model to cow-lactation data set i, modelled as a single lactation: $y_{ij} = W(t_{ij}; k_i, b_i, c_i) + \varepsilon_{ij}$, and save the Residual SS (=RSS_0, reduced model).
2. Fit a two-component Wood model for lactation data set i, modelled as a double lactation, assuming the second lactation commenced 365 days after the first:

$$y_{ij} = \begin{cases} W(t_{ij}; k_i, b_i, c_i) + \varepsilon_{ij} & 0 < t_{ij} < 365 \\ W(t_{ij} - 365; k_i, b_i, c_i) + \varepsilon_{ij} & t_{ij} \geq 365 \end{cases}$$

   and save the Residual SS (=RSS_1, full model). This model assumes that the second curve has the same shape as the first, just separated by 12 months.
3. Calculate an F statistic and P-value based on comparing RSS_0 and RSS_1.
4. If $P < 0.1$, that could be sufficient evidence for a second lactation.

However, the specific threshold P-value needs to be evaluated to achieve a balance of false positives / false negatives. To assist this process, a false discovery rate approach can be used, and the q-value method of Storey and Tibshirani (2003) has been adopted here.

In the present study, the fitted Wood model was then used to summarise various characteristics of the lactation curve shape, such as persistence and extended lactation (Abdelsayed *et al*. 2013), and these are derived from estimates of k, b and c from each cow-lactation. Consequently cow-lactations with extreme or infeasible estimates of these were excluded, based either on the outlier method mentioned above, or in the case of c, excluding any cow-lactation with negative estimates (which would imply ever-increasing yield over a lactation).

RESULTS AND DISCUSSION

Figure 1 shows the average milk yield for all test-day records observed on each calculated DIM. The form of the curve is as expected for a lactation curve up to approximately day 270. Beyond that however, the smaller second peak is evidence of a number of cows not having their calving dates recorded, resulting in a false 'extended lactation'. This is also supported by the increase in variability of these means, not entirely explained by sampling fluctuations of fewer lactations recorded at the particular DIM. The wide variation in mean yield beyond 600 days however is a reflection of the fewer records extending that far.

For each cow-lactation that extended beyond 365 days, single- and double-lactation models were fitted to the test-day data, and the *P*-value calculated as a means of assessing if the fit of a double-lactation model was a better fit than a single lactation model. Lactations with $P < 0.1$ were considered for possibility of being a double lactation: two sample plots are shown in Figure 2, one being almost certainly a double lactation ($P \approx 0$), the other probably better considered a single lactation (with $P = 0.084$). Sample lactation curves were scrutinised, and $P = 0.06$ threshold was adopted: this corresponds to a false discovery rate of just in excess of 5% ($q = 0.054$); a distribution of *P*-values of 1,150 sample cow-lactations is shown in Figure 3.

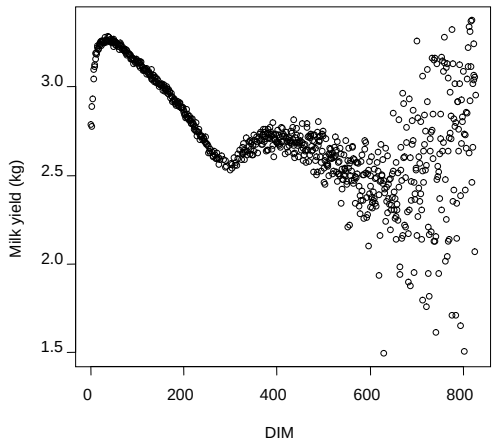


Figure 1. Mean milk yield (kg) vs days in milk (DIM). Evidence for calving date not being recorded is indicated by the second smaller peak around Day 400.

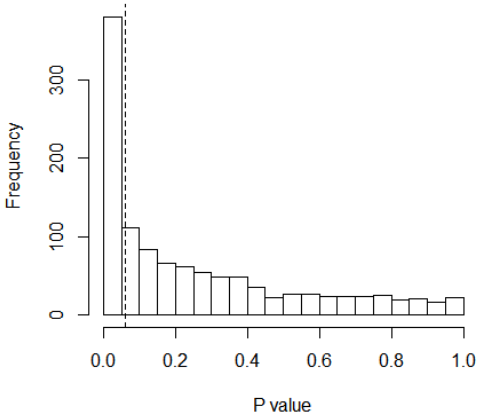


Figure 3. Histogram of *P*-values for testing double- vs single-lactations. The vertical dashed line is drawn at the adopted threshold of $P = 0.06$.

Using this method to classify double lactations, i.e. cows whose second calving was not recorded, there will undoubtedly be some false classifications. It should also be noted that for extended lactations, seasonal influences are apparent, and this has led some authors to model extended lactation with a di-phasic model, even in the absence of "double lactations" (e.g. Vargas *et al*. 2000; Grossman and Koops 2003).

As a result, after such cleaning processes, reliable estimates of the lactation curve parameters (*k*, *b*, *c*) for each cow-lactation have been obtained, and persistency and extended lactation traits can then be derived. These traits can then be used for the standard quantitative genetics analysis (Abdelsayed *et al*. 2013). However, due to computational limitations, when data volumes are large, it may not be possible to fit a large-scale linear mixed model, so it is necessary again to

randomly split the data into further subsets, perhaps using the same data subsets as used in the initial data screening. Because the allocation is random, overall genetic and fixed effect estimates can be obtained by simply averaging across those produced from each data subset, with appropriate weighting.

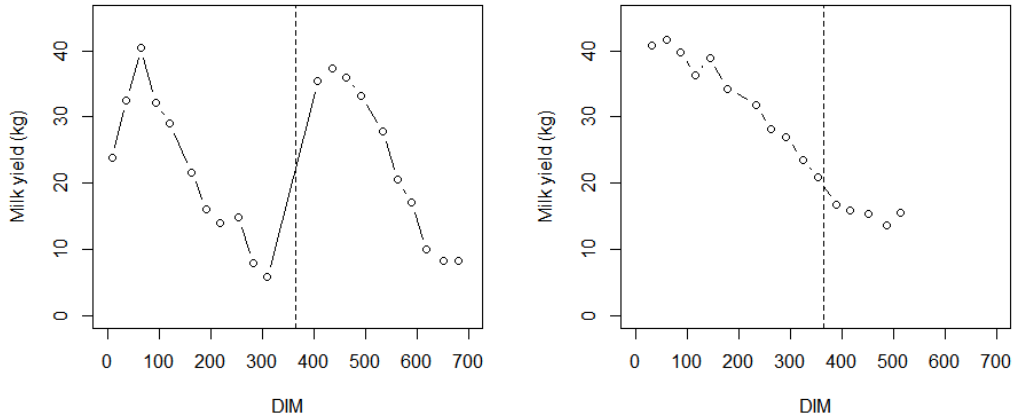


Figure 2. Test-day milk yields of two cow-lactations flagged as possible double lactations. The curve on the LHS, very clearly a double lactation, has $P = 2\times10^{-11}$, whereas the curve on the RHS has $P = 0.084$. The vertical dashed line is drawn at 365 DIM, the potential start of a second lactation.

CONCLUSIONS

Large-scale test-day datasets and other industry data are being routinely collected and in turn new computational and statistical approaches need to be developed to handle these "big data" sets. This paper has described some approaches to this, in the context of test-day records from Australian dairy cattle to assess lactation persistency and extended lactation. In particular, a method for screening for missed second lactations has been outlined, but other data screening and analysis aspects have also been considered. While a methodology has been outlined in this paper to address the problems encountered with extended lactation data, the process is not perfect, and there would be great benefit to investigate how all industry calving data information could be captured, ensuring that lactation length could be accurately evaluated.

REFERENCES

Abdelsayed M., Thomson P.C. and Raadsma H.W. (2013) *Proc. Assoc. Advmt. Anim. Breed. Genet.* **20** (this publ).

Grossman M., and Koops W.J. (2003) *J. Dairy Sci.* **86**: 988.

Howe D., Costanzo M., Fey P., Gojobori T., Hannick L., Hide W., Hill D.P., Kania R., Schaeffer M., St Pierre S., Twigger S., White O., and Rhee S.Y. (2008) *Nature* **455**: 47.

Jacobs A. (2009) *Commun. ACM* **52**: 36.

Storey J.D., and Tibshirani R. (2003) *Proc. Natl. Acad. Sci. USA* **100**: 9440.

Wood P.D.P. (1967) *Nature* **218:** 894.

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:523-526

SAMPLING BASED APPROXIMATION OF CONFIDENCE INTERVALS FOR FUNCTIONS OF GENETIC COVARIANCE MATRICES

Karin Meyer[1] and David Houle[2]

[1]Animal Genetics and Breeding Unit*, University of New England, Armidale, NSW 2351
[2]Department of Biological Science, Florida State University, Tallahassee, FL 32306-4295

SUMMARY

Approximate lower bound sampling errors of maximum likelihood estimates of covariance components and their linear functions can be obtained from the inverse of the information matrix. For non-linear functions, sampling variances are commonly determined as the variance of their first order Taylor series expansions. This is used to obtain sampling errors for estimates of heritabilities and correlations, and these quantities can be computed with most software performing such analyses. In other instances, however, more complicated functions are of interest or the linear approximation is difficult or inadequate. A pragmatic alternative then is to evaluate sampling characteristics by repeated sampling of parameters from their asymptotic, multivariate normal distribution, calculating the function(s) of interest for each sample and inspecting the distribution across replicates. This paper demonstrates the use of this approach and examines the quality of approximation obtained.

INTRODUCTION

Maximum likelihood (ML) theory indicates that ML estimates asymptotically have a multivariate normal (MVN) distribution with covariance matrix given by the inverse of the information matrix, i.e. the inverse of the matrix of second, partial derivatives of the likelihood function. Hence lower bound sampling errors of ML estimates are usually obtained from the diagonal elements of this matrix, and pertaining confidence limits are determined multiplying these values with the appropriate intercepts of a standard normal distribution. Corresponding statistics for linear functions of the parameters estimated are readily derived. For a non-linear function, the standard procedure is to replace the function with its first order Taylor series expansion and to calculate the variance of this linear approximation, a strategy sometimes referred to as the Delta method (e.g. Oehlert 1992). In genetic parameter estimation, this is used to approximate sampling errors of variance ratios and correlations, and is implemented in most restricted maximum likelihood (REML) software available.

In some cases, however, more complicated functions and their sampling distribution are of interest, which may not be approximated closely enough by a linear expansion. Others may involve variables afflicted by constraints on the parameter space or may simply not be accommodated by the facilities to calculate approximate variances of 'user-defined' functions of covariances available in software such as `ASReml` (Gilmour *et al.* 2009) and WOMBAT (Meyer 2007). A simple alternative then is to evaluate asymptotic sampling characteristics for such functions by repeated sampling of parameter estimates from their asymptotic, MVN distribution, calculating the function(s) of interest for each sample and inspecting their distribution(s) across replicates. This paper describes a suitable sampling strategy and examines the quality of approximation of sampling distributions obtained.

SAMPLING STRATEGY

Newton-Raphson type algorithms to maximise the REML (log) likelihood ($\log \mathcal{L}$) function utilize second derivatives of $\log \mathcal{L}$ and are well established as the most efficient methods available, especially the so-called average information variant (Gilmour *et al.* 1995). However, these involve an unconstrained optimization. Hence, estimation of covariance components is generally performed

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

employing a re-parameterisation to functions which do not require constraints to ensure positive (semi-) definite estimates of covariance matrices. A common choice for covariance matrices is to estimate the elements of their Cholesky factors, transforming diagonal elements to logarithmic scale (Meyer and Smith 1996). Furthermore, performing the factorization with pivoting on the largest diagonal readily facilitates reduced rank analyses (Meyer and Kirkpatrick 2005).

In addition, such parameterisation directly allows sampling of estimates of covariance matrices which are guaranteed to be in the parameter space, mimicking the constraints imposed in REML estimation. Let $\hat{\boldsymbol{\theta}}$, of length p, denote the vector of parameter estimates and $\mathbf{H} = \mathrm{Var}(\hat{\boldsymbol{\theta}})$ the corresponding inverse of the information matrix at convergence. Samples of parameters from $N(\hat{\boldsymbol{\theta}}, \mathbf{H})$ are obtained as $\tilde{\boldsymbol{\theta}} = \hat{\boldsymbol{\theta}} + \mathbf{L}_H\mathbf{d}$ with $\mathbf{L}_H$ the Cholesky factor of $\mathbf{H}$ and $\mathbf{d}$ a vector of standard normal deviates, $d_i \sim N(0, 1)$. Samples of covariance matrices can then be constructed from $\tilde{\boldsymbol{\theta}}$ by reversing the transformation.

APPLICATION

Data for 6 traits recorded on 4000 individuals in 500 independent families of size 8 were simulated for the design of Bondari *et al.* (1978). Population parameters assumed all residual correlations were equal to 0.3. Heritabilities were 0.2, 0.3 and 0.4 for two traits each, and all phenotypic variances were equal to 100. For Case I, all genetic correlations were assumed to be equal to 0.5, while for case II values for traits i and j were $0.7^{|i-j|}$.

REML estimates of genetic and residual covariances matrices were obtained fitting an animal model, using an average information algorithm. Three estimates of sampling (co)variances for covariance components and functions thereof were contrasted:

A) Values from the REML analysis, obtained from $\mathbf{H}$ using the Delta method. Let σ_{ij} denote the elements of a covariance matrix $\boldsymbol{\Sigma} = \mathbf{L}\mathbf{L}'$, with $\mathbf{L} = \{l_{ij}\}$ its Cholesky factor. For $\mathrm{Cov}(\hat{l}_{ij}, \hat{l}_{km})$ given by the corresponding element of $\mathbf{H}$, $\mathrm{Cov}(\hat{\sigma}_{ij}, \hat{\sigma}_{kl})$ is approximated as
$$\sum_{t=1}^{f(i,j)} \sum_{s=1}^{f(k,m)} \left[\hat{l}_{jt}\hat{l}_{ms}\,\mathrm{Cov}\left(\hat{l}_{it}, \hat{l}_{ks}\right) + \hat{l}_{jt}\hat{l}_{ks}\,\mathrm{Cov}\left(\hat{l}_{it}, \hat{l}_{ms}\right) + \hat{l}_{it}\hat{l}_{ms}\,\mathrm{Cov}\left(\hat{l}_{jt}, \hat{l}_{ks}\right) + \hat{l}_{it}\hat{l}_{ks}\,\mathrm{Cov}\left(\hat{l}_{jt}, \hat{l}_{ms}\right) \right]$$
with $f(i, j) = \min(i, j, r)$, and r the rank at which $\boldsymbol{\Sigma}$ is estimated. Similar formulations apply when diagonal elements l_{ii} are transformed to logarithmic scale or for covariances among components belonging to matrices $\boldsymbol{\Sigma}$ pertaining to different sources of variation.

B) Empirical values obtained by repeatedly sampling data for the given structure from appropriate normal distributions with population values equal to the estimates of covariances, and carrying out a REML analysis for each sample. A total of 10,000 analyses were performed, and sampling variances determined as the variances across replicates.

C) Approximate values obtained as covariances across 200,000 samples drawn from a MVN distribution as described above.

For both empirical and MVN samples, 95% confidence intervals were obtained after sorting samples in numerical order as the mid-points between the 2.5% top and bottom samples and the remainder. REML estimation and sampling from the MVN distribution were carried out using WOMBAT.

Results. Estimates of sampling covariances among the distinct elements of the genetic covariance matrix ($\hat{\boldsymbol{\Sigma}}_G$) for case I are contrasted in Figure 1, showing excellent agreement between all three values [■ depicting variances $\mathrm{Var}(\hat{\sigma}_{G\,ij})$ and • covariances $\mathrm{Cov}(\hat{\sigma}_{G\,ij}, \hat{\sigma}_{G\,kl})$]. For case II, the estimate of the smallest genetic eigenvalue was not significantly different from zero, i.e. a full rank estimate of $\boldsymbol{\Sigma}_G$ represented an over-parameterised model. As illustrated in Figure 2, this resulted in an overestimate of $\mathrm{Var}(\hat{\sigma}_{G\,ij})$ obtained from the MVN approximation. The component affected pertained to the variance of the trait considered last in the Cholesky decomposition of $\boldsymbol{\Sigma}_G$, i.e. the overestimate reflected

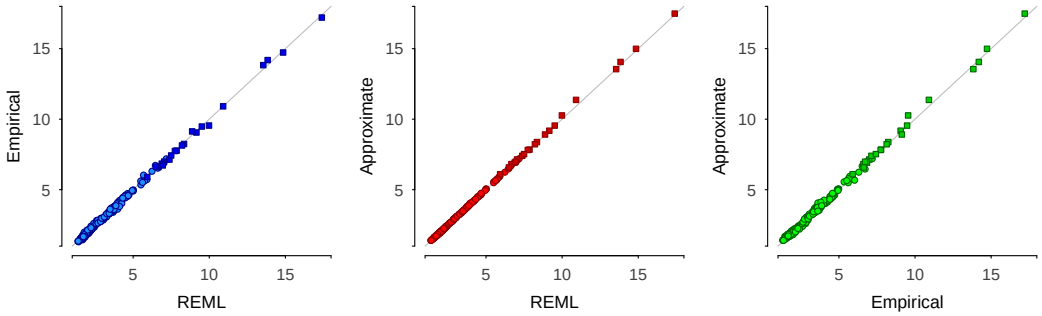


Figure 1. REML, empirical and approximate sampling (co)variances for case I

accumulation of errors for a redundant parameter. Reducing the number of parameters by estimating $\mathbf{\Sigma}_G$ at reduced rank again yielded very good agreement between empirical and approximated values.

Empirical and approximated sampling distributions for selected functions of covariances are compared in Figure 3, with left and right solid vertical bars marking the 95% confidence limits obtained as truncation points between the top and bottom 2.5% of samples and dashed bars showing their 'standard' counterparts, 1.96 standard deviations either side of the mean. Again, there was close agreement between empirical results obtained by re-sampling data and the MVN approximation. For functions at the boundary of the parameter space, such as the genetic correlation between traits 1 and 2, sampling distributions tend to be skewed and confidence intervals derived directly from the distribution tend to be more appropriate than those calculated from sampling errors and normal intercepts. Estimates of genetic eigenvalues are generally reported without any measure of their precision. Similarly, canonical eigenvalues and the number of effective dimensions, $\sum_i \lambda_i/\lambda_1$ (with λ_i the eigenvalues of the matrix of (co)heritabilities and λ_1 the largest value; Kirkpatrick (2009)) are functions of both genetic and phenotypic covariance matrices, and calculation of sampling variances using the Delta method would be, at the least, tedious while it is straightforward using MVN sampling.

DISCUSSION

By definition, REML estimation of covariance components involves the solution of a constrained optimisation problem. Fortunately, this task can be made easier by a transformation to parameters which do not require constraints to yield valid estimates of covariance matrices. Sampling from the asymptotic distribution of these parameters has been shown to yield numerical estimates of sampling covariances, distributions and confidence intervals in close agreement with those obtained by resampling data. It has to be emphasized though that for this to hold, large sample properties need to apply, i.e. the inverse of the information matrix has to provide an adequate description of

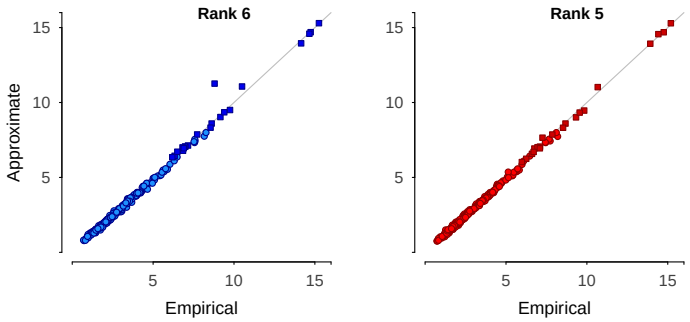


Figure 2. Approximate vs. empirical sampling (co)variances for case II

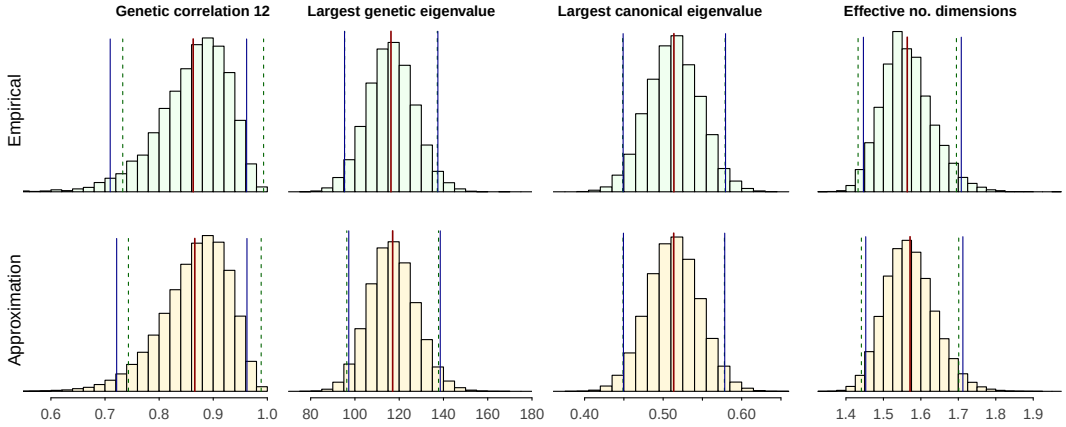


Figure 3. Sampling distributions and confidence intervals for selected functions (Case II)

sampling covariances among the parameters estimated. If this is not the case, estimates of confidence limits derived from the profile likelihood for individual parameters may be more preferable though computationally considerably more demanding (Meyer 2008). In addition, the sampling procedure was found to be sensitive to an overparameterised model, yielding overestimates of sampling variances for redundant parameters, and care needs to be taken for multivariate analyses of more than a few traits to estimate covariance matrices at the appropriate rank.

To facilitate use of the approach described, an option to invoke sampling of parameters from their asymptotic distribution together with the transformation to estimates of covariance matrices has been implemented in our REML package WOMBAT (Meyer 2007) as a post-estimation step. This yields a file with samples of covariance matrices suitable for input to a package such as R (R Core Team 2012) to evaluate the functions of interest and compute summary statistics.

CONCLUSIONS

Sampling of REML estimates from their asymptotic MVN distribution, specified by the inverse of the information matrix, offers a straightforward and computationally undemanding way to derive sampling distributions and confidence intervals for estimates of covariance components and 'non-standard' functions thereof numerically. It is a small but useful addition to our toolkit for estimation.

REFERENCES

Bondari K., Willham R.L. and Freeman A.E. (1978) *J. Anim. Sci.* **47**:358.
Gilmour A.R., Thompson R. and Cullis B.R. (1995) *Biometrics* **51**:1440.
Gilmour A.R., Gogel B.J., Cullis B.R. and Thompson R. (2009) *ASReml User Guide Release 3.0*. VSN International Ltd, Hemel Hempstead, HP1 1ES, UK.
Kirkpatrick M. (2009) *Genetica* **136**:271.
Meyer K. (2007) *J. Zhejiang Uni. SCIENCE B* **8**:815.
Meyer K. (2008) *Heredity* **101**:212.
Meyer K. and Kirkpatrick M. (2005) *Genet. Sel. Evol.* **37**:1.
Meyer K. and Smith S.P. (1996) *Genet. Sel. Evol.* **28**:23.
Oehlert G.W. (1992) *Amer. Stat.* **46**:27.
R Core Team (2012) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

This work was funded by Meat and Livestock Australia under grant B.BFG.0050. Support from the National Institute for Mathematical and Biological Synthesis, University of Tennessee, for working group meetings stimulated the collaboration.

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:527-530

A PRELIMINARY EVALUATION OF A METHOD FOR INCORPORATING GENETIC INFORMATION INTO PHENOTYPIC PREDICTION MODELS

B.J. Walmsley[1], B.P. Kinghorn[2], V.H. Oddy[1], M.J. McPhee[1] and W.A. McKiernan[1,3]

Cooperative Research Centre for Beef Genetic Technologies, Australia. [1]NSW Department of Primary Industries, University of New England, Armidale, NSW 2351. [2]University of New England, Armidale, NSW 2351. [3]71 Fishing Point Road, Rathmines, NSW 2283

SUMMARY

Genetic merit and phenotypic performance are generally predicted independently. Achieving further gains in production efficiency requires genetic and non-genetic fields of expertise work in an integrated manner to deliver new technologies. A study was undertaken using beef cattle to examine a method for incorporating genetic information into phenotypic prediction models. Relationships between fat deposition parameters in a modified version of the Meat Animal Research Centre (MARC) model and Rib Fat EBVs from the Angus cattle breed were explored. These relationships were incorporated into the MARC model and subsequent predictions of P8 fat depth were compared to a scenario where Rib Fat EBVs were not used. Generally small improvements in most measures of predictive accuracy were found. The full commercial benefit of integrating genetic information into phenotypic prediction models will be gained when genetic information is delivered across multi-breed platforms.

INTRODUCTION

The importance of matching animal genotype and management practices with the prevailing production environment has been evident for quite some time. However, the prediction of genetic merit (breeding value) and phenotypic performance still occur relatively independently. Bourdon (1998) suggested predictions of genetic merit are presented without context making it difficult for breeders to effectively use them, particularly in diverse environments. Generally, simulation modelling designed to predict phenotypic performance suffers from the limitation of not giving due consideration to genetic merit. Most simulation models consider genetic merit is described by either breed or breedtype. Both represent impediments to increasing production efficiency. Recognising this Oddy (2009) argued genetics and nutrition need to be more integrated in their approach to delivering new technologies for use in animal production. Bourdon (1998) coined the term 'physiological breeding value' to describe genetic values for inclusion in simulation modelling. Kinghorn (2012) has described various ways genetic information can be used to assist livestock management beyond breeding decisions.

The purpose of this paper is to describe a method for incorporating genetic information into phenotypic prediction models. This will increase the utility of such tools for on-farm decision making and increase the acceptance of phenotypic and genetic prediction in animal production.

MATERIALS AND METHODS

Data. Data used in this study originated from the New South Wales Department of Primary Industries muscling selection herd. This herd is described further elsewhere (McKiernan and Robards 1996, 1997; Walmsley and McKiernan 2011). Briefly, this herd originates from a group of Hereford cows mated to Angus bulls selected from industry herds in 1991 (McKiernan and Robards 1996, 1997) based on visual muscle score (McKiernan 1990). Heifer selection in subsequent generations was based on visual muscle score and single-sire matings were made to Angus bulls selected from industry for high or low muscling to increase/maximise divergence in muscle score between lines.

All progeny have had regular live weights, scanned fatness (P8, rib and IMF) and eye muscle areas recorded by a BREEDPLAN accredited scanner using real-time ultrasound following BREEDPLAN protocols (Graser *et al.* 2005). All pedigree information and recorded data have been submitted to the Angus group BREEDPLAN database and been used to estimate breeding values using the national genetic evaluation system, BREEDPLAN (Graser *et al.* 2005). Data were taken from steer cohorts born in 2006 and 2007 to evaluate the inclusion of EBVs in phenotypic prediction models. The 2006 cohort were used to develop and evaluate relationships between model parameters and EBVs with the 2007 cohort only being used to evaluate these relationships.

Predicting Phenotypes. A dynamic growth modelling system, called 'BeefSpecs', is used to predict animal phenotypes; specifically P8 fat depth (mm). A description of the development of BeefSpecs can be found in Walmsley *et al.* (2010a; 2011). BeefSpecs uses easy to obtain on-farm inputs (e.g. breed type, sex, frame score, live weight, P8 fat depth) in combination with differential equations to describe the pattern of lean and fatty tissue deposition in the empty body of animals (MARC model; Willams and Jenkins 1998). Total body fat is used to predict rib fat depth (mm) and in turn P8 fat depth (mm) (Walmsley *et al.* 2010b).

Integrating Genetic Information. The inclusion of genetic information in phenotypic prediction models was explored using the method outlined by Doeschl-Wilson *et al.* (2007) and Kinghorn (2012). In brief, this involves using differential evolution (DE; Price and Storn 1997) to manipulate parameters in the modified MARC model to achieve the best agreement between observed and predicted P8 fat depths. The first parameter modulates the impact weight gain has on body composition (θ; Keele *et al.* 1992) and the other term describes the relationship between total body fat and rib fat depth (ω; Williams *et al.* 1992). Regressions of estimated growth model parameters on Rib Fat EBV were then developed for inclusion in the modified MARC model.

Statistical Analysis. Sums of squares of difference are used by DE to compare observed and predicted P8 fat depths. Regressions between Rib Fat EBVs and model parameters were developed using the linear model function in the R statistical package (R Development Core Team 2011). Model predictions were evaluated using a customised procedure in R that included mean bias, [∑(observed – predicted)/n], mean square error of prediction (MSEP) and the decomposition of MSEP into bias, slope and random components as proportions (Tedeschi 2006), as well as the regression slope and correlation between observed and predicted values. The root of MSEP (RMSEP) is used to present the prediction error on the same scale as fat depth.

RESULTS AND DISCUSSION

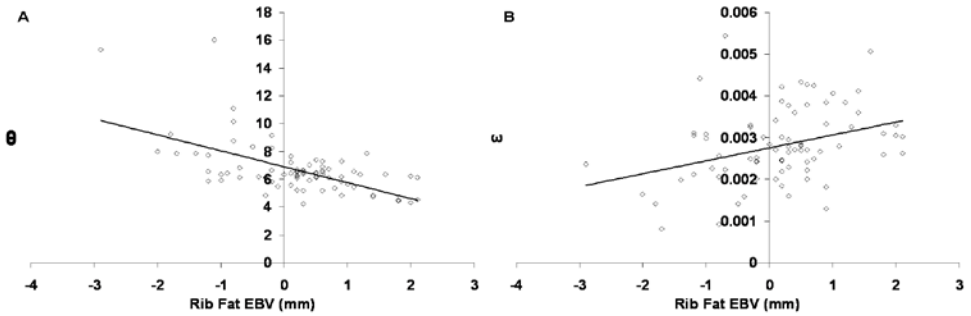


Figure 1. Rib Fat EBV relationship with fat deposition parameters (θ and ω) in the modified Meat Animal Research Centre model for the 2006 steer cohort from the New South Wales Department of Primary Industries muscling herd.

Fat deposition parameters (θ and ω) that produced the highest agreement with observed P8 fat depths in the 2006 steer cohort were determined using DE. The relationships Rib Fat EBVs have with these parameters are demonstrated in Figure 1. The regression of θ on Rib Fat EBV was: θ = 6.90 (s.e. = 0.18) – 1.15*EBV (s.e. = 0.18) (R^2 = 0.34). The regression of ω on Rib Fat EBV was: ω = $2.75e^{-3}$ (s.e. = $9.55e^{-5}$) + $3.10e^{-4}$*EBV (s.e. = $9.64e^{-5}$) (R^2 = 0.12). Regressions containing both Rib and Rump Fat EBVs were tested and resulted in similar accuracy.

Table 1. Assessment of differences between observed and predicted P8 fat depths when not using (base) and using (EBV) Rib Fat EBVs to assist predictions in the 2006 and 2007 steer cohorts from the New South Wales Department of Primary Industries muscling herd.

	2006 Cohort		2007 Cohort	
Descriptor	Base	EBV	Base	EBV
n	80		78	
Observed (O) P8 fat, mm	9.68		3.68	
Predicted (P) P8 fat, mm	9.61	9.69	4.09	4.06
Mean Bias, mm	0.06	-0.01	-0.41	-0.38
Slope of O vs. P, b	0.78	0.71	0.56	0.63
Correlation between O and P, r	0.65	0.87	0.37	0.45
RMSEP, mm	1.72	1.35	1.48	1.41
Bias, %	0.13	0.01	7.56	7.35
Slope, %	5.77	32.65	8.47	7.13
Random, %	94.10	67.34	83.97	85.51

The predictive accuracy of including Rib Fat EBVs in the MARC model is demonstrated in Table 1 in comparison to base scenarios where EBVs were not used. In the 2006 cohort there was a slight improvement in mean bias, correlation between observed and predicted, RMSEP and MSEP due to bias. However, the base scenario had a slope between observed and predicted that was closer to 1 and a lower proportion of MSEP was due to errors in the slope component. All measures of predictive accuracy improved slightly in the 2007 cohort compared to the base scenario when Rib Fat EBVs were incorporated to make predictions of P8 fat depth.

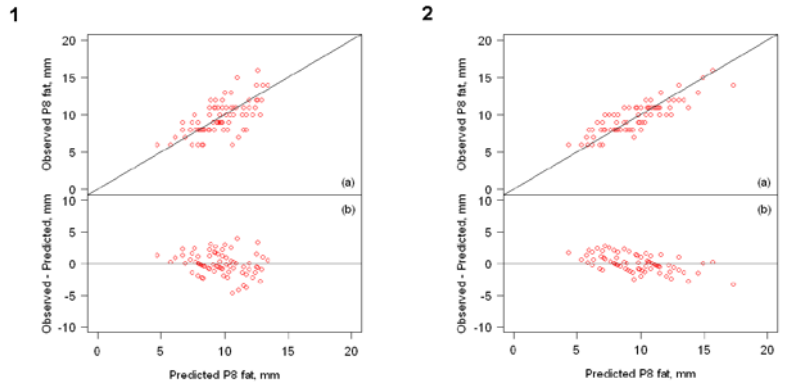


Figure 2. Plot of (a) observed vs. predicted and (b) residual P8 fat depth when (1) not using and (2) using Rib Fat EBVs to assist predictions of P8 fat depth in the 2006 steer cohort from the New South Wales Department of Primary Industries muscling herd.

Differences between observed and predicted P8 fat depth in the 2006 cohort are shown in Figure 2. The improvements in mean bias and RMSEP due to the inclusion of Rib Fat EBVs in predictions of P8 fat depth seen in Table 1 are evident in Figure 2. The slight change in the slope between observed and predicted P8 fat depth seen in Table 1 is also recognisable. These results suggest there is scope for EBVs to be incorporated into phenotypic prediction models. However, some issues have arisen during this process. Some θ values obtained from the linear regressions between θ and Rib Fat EBV are beyond the parameter range considered realistic and thus these regressions need further investigation. Another important issue hindering the inclusion of EBVs in phenotypic prediction models is EBVs are currently derived on a breed specific basis. The modified MARC model functions across breeds by specification of breed type (British, European, *Bos Indicus*). Generation of EBVs applicable across breeds would compliment this and simplify their inclusion in phenotypic prediction models not only in beef cattle but other livestock species.

CONCLUSIONS

The inclusion of Rib Fat EBVs in a modified MARC model slightly increased the accuracy of predicting P8 fat depth. This result indicates that EBV inclusion in phenotypic prediction models should be further explored over a wider range of EBVs. Some consideration should however be given to which model parameters are involved and their biological interpretation as well as the method of EBV delivery to industry (i.e. multibreed EBVs in preference to breed-specific EBVs).

ACKNOWLEDGMENTS

The many people involved in data collection, animal management and other duties associated with the New South Wales Department of Primary Industries muscling herd are acknowledged. The authors gratefully acknowledge the funding support from the Cooperative Research Centre for Beef Genetic Technologies during this study.

REFERENCES

Bourdon R.M. (1998) *J. Anim. Sci.* **76**: 2308
Doeschl-Wilson A.B., Knap P.W., Kinghorn B.P. and van der Steen H.A.M. (2007) *Anim.* **1**: 489
Graser H-U., Tier, B., Johnston D.J., and Barwick S.A. (2005) *Aust. J. Exp. Agric.* **45**: 913
Keele J.W., Williams C.B. and Bennett G.L. (1992) *J. Anim. Sci.* **70**: 841
Kinghorn B.P. (2012) *Anim. Prod. Sci.* **52**: 78
McKiernan W.A. (1990) *Proc. Ass. Adv. Anim. Breed. Genet.* **9**: 447
McKiernan W.A. and Robards G.E. (1996) *Proc. Aust. Soc. Anim. Prod.* **21**: 151
McKiernan W.A. and Robards G.E. (1997) *Proc. Ass. Adv. Anim. Breed. Genet.* **12**: 77
Oddy V.H. (2009) *Rec. Adv. Anim. Nut.* **17**: 103
Price K. and Storn R. (1997) *Dr Dobb's J.* **264**: 18
R Development Core Team (2011) R Foundation for Statistical Computing, Vienna, Austria. URL: http://www.R-project.org
Tedeschi L.O. (2006) *Ag. Sys.* **89**: 225
Walmsley B.J. and McKiernan W.A. (2011) *Proc. Ass. Adv. Anim. Breed. Genet.* **19**: 123
Walmsley B.J., Oddy V.H., McPhee M.J., Mayer D.G. and McKiernan W.A. (2010a) *Aust. Farm Bus. Man. J.* **7**: 29
Walmsley B.J., Wolcott M.L. and McPhee M.J. (2010b) *J. Anim. Sci.* **88**: 1848
Walmsley B.J., Oddy V.H., McPhee M.J., Mayer D.G. and McKiernan W.A. (2011) *ODSIM2011, 19th International Congress on Modelling and Simulation* **19**: 898
Williams C.B. and Jenkins T.G. (1998) *J. Anim. Sci.* **76**: 980
Williams C.B., Keele J.W., and Bennett G.L. (1992) *J. Anim. Sci.* **70**: 858

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:531-533

IMPROVED REPORTING METHOD FOR GENETIC CONNECTEDNESS

B. Visser[1], T.J. Byrne[1], M.J. Young[2] and P.R. Amer[1]

[1]AbacusBio Limited, PO Box 5585, Dunedin
[2]Beef + Lamb New Zealand Ltd., PO Box 39-085, Harewood, Christchurch

SUMMARY

In order to compare the genetic merit of sheep from different flocks in different environments, the genetic connectedness (or linkage) between those flocks has to be adequate. Connectedness between flocks can be maintained or improved by using a common sire that has progeny with performance records in both flocks. A dendrogram was used previously by Sheep Improvement Ltd, NZ (SIL) as a tool to visualise across-flock connectedness. Due to its technical appearance, this reporting method has been found to be poorly understood. This has restricted breeder commitment to building and maintaining connectedness between flocks.

An improved and simplified reporting method for across-flock connectedness was developed (derived from dendrogram results), where a connectedness value is represented by a number from 3 to 0, strong to weak respectively. Traffic lighting — colour coding these numbers green, amber or red — is used to instantly show flocks where connectedness is weakening or is insufficient. In addition to traffic lighting, a smiley face (😐 or ☹) indicates a trend, i.e. when flocks will 'disconnect' in the future. The smiley face functions as an early warning signal to the flock manager, so action can be taken to maintain or add to that flock's connectedness in the future.

A new tool comprising the improved reporting method has been implemented by SIL. The new output format simplifies presentation of across-flock connectedness to ram breeders, and is designed to increase understanding and awareness, giving breeders more incentive to maintain or increase genetic connectedness of their flocks.

INTRODUCTION

In order to validly compare the genetic merit of animals from different flocks, genetic connectedness between these flocks has to be sufficient. Currently across-flock connectedness is already calculated, however, it has been found that numeric and graphical results from the analysis require technical knowledge to interpret and are therefore difficult to understand for many users. Specifically if multiple traits per flock are considered, it is not straightforward to get an overview of the situation. As a result, many breeders have not paid enough attention to maintaining good across-flock connectedness.

An improved reporting tool for across-flock connectedness was developed for and implemented by SIL, where a connectedness value is quantified and depicted visually by "traffic lights". This method is expected to greatly increase understanding of the across-flock connectedness analysis results.

Alongside the traffic lighting, another statistic is depicted that warns when a flock may lose connectedness in the following years. This assists bureaus and breeders in determining flocks and the traits they measure that need addressing to maintain or enhance connectedness.

This report describes the new method for presenting the results from an across-flock connectedness analysis, along with some improvements to the existing calculations of connectedness.

METHODS

The new across-flock connectedness tool is based on SAS code already in use by SIL. The following improvements to the earlier implementation were made:

1. Connectedness values are rescaled to correspond with their actual meaning (e.g. a high value corresponds to a strong connectedness);
2. Across-flock connectedness per flock and per trait is collected in a table, allowing for a compact overview;
3. Values in the table are coded to a colour (red, amber, or green), to indicate the quality of connectedness they represent (weak, adequate, or strong respectively), directing attention to flocks that need it;
4. When a flock's connectedness is adequate, but mostly from sires used in earlier years rather than recent years, it is marked with a warning sign, to flag when the flock will lose connectedness if no action is taken.

Calculating connectedness. A SAS program calculates parameters for across-flock connectedness for a trait by counting sire offspring in different flocks in the most recent years containing relevant data. Then the relative distances between flocks with respect to number of offspring from common sires is determined and the flocks are clustered according to the nearest neighbour algorithm. Connectedness per trait is defined as the distance between those clusters (Young and Newman 2009).

In the old method, all distances were normalized to the mean distance per trait by the SAS clustering procedure. This meant that the connectedness values were specific to a single trait analysis, making comparisons between traits unfair. The new implementation removes cluster normalization and this drawback.

A main cluster was defined as the cluster that held the highest number of strongly and adequately connected flocks. All other clusters then have to be connected to the main cluster in order to retain their connectedness value. This meant that although two flocks might be strongly connected to each other, unless they were connected to the main cluster, their original connectedness value would be discarded.

Rescaling. For the old method, a connectedness value was defined within a range from 0.0 to 0.5, where a value closer to 0 meant that a flock's connectedness is stronger. It has been found that this was counter-intuitive for most users, because a *high* connectedness value meant that a flock's connectedness was *weak*, and vice versa. Therefore, after calculating the connectedness values, they were rescaled so that high values meant stronger connectedness. This aims to clarify across-flock connectedness for users.

Traffic lighting. Colour coding can be used to show a reader if a result is good or bad without the need to assess the relative size of values. This practice is referred to as "traffic lighting".

For the across-flock connectedness summary, traffic lighting was applied to connectedness values. Numbers were tabulated and depicted as coloured symbols. This meant that for a trait, flocks have either a green, amber or red value, quickly highlighting strong, adequate and weak connectedness respectively.

It is worth noting that a value and colour essentially show the same information, i.e. how well the flock is connected to other flocks for a specific trait. Traffic lighting was added to further simplify the meaning of a value in the table. However, when the connectedness table is printed in black and white, it is not possible to distinguish between the colours, so numbers are the primary indicators in this situation.

Warning sign. Another feature of the new connectedness tool is the ability to flag the connectedness trend, an early warning for flocks that would lose connectedness in the later years if no action was undertaken by the breeder.

Because connectedness is calculated for a relatively short range of years, the contribution of a single year is large and can possibly skew the connectedness value. This becomes a problem when the oldest year in the range has a lot of strongly connected flocks, while the newer years in the range show a weaker connectedness per year. Overall the connectedness over a range of years

might still be adequate or even strong, but when connectedness is calculated in the future, say a year later, the oldest data is no longer included. This causes a steep (and unexpected) drop in the across-flock connectedness value.

The warning symbol in the form of a smiley face, was added to the connectedness table for flocks in danger of losing connectedness in the subsequent years. It is issued when the connectedness of the most recent years of the window are weak.

Table 1 Connectedness values are aggregated by flock and trait, providing a compact overview of the situation. Numbers in a range from 0 to 3 represent the level of current connectedness from weak to strong. Colours correspond to connectedness levels and (when colour printing is available) draw attention to flocks that need it. Smiley faces are used to warn of flocks that will have a weak connectedness in the next (☹), or year after next (😐) respectively.

Flock	Growth	Wool	Reproduction
A	3 *(g)*	3 *(g)*	2 *(y)*
B	3 *(g)*	1 *(r)*	3 😐 *(g)*
C	1 ☹ *(r)*	2 ☹ *(y)*	0 *(r)*
D	2 😐 *(y)*	3 *(g)*	2 ☹ *(y)*

(g) green, *(y)* yellow, *(r)* red – for interpretation here with grey scale printing

CONCLUSION

SIL's connectedness tool has been updated to include a simplified connectedness summary. Previously across flock connectedness was calculated and presented as a single dendrogram (graph) per trait; a fairly technical method of displaying results. This meant that several graphs had to be consulted for an overview of the connectedness across multiple traits. In addition, due to the method used to calculate connectedness, *low* values corresponded to *strong* connectedness and vice versa. This was thought to be counter-intuitive for most users.

Firstly connectedness values were aggregated for all traits in the analysis and presented in a single table, providing a compact overview. Secondly, connectedness values were rescaled so they correspond to the level of connectedness: high values for strong connectedness, and low values for weak connectedness. Lastly, traffic lighting (assigning colours to numerical values) is used to help interpret the numbers in the table, without the need to understand the underlying scale.

To give users an early warning, the connectedness trend is calculated separately and depicted by smiley faces. It shows whether connectedness will be weak in the next (☹), or year after next (😐) respectively.

It is expected that the new flock connectedness tool will increase awareness of connectedness and proactive use of link sires by ram breeders. As a consequence across-flock connectedness should strengthen, increasing the accuracy of across flock genetic evaluations.

ACKNOWLEDGEMENTS

The authors would like to thank Ovita Ltd for initial funding for this work, Sheep Improvement Ltd for assisting in provision of the data and Sheryl-Anne Newman for helpful input and comments in developing the tool.

REFERENCES

Young M.J. and Newman S-A.N. (2009) *Proceedings of the New Zealand Society for Animal Production* **69**: 153.

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:534-537

A METHOD FOR MAXIMISING AVERAGE FLOCK REPRODUCTION BY OPTIMISING CULLING POLICIES ACROSS AGE GROUPS

J.S. Richards[1,2,3], K.D Atkins[1], B.P. Kinghorn[1] and J.H.J. van der Werf[1,2]

[1]University of New England, Armidale, NSW 2351, Australia.
[2]Australian Cooperative Research Centre for Sheep Industry Innovation, Homestead Building, University of New England, Armidale, NSW 2351, Australia.
[3]NSW DPI, Orange Agricultural Institute, Forest Road, Orange, NSW 2800, Australia.

SUMMARY

Increasing the average reproduction of a sheep flock is beneficial for most commercial sheep producers, and especially those focussing on meat production. Not only does it increase selection intensity for further selection if required, but it also increases the quantity of meat sales, and therefore the value of the flock. Currently some producers cull ewes if they are dry once (or sometimes dry twice). The aim of this study was to optimise culling strategies based on expected lifetime reproductive performance across age groups to achieve a desired higher average flock reproduction. If commercial sheep producers were able to use information readily available to them to improve their average flock production and reproduction it would provide a simple way to increase the value of their flock.

INTRODUCTION

There are many benefits of increasing the average reproduction of a flock of sheep. Benefits such as increased selection intensity (with more lambs to select replacements from), more excess lambs (kg of meat) for sale and the need to support fewer ewes for the same number of lambs produced. These and other benefits can result in both immediate financial advantage as well as continued genetic improvement.

Reproduction is lowly heritable and low-moderately repeatable. This suggests that selection on lifetime performance may be a useful way to increase flock reproduction levels. Lee and Atkins (1996) found that ewes that produced lambs in the first two joinings subsequently reared twice as many as those that didn't rear any lambs in their first two joinings. In an attempt to increase reproduction in flocks some commercial sheep producers cull ewes that are dry once or sometimes dry twice.

There is potentially a better approach that could increase the average reproduction of the flock by recording all the previous lambings of each ewe and using this information to predict later lifetime performance for making management decisions. Optimal culling policies could be based on the historical lambing information for each ewe. This information is easily recorded by lambing rounds where lambs are manually identified and recorded with their dam or by using more recent technology such as Pedigree MatchMaker (Richards and Atkins 2007) which identifies which lambs belong to which ewe and subsequently how many lambs each ewe raised each lambing.

It would be impractical waiting several years to examine various culling policies in trials on farm. Instead, simulation can be used to examine various scenarios to determine the best culling policy. We used a model previously developed for optimising culling strategies across age groups for continuous traits and this model was adapted for non-continuous traits, in this case fertility.

METHOD

Method and code was previously developed for simulating flock changes over age classes for a single continuous trait, such as fibre diameter. This also had the option of optimising culling strategies for increased profitability (or a similar objective) (Richards, unpublished). This was

adapted to handle reproduction, with the additional challenge of handling discrete data rather than continuous data. This original model used a normal distribution with a mean and SD to simulate the base flock group of sheep for a continuous trait. Later years were simulated using a measure of growth as well as correlations between trait expressions in different years. This method would not by itself work for a discrete trait, such as fertility.

To adapt this model to be used for discrete traits, a liability distribution was created for predicting realised reproduction thereby allowing it to be based on the continuous model approach. Thresholds were used to determine the number of lambs for ewes within the flock given their liability values. Figure 1 shows a normal distribution of liability scores and thresholds for singles and twins. When culling a proportion of the flock with continuous data a truncation point along the x-axis can be used to split the animals into keep and cull groups. With liability scores for predicting discrete traits it is more complicated. When knowing the proportion of animals required the best animals can be determined according to liability scores. However, within a lambing subclass, we cannot observe liability as all have the same discrete phenotypic value. For example, in Fig. 1 everything from 'a' to 'b' is only observed to be single. The proportion of singles needed will then be taken equally from all liability values within the subclass; shown by the shaded (selected) section of singles. All of the twins are selected in this case. This approach of using a liability score allows discrete traits to be modelled via an underlying continuous trait.

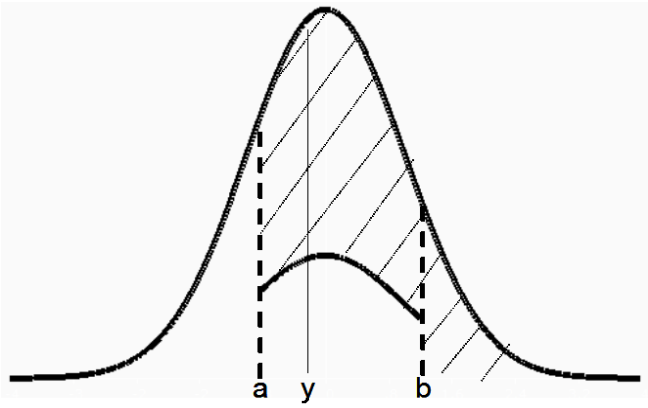


Figure 1: Liability distribution showing thresholds for realised dry (<a), single (a-b) and twin (>b) ewes. A cutoff at point y liability score would result in shaded animals selected based on observed reproductive phenotype

The distributions of subsequent years are then simulated based on the selection in the previous year. The age class means are adjusted by age adjustment factors (Table 1) and correlations between performance in subsequent years are presented in Table 2. The distribution is divided into 100 segments, with the animals in each segment having a similar liability score. The mean value of a given segment in the next (i^{th}) year will be predicted from the regression of the current mean on the mean of the next year as follows:

$$y_i = x_i + \frac{\left(y_{i-1} - x_{i-1}\right).SD_i.r_{i,i-1}}{SD_{i-1}}$$

where *y* is the liability as in Figure 1, *x* is the average flock liability, *r* is the phenotypic correlation between the current year and next year's performance as in Table 2, and *SD* is the standard deviation of y_i for if there was no selection. Values for SD are derived from the mean of each age group using a coefficient of variation of 7%.

The variation around the segment is generated as follows:

$$RSD_i = \sqrt{(1-r^2)}.CV_{i.}.y_i$$

where *RSD* is the residual *SD* and *CV* is coefficient of variation. The full distribution of liability for a given year is found by summing these distributions across segments. These liability scores are then used for selection as described above for the first year (Figure 1). For fixed thresholds, this process is equivalent for "across age groups within year" and "for the same cohort across years", but the latter has been used for description here.

Optimisation seeks to set the best set of truncation points (y in Fig 1.) for each year, with the objective of maximising predicted overall flock reproductive performance.

The age adjustments and the correlations used in the model are shown in Tables 1 and 2.

Table 1: Age adjustments for reproduction (fertility) of animals (age 1 to 6 yrs) used in simulation model Source: (Turner and Dolling, 1965)

	Age 2	Age 3	Age 4	Age 5	Age 6
Reproduction adjustment	0.838	0.940	0.990	1.072	1.108

Table 2: Phenotypic correlations for litter size between subsequent ages used in simulation model Source: (Atkins, 1990)

	Age 2	Age 3	Age 4	Age 5	Age 6
Correlation to previous age		0.14	0.15	0.12	0.17

Adjustment of the thresholds (points a and b in Figure 1) can be made to reflect conditions (eg. genetics, nutrition, stocking rate) that affect reproduction, and change the proportion of each class expected within a flock or age class in a particular year.

In order to optimise the culling policies an objective needs to be set. We assumed that profit was determined by a fixed price per lamb produced. In further development it would be useful to also account for a lamb trait such as weaning weight or carcass value, and maximise overall profit, accommodating the correlation between fertility and the lamb trait.

When optimising culling strategies for increased value across multiple age groups additional considerations need to be made, such as the benefit of younger animals having higher genetic merit due to genetic trend, a survival rate per age group that decreases for the older animals and a different accuracy between younger and older ewes to predict lifetime performance because young animals have less information available whereas the older ewes have proven reproduction rates. By setting the objective to achieve the highest dollar value the optimal culling policy will be determined (as this would be highest number of lambs).

A software package was developed based on the model described, including an evolutionary algorithm to optimise culling decisions for the prevailing parameters. It enables scenarios such as

using one, two, and three years of fertility records for selection to be compared by recording the total dollar value as well as average lifetime reproduction of the flock. This can be used to compare the value of optimising across age groups against the more traditional methods (by setting the selection cutoff manually to the same as the single threshold), so all animals that were dry in one, then one and two, and lastly one to three age groups were culled.

DISCUSSION

The model does not only allow a comparison of various culling strategies but it also can be used to identify whether it would be worth the extra labour and time to record fertility records of ewes for at least 1, 2 or 3 joinings and using this information to better select the animals to retain. The resulting increase in flock reproduction would have added benefits to the flock than just the extra meat value, so the benefits would actually be greater than the dollar value shown from the output. If there is a large increase in value from using this approach then it suggests this process would be a suitable alternative for increasing average flock reproduction with little extra labour and time (if using a low labour intensive method for collecting these records such as Pedigree MatchMaker or DNA tests for pedigree) or alternative manually recording information on daily lambing rounds. If it is a similar value to the traditional approach then it may not be worth the extra effort required to achieve it.

The addition of other traits to the model would be desirable as it would make the results more realistic. Once additional traits are added for examining the results of selecting on one trait, it would be very beneficial to then be able to combine one or more traits into the selection process and in calculating overall flock profit. This may be examined in future studies. The ability to combine continuous traits with discrete traits will be very useful and should be possible with more work on the model also.

CONCLUSION

The results of this study show that data on fertility performance of ewes can be used to establish better culling policies for increasing average flock reproduction in commercial sheep flocks. We described a model that can compare different culling policies and determine the value of optimising culling strategies as well as the value of recording performance data in commercial flocks. Liability scores are a useful way of simulating and examining discrete traits and this model could easily be adjusted for other traits, including traits in other species as long as the age effects, correlations to previous year and potentially other traits have been determined.

REFERENCES

Atkins K.D. (1990) *Proceedings of the 4th World Congress on Genetics Applied to Livestock Production* **17**: 26.

Lee G. and Atkins K.D. (1996) *Australian Journal of Experimental Agriculture* 36:123.

Turner, Helen Newton and Dolling C.H.S. (1965) *Australian Journal of Agricultural Research* **16**: 699.

Richards J.S., Atkins K.D. (2007) Proceedings of the Association for the Advancement of Animal Breeding and Genetics **17**: 403.

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:538-541

GENETIC CONTROL OF RESIDUAL VARIANCE FOR TEAT NUMBER IN PIGS

M. Felleki[1,2] and N. Lundeheim[2]

[1]School of Technology and Business Studies, Dalarna University, 79188 Falun, Sweden
[2]Department of Animal Breeding and Genetics, Swedish University of Agricultural Sciences, 75007 Uppsala, Sweden

SUMMARY

The genetic improvement in litter size in pigs has been substantial during the last 10-15 years. The number of teats on the sow must increase as well to meet the needs of the piglets, because each piglet needs access to its own teat. We applied a genetic heterogeneity model on teat number in sows, and estimated medium-high heritability for teat number (0.5), but low heritability for residual variance (0.05), indicating that selection for reduced variance might have very limited effect. A numerically positive correlation (0.8) between additive genetic breeding values for mean and for variance was found, but because of the low heritability for residual variance, the variance will increase very slowly with the mean.

INTRODUCTION

For pigs the genetic improvement in litter size has been substantial during the last 10-15 years. The number of teats on the sow must increase as well to meet the needs of the piglets, because each piglet needs access to its own teat (Chalkias *et al.* 2013).

Genotypes differ not only in mean for a trait but also in variation around the mean (Mulder *et al.* 2007). The possibility to select for uniform individuals by selecting animals expressing a small response on environment has been studied extensively in animal breeding. Considerable support for a heritable component in the environmental variation has been found (Hill and Mulder 2010).

The term genetic heterogeneity is used for models including genetically structured differences in the residual variance. It is difference in residual variance among individuals maintained in similar environments, caused by genetic interaction with unknown environmental differences. Having fitted fixed effects such as herd and sex, the remaining unknown environmental differences among individuals are assumed to be negligible, therefore referred to as micro-environmental changes (Mulder *et al.* 2013, Rönnegård *et al.* 2013). Genetic heterogeneity is not to be confounded with the topic of robustness; reaction on macro environmental differences.

Rönnegård *et al.* (2010) and Felleki *et al.* (2012) proposed an algorithm for estimation of genetic heterogeneity, which builds on the theory of Double Hierarchical Generalized Linear Models (Lee and Nelder 2006). The algorithm has previously been used for analysing data on litter sizes in pigs, and for analysing data on milk yield and somatic cell counts in dairy cattle (Rönnegård *et al.* 2013).

The aim for this paper is to study genetic heterogeneity for teat numbers in pigs, and thereby discuss the feasibility for genetic increase of number of teats in sows, that is, the possibility to select for an increasing stabile number of teats.

MATERIAL AND METHODS

Data were obtained from the Swedish pig breeding organisation Nordic Genetics, and included data on teat number (recorded at three weeks of age, both genders) on 47866 purebred Yorkshire pigs and their pedigree (in total 52817 individuals). The teat number is total teat number including non-functional teats. The pigs were born between January 2007 and April 2009. Variables in the data set were number of teats at three weeks of age, litter identity, year-month of birth, herd, gender, litter size, and birth parity number. Analyses were restricted to nucleus herds with at least

a thousand animals recorded during the time period considered. Teat number observations below ten and above nineteen (totally 25 observations) were removed from the data set.

The mean of teat number was 14.49, and the standard deviation was 0.94. Most pigs (24147) had 14 teats, 12355 had 15 teats, 6708 had 16, 3017 had 13, 825 had 12, 642 had 17 teats, and the rest (totally 172) had 10, 11, 18 or 19 teats. Number of dams was 3403 and number of sires was 337. Dams had between 1 and 67 offspring with median 11, and sires had between 2 and 717 offspring with median 87.

Four models were fitted. For the first model, the Animal model, teat numbers *y* were modelled

$$y = \mu + X\beta + Za + Wpe + e,$$

where μ was an intercept, β was a vector of fixed effects of year-month of birth, herd, gender, and birth parity number, X was a known design matrix, $a \sim N(0, A\sigma_a^2)$ was the random effect of animal, A was the additive genetic relationship matrix, $pe \sim N(0, I\sigma_{pe}^2)$ was the random effect of litter identity, Z and W were known coincidence matrices, and $e \sim N(0, I\sigma_E^2)$ was the residual.

Three models included individually structured genetically differences in the residual variance. Same additive genetic structure, either sire, dam, or sire-dam, was used for mean and variance, and the models were named Sire, Dam, and Sire-dam referring to the common structure of the additive genetic effects s for the mean model and s_d for the variance model. The coincidence matrix Z had a 1 in the column for sire, dam or both, and the mean part was otherwise similar to the animal model, $y = \mu + X\beta + Zs + Wpe + e$.

The residuals were assumed to be heterogeneous, $e \sim N(0, \Phi)$, Φ was a diagonal matrix with diagonal φ, and it was assumed that φ was linear on logarithmic scale, $\log \varphi = \log \sigma_{E,exp}^2 + X\beta_d + Zs_d + Wpe_d$.

It was moreover assumed that s and s_d were correlated,

$$\begin{pmatrix} s \\ s_d \end{pmatrix} \sim N(0, \begin{pmatrix} \sigma_s^2 & \rho\sigma_s\sigma_{s_d,exp} \\ \rho\sigma_s\sigma_{s_d,exp} & \sigma_{s_d,exp}^2 \end{pmatrix} \otimes A),$$

while the random effects for litter identity pe and pe_d were assumed independent, $pe_d \sim N(0, I\sigma_{pe_d,exp}^2)$. Fixed effects, β_d, were same as for the mean model.

The genetic heterogeneity models were fitted using the algorithm from Felleki *et al.* (2012). The statistical principle used is that of extended likelihood, or hierarchical likelihood. The joint likelihood of trait values and random effects is used for estimation of mean effects, and adjusted profile likelihoods are used for estimation of effects for the residual variance, and for estimation of the variance components. The resulting algorithm is feasible for large data sets, and necessary commands are implemented in ASReml 4.0.

Mulder *et al.* (2007) gave formulas for the heritability for residual variance, which is modified to be used for the sire, dam, and sire-dam models with permanent environmental effect,

$$h_v^2 = \frac{4\sigma_{s_d}^2}{2\sigma_P^4 + 3(\sigma_{s_d}^2 + \sigma_{pe_d}^2)},$$

where

$$\sigma_P^2 = \sigma_a^2 + \sigma_{pe}^2 + \sigma_E^2,\ \sigma_E^2 = \sigma_{E,exp}^2 \exp\left(\frac{\sigma_{a_d,exp}^2}{2}\right) \exp\left(\frac{\sigma_{pe_d,exp}^2}{2}\right),$$

$$\sigma_{s_d}^2 + \sigma_{pe_d}^2 = \sigma_{E,exp}^4 \exp\left(2\sigma_{s_d,exp}^2\right) \exp\left(2\sigma_{pe_d,exp}^2\right) - \sigma_E^4,\ \sigma_{s_d}^2 = (\sigma_{s_d}^2 + \sigma_{pe_d}^2) \cdot \left(\frac{\sigma_{s_d,exp}^2}{\sigma_{pe_d,exp}^2 + \sigma_{s_d,exp}^2}\right).$$

The genetic coefficient of correlation was calculated by $\mathrm{GCV_E} = \sqrt{4\sigma_{s_d}^2}/\sigma_E^2$.

RESULTS AND DISCUSSION

Estimated variance components with standard errors for the four models are found in Table 1. The genetic variance component, σ_a^2, for the animal effect for the mean part of the genetic

heterogeneity models is calculated by $\sigma_a^2 = 4\sigma_s^2$, and these values are between 0.33 and 0.38 in agreement with 0.35 for the animal model.

Table 1. Estimate(standard error) of variance components for an animal model, and variance components and correlations for three genetic heterogeneity models with identical genetic structure for mean and variance (sire, dam, or sire-dam)

Model name	σ_{pe}^2	σ_a^{2*}	$\log \sigma_{E,exp}^2$	$\sigma_{pe_d,exp}^2$	$\sigma_{s_d,exp}^2$	ρ
Animal[†]	0.03(0.003)	0.35(0.017)	-0.67(0.008)			
Sire	0.10(0.003)	0.33(0.036)	-0.64(0.064)	0.12(0.008)	0.03(0.005)	0.85(0.047)
Dam	0.09(0.004)	0.38(0.026)	-0.59(0.068)	0.09(0.008)	0.07(0.008)	0.86(0.039)
Sire-dam	0.02(0.003)	0.34(0.018)	-0.55(0.074)	0.07(0.008)	0.04(0.004)	0.81(0.035)

[*] For the three latter models, $\sigma_a^2 = 4\sigma_s^2$.

[†] Residual variance for Animal model is estimated on logarithmic scale, $\sigma_E^2 = \exp(\log \sigma_{E,exp}^2)$, where $\log \sigma_{E,exp}^2$ is estimated.

The correlation ρ is positive and numerically high (0.81-0.86) indicating a close connectedness between breeding values for mean and for residual variance, hence a Poisson model for teat data might be more appropriate.

Phenotypic variance, σ_P^2, is 0.89 for the animal model, and between 0.72 and 0.78 for the genetic heterogeneity models (Table 2). The difference in values among models might be due to the assumption that the random effects are independent, or the lower values for the genetic heterogeneity models might be caused by the fixed effects in the variance part.

Heritability, h^2, for teat number is found in the medium-high range between 0.39 and 0.48, as previously reported (Chalkias *et al.* 2013). For the animal model, h^2 is 0.39, while h^2 is slightly higher for the genetic heterogeneity models (0.45-0.48).

The heritability for residual variance, h_d^2, takes the values 0.03, 0.05, and 0.07. These are in the higher range of common reported values (Hill and Mulder 2010). As heritability values, however, these values are negligible. The closely connected genetic coefficients of variation GCV_E, with values between 0.34 and 0.59, are also found in the higher range of common values.

Table 2. Heritability and genetic coefficient of variation

Model name	σ_P^2	h^2	$\sigma_{a_d}^2$	h_d^2	GCV_E
Animal	0.89(0.009)	0.39(0.016)			
Sire[‡]	0.75	0.45	0.04	0.03	0.34
Dam[‡]	0.78	0.48	0.10	0.07	0.54
Sire-dam[‡]	0.72	0.48	0.06	0.05	0.39

[‡] Standard errors could not be found.

Inferences under the genetically structured heterogeneous variance model can be misleading when the data are skewed (Yang *et al.* 2011). Therefore data should be checked for scale effects before fitting a genetic heterogeneity model, which has not been done in this study.

Functionality (not inverted, blind, small or inserted) of the teats is a necessity. In this study the genetic components for mean and for variance of total number of teats are estimated, leaving out correlation between functional, non-functional, and total teat number. The data for this study is collected at three weeks of age; hence the counts of non-functional teats might not be accurate.

Chalkias *et al.* (2013) found a favourable (that is positive) correlation between number of functional and total number of teats, and concluded that the genetic increase of teats, will give

increase in functional teats as well. They did, however, mention the possible consequence of a non-functional teat for a piglet using the crucial first hours of life suckling it. We suppose that also the sow could be stressed of this with consequences for nursing behaviour. 13% of all tested pigs had at least one non-functional teat.

Traits important for pig production are many: litter size and uniformity, piglet survival, weight and growth, milk production, teat number, ability to become pregnant, and behaviour (Rydhmer 2000). Many of these traits are genetically connected such that selection on one, as practised on teat number, might give undesired results for other traits (Chalkias *et al.* 2013). These correlations are to be studied.

The heritability for the residual variance, and the correlation between breeding values for mean and variances, can be tools to determine if a trait can be controlled under selection, or if fluctuation of the trait values will increase. In this study we find a considerable correlation, thus variances are expected to increase with increased mean values, and we also find a low value for heritability of residual variance, indicating that selection for reduced variance might have very limited effect. Hill and Mulder (2010) reported that no convincing results have been reported this far on selection for reduced variance in any study. It would be interesting to repeat such an experiment on a trait with a numerically small mean-variance correlation (close to zero) and high variance heritability, if such a trait is found.

CONCLUSION

For teat number in pigs, we find breeding values for mean and variance to be highly correlated indicating a Poisson distribution. Hence selecting for an increased mean number, the variance might increase as well. We also find heritability of breeding values for residual variance to be low; hence selection for decreased residual variance might give negligible response.

As long as the new teats are mainly functional, one way to go around the problem is selection of sows with many functional teats for production as already practised. The low heritability for residual variance indicates that the variance will increase very slowly with the mean. However the piglet's and sow's reactions on non-functional teats are to be investigated.

ACKNOWLEDGEMENTS

The authors are grateful to Lars Rönnegård, Helena Chalkias and Han Mulder for discussions on the study. Data was provided by the breeding company Nordic Genetics.

REFERENCES

Chalkias H., Rydhmer L. and Lundeheim N. (2013) *Livest. Sci.* **152**:2
Felleki M., Lee D., Lee, Y., Gilmour A.R. and Rönnegård L. (2012) *Genet. Res.* **94**:6
Hill W.G. and Mulder H.A. (2010) *Genet. Res.* **92**:5
Lee Y. and Nelder J.A. (2006) *J. R. Stat. Soc.* **55**:2
Mulder H.A., Bijma P. and Hill W.G. (2007) *Genetics* **175**:4
Mulder H.A., Rönnegård L., Fikse W.F., Veerkamp R.F. and Strandberg E. (2013) *Genet. Select Evol*. **45**:23
Rydhmer L. (2000) *Livest. Prod. Sci.* **66**:1
Rönnegård L., Felleki M., Fikse W.F., Mulder H.A. and Strandberg E. (2010) *Genet. Select Evol*. **42**:8
Rönnegård L., Felleki M., Fikse W.F., Mulder H.A. and Strandberg E. (2013) *J. Dairy Sci.* **96**:4
Yang Y., Christensen O. and Sorensen D.A. (2011) *Genet. Res.* **93**:1

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:542-545

ACROSS- AND WITHIN-BREED IMPUTATION ACROSS SEVERAL GENOTYPING DENSITIES IN DAIRY AND BEEF CATTLE

D.P. Berry[1], M.P. Mullen[2] and A.R. Cromie[3]

[1]Animal & Grassland Research and Innovation Centre, Teagasc, Moorepark, Fermoy, Co. Cork, Ireland. [2]Animal & Grassland Research and Innovation Centre, Teagasc, Athenry, Co. Galway, Ireland. [3]Irish Cattle Breeding Federation, Bandon, Co. Cork, Ireland.

SUMMARY

Illumina high density genotypes (777,962 SNPs) were available on 3,122 dairy and beef bulls. Animals were partitioned into either a calibration or validation dataset to test the accuracy of imputation. All animals, irrespective of breed, born after 2005 (n=698) were assumed to represent the validation bulls. The high density genotypes were masked in the validation animals to represent a low density (n=6,501) or medium density (n=47,770) genotyping platform. The accuracy of within breed imputation (i.e., correlation between actual and imputed genotype) from medium density to high density (0.987) was superior to that from low-density to high density (0.938) with the difference diminishing as the proportion of back-pedigree genotyped on the high-density platform increased. Using multiple breeds in the calibration dataset for imputation did not improve the accuracy of imputation.

INTRODUCTION

Genomic selection (Meuwissen *et al.* 2001) exploiting genome wide information on a large population of animals is the method of genetic evaluations in many dairy (Hayes *et al.* 2009) and some beef (Saatchi *et al.* 2012) populations. The accuracy of the genomic predictions is a function of the size of the population of animals with both phenotypes and genotypes. Greater prediction accuracy is achievable with larger reference populations (Daetwyler *et al.* 2008). There is nonetheless a cost to genotyping large populations of animals especially for higher density genotypes. This cost could be reduced by genotyping using a lower density (i.e., lower cost) genotype panel and imputing to a higher density. Imputation still requires a population of animals genotyped on the higher density genotype panel but has been shown to be accurate within dairy (Weigel *et al.* 2009; Berry and Kearney 2011) and beef cattle (Dassonneville *et al.* 2012; Huang *et al.* 2012). These studies have primarily only imputed from low to medium density genotype panels although studies on imputation to high density genotype panels also exist (Erbe *et al.* 2012; VanRaden *et al.* 2013). The cost of acquiring higher density genotypes could potentially be further reduced if the reference population of animals genotyped on the higher density could be generated from multiple breeds. Nevertheless, there is little information on the usefulness of across-breed imputation in cattle (Brøndum *et al.* 2012), especially genetically diverse breeds like beef and dairy breeds.

The objective of the present study therefore was to evaluate the accuracy of imputation from lower density genotyping panels to higher density genotyping panels in dairy and beef cattle using a single-breed reference population or multi-breed reference population.

MATERIALS AND METHODS

Genotype data. Illumina high-density (HD) genotypes (777,962 single nucleotide polymorphisms; SNP) were available on 3,122 dairy and beef bulls. The number of bulls per breed was 269, 196, 710, 234, 719, 730, and 264 for Angus, Belgian Blue, Charolais, Hereford, Holstein-Friesian, Limousin and Simmental, respectively. Mendelian inconsistencies were used to validate animal identification but also to discard autosomal SNPs that did not adhere to Mendelian

inheritance. Only autosomal SNPs with a UMD 3.1 genomic location were retained.

Two alternative SNP density panels were generated to represent the Illumina Bovine50 beadchip (50K) and Illumina Low Density (LD) genotyping panel. A total of 47,770 of the autosomal SNPs on both the HD panel and 50K genotyping were retained. Additionally 6501 autosomal SNPs on both the HD and LD panels were retained.

Imputation. Animals were partitioned into either a reference or a validation population to test the accuracy of imputation. All animals, irrespective of breed, born after 2005 (n=698) were assumed to represent the validation bulls; all other bulls were included in the reference population. Imputation was from lower to higher density genotypes. In all analyses the full complement of higher density genotypes were retained in the reference animals. Genotypes were masked in the validation animals to represent the lower density panels. Imputation to the higher density genotypes was undertaken for each chromosome separately using the freely available software Beagle Version 3.1.0 (Browning and Browning 2007; 2009). Imputation was undertaken within and across breeds. In all analyses the same animals were included in the validation population. However, when the analysis was within breed, only the animals of that breed were included in the reference population. The accuracy of imputation was determined based on the correlation between the actual and imputed genotypes. In all instances, the accuracy of imputation was calculated by including in the arithmetic the non-masked genotypes. This was to generate results that are therefore applicable in the real life situation; most studies only report the accuracy of imputation for the masked genotyped so therefore results in the present study are likely to be slightly better.

RESULTS AND DISCUSSION

Summary statistics for the accuracy of within-breed and across-breed imputation across the different genotyping platforms are in Table 1. Mean accuracy of imputation per chromosome was similar although variation in imputation accuracy did exist across the genome and the genomic locations of the reduced accuracy were comparable with documented elsewhere (Erbe *et al.* 2012). Erbe *et al.* (2012) reported that 1,231 of the HD SNPs in their population had a genotype concordance rate of <0.80 while the equivalent statistic in the present study when evaluating the accuracy of across-breed imputation from 50K to HD was 2,234 SNPs.

The accuracy of imputation was, on average, greatest when imputing from 50K to HD and was poorest when imputing from LD to HD (Table 1). Minor allele frequency of the different genotype platforms may affect the accuracy of imputation. The minor allele frequency for the LD, 50K and HD genotype panel across all animals in the present study was 0.39, 0.24 and 0.25, respectively. On an individual animal basis, the mean accuracy of imputation from 50K to HD was always superior to the mean individual accuracy of imputation from LD to either 50k or HD. The same conclusion was evident irrespective of whether the imputation was undertaken within or across breed.

Mean imputation accuracy per breed was always superior when undertaken within-breed compared to undertaken across-breed with the exception of the 50K to HD imputation scenario when undertaken in Angus and Belgian Blue cattle although the difference was minuscule.

Despite the differences in reference population sizes of the breeds, there were no obvious breed differences in mean imputation accuracy across genotype platforms when imputation was undertaken within or across breeds; the reference population size of the Holstein-Friesian population was 688 compared to 140 for Belgian Blues.

Irrespective of whether the imputation was undertaken within breed or across breed, the proportion of correctly imputed homozygous genotypes was always poorest when imputing from LD to HD and was always greatest when imputing from 50K to HD (Table 2). A similar conclusion was evident for the imputation of heterozygous genotypes. The accuracy of imputation

of heterozygous genotypes was lower than the accuracy of imputation of homozygous genotypes.

Table 1. Correlation between true and imputed genotypes for each breed both within and across breeds for the different imputation scenarios

Breed	LD to 50K		LD to HD		50K to HD	
	Across	Within	Across	Within	Across	Within
AA	0.942	0.962	0.940	0.951	0.988	0.988
BB	0.931	0.950	0.918	0.933	0.981	0.980
CH	0.952	0.964	0.948	0.960	0.990	0.990
HE	0.949	0.970	0.949	0.960	0.990	0.991
HF	0.928	0.943	0.920	0.937	0.981	0.982
LM	0.943	0.959	0.941	0.955	0.987	0.989
SI	0.927	0.951	0.922	0.940	0.981	0.983

Results from this study suggest that, in this population at least, and in the scenarios investigated (including the imputation algorithm used) there is no benefit for imputation of a particular breed of exploiting higher density genotypes from multiple breeds. This is likely due to a lack of linkage phases between SNPs across breeds and this hypothesis was substantiated here by the difference between across-breed and within-breed being almost negligible when imputing from 50K to HD. The linkage disequilibrium among breeds between adjacent SNPs in the 50K is likely to be greater than between SNPs on the LD because of the greater marker density in the former. This therefore suggests that there may indeed be some benefit of across breed imputation from HD to sequence data since linkage disequilibrium between adjacent SNPs is likely to be stronger. In an assessment of African-American human subjects for over 500,000 SNPs, Hancock *et al.* (2010) reported reduced imputation accuracy (across different imputation algorithms) when more distantly related individuals were added to the reference population.

Table 2. Proportion of genotypes correctly imputed for the different genotype platform imputation scenarios when the true genotype is homozygous or heterozygous and the imputation is undertaken within breed (Within) or across breeds (Across)

Genotype Platforms	Homozygotes		Heterozygotes	
	Within	Across	Within	Across
LD to 50K	0.962	0.944	0.907	0.879
LD to HD	0.955	0.939	0.900	0.882
50K to HD	0.989	0.987	0.972	0.972

CONCLUSIONS

Imputation accuracy from the medium density genotype panel (50K) to the HD panel was superior to that of imputation from lower density genotype panels. On average the accuracy of imputation was very high. There was, on average, no benefit in imputation accuracy from exploiting a multi-breed reference population and in most instances the accuracy of imputation

reduced when imputation was undertaken using a multi-breed reference population as opposed to a single breed reference population.

REFERENCES

Berry D.P. and Kearney J.F. (2011) *Ani.l* **5:** 1162.
Brøndum R.F., Ma P., Lund M.S. and Su G. (2012). *J. Dairy Sci.* **95:** 6795.
Browning S.R. and Browning B.L. (2007) *Am. J. Human Genet.* **81:** 1084.
Browning B.L. and Browning S.R. (2009) *Am. J. Human Genet.* **84:**210.
Daetwyler H.D., Villanueva B. and Woolliams J.A. (2008) *PLoS ONE* **3**: e3395.
Dassonneville R., Fritz S., Ducrocq V. and Boichard D. (2012) *J. Dairy Sci.* **95:** 4136.
Erbe M., Hayes B.J., Matukumalli L.K., Goswami S., Bowman P.J., Reich C.M., Mason B.A. and Goddard M.E. (2012) *J. Dairy Sci.* **95:** 4114.
Hancock D.B., Levy J.L., Gaddis N.C., Bierut L.J., Saccone N.L., Page G.P. and Johnson E.O. (2012) *PLoS ONE* **7:** e50610.
Hayes B.J., Bowman P.J., Chamberlain A.J. and Goddard M.E. (2009) *J. Dairy Sci.* **92:**433.
Huang Y., Maltecca C., Cassady J.P., Alexander L.J., Snelling W.M. and MacNeil M.D. (2012) *J. Anim. Sci.* **90**: 4203.
Meuwissen T.H.E., Hayes B.J., Goddard M.E. (2001) *Genetics* **157:**1819.
Saatchi M., Schnabel R.D., Rolf M.M., Taylor J.F. and Garrick D.J. (2012). *Genet. Sel. Evol.* **44:** 38.
VanRaden P.M., Null D.J., Sargolzaei M., Wiggans G.R., Tooker M.E., Cole J.B., Sonstegard T.S., Connor E.E., Winters M., van Kaam J.B.C.H.M., Valentini A., Van Doormaal B.J., Faust M.A. and Doak G.A. (2013). *J. Dairy Sci.* **96:** 668.
Weigel K.A., de los Campos G., González-Recio O., Naya H., Wu X.L., Long N., Rosa G.J.M. and Gianola D. (2009) *J. Dairy Sci.* **92:** 5248.

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:546-549

EFFECT OF GENOTYPE AND PEDIGREE ERROR ON DETECTION OF RECOMBINATION EVENTS, SIRE IMPUTATION AND HAPLOTYPE INFERENCE USING THE HSPHASE ALGORITHM

Mohammad H. Ferdosi[1], Brian P. Kinghorn[1], Julius H.J. van der Werf[1, 2] and Cedric Gondro[1]

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

SUMMARY

HSPhase is a fast and accurate algorithm for detection of recombination events, sire imputation and haplotype inference of half-sib families. It can be used on data for half-sib families with as few as 4 individuals in a family. The robustness of this algorithm in relation to genotype and pedigree errors was evaluated. If there were more than 20 half-sibs in a family, the performance of the algorithm with 5% pedigree or genotyping errors was still reliable with the accuracy of phasing and imputation above 0.87. These error rates are above those commonly observed in industry data which indicates the algorithm is sufficiently robust for deployment in real world settings. An R package implementing the method is freely available and includes a function to generate diagnostic plots which are very useful to rapidly identify problems in the dataset.

INTRODUCTION

The availability of genotype information on large numbers of dense molecular markers (usually Single Nucleotide Polymorphisms - SNPs) or even full genome sequences has provided interesting challenges with regard to the best use of all available information. One way to start addressing this is by haplotype reconstruction. Although with current technology it is possible to generate phased data directly, it is still expensive and not suitable for routine usage (Browning *et al.* 2011). Alternatively, computational methods can be used to reconstruct haplotypes from genotypes. The most common approaches make use of population wide data and use a hidden Markov model, e.g. as implemented in BEAGLE (Browning *et al.* 2011). These methods and algorithms were mainly developed for the human population structure and few algorithms have been developed specifically for livestock populations that consist of complex pedigree and for which large half-sib families are usually available (Hickey *et al.* 2011*;* Boettcher *et al.* 2004*).*

We developed the HSPhase algorithm to create block structures of haplotype relationships, which are then used to impute/phase sire and phase genotype data specifically for half-sib family groups (Ferdosi *et al.* 2013). In the real world, data on pedigree and genotypes contain errors which could well affect the performance of phasing methods. In this paper we evaluate the robustness of HSPhase to genotype and pedigree errors.

METHODS

The HSPhase algorithm uses opposing homozygotes to create a block structure and finds a parental origin for each SNP allele; therefore, for each individual the haplotype of the strand that belongs to the sire becomes evident (Figure 1-A, 1-B and 1-C). As the parental origin at multiple SNPs becomes evident, sire haplotypes can be imputed by using the block structure and by calculating the average of SNPs that belong to the first or second sire haplotype (Figure 1-D). The haplotype can simply be reconstructed by replacing the haplotype of the sire with the corresponding block in the half-sib structure (Ferdosi *et al.* 2013).

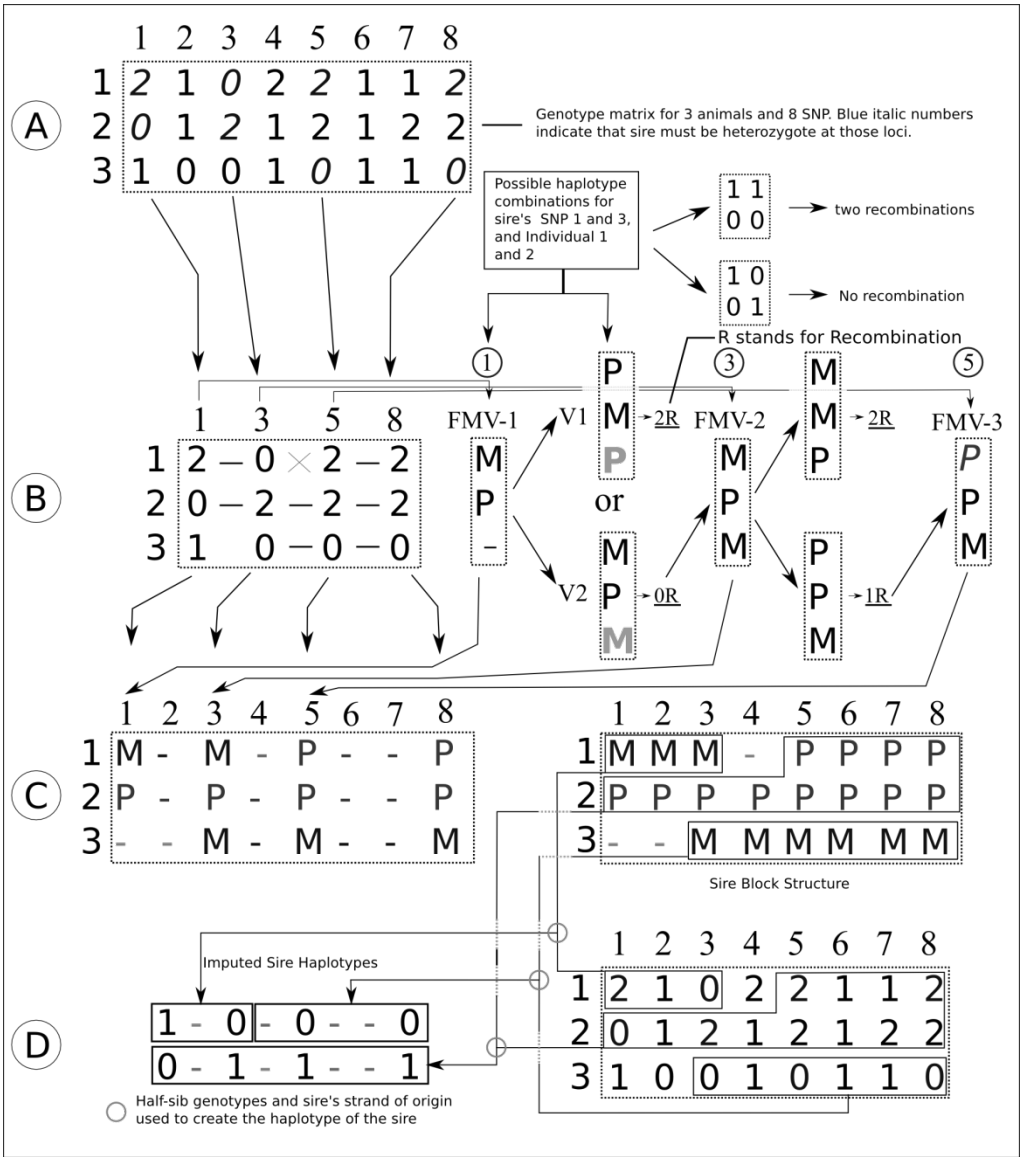


Figure 1. A: Selection of heterozygote loci in sire based on the genotype of the offspring, B: detection of the parental strand of origin of each offspring, C: Filling the gap between different heterozygote, if two loci in one individual have the same parental origin, the SNPs between them are deemed to come from the same haplotype, D: By using the strand origin and genotype the sire's haplotype becomes evident.

The QMsim (Sargolzaei *et al*. 2009) program, which simulates genotype data based on the population structure found in commercial livestock, was used to generate a dataset. A single chromosome of 500 cM in length and 10,000 markers was simulated. For each dataset 20 males were mated to 400 females each generation and genotypes for the last 7 generations were recorded. Each of the final datasets consisted of 120 half-sib groups with 40 offspring in each. Smaller

half-sib families (4, 6, 8, 10, 20 and 40) are sampled randomly from this population by using *sample* function in R (R Core Team 2013).

To evaluate the accuracy of the algorithm with varying rates of genotyping errors, a random proportion of SNP errors were added to the genotype by using the *sample* function. The accuracy of the algorithm was evaluated for 1, 5 and 10 percent genotyping errors. To evaluate the effect of pedigree errors, an individual in the half-sib family was replaced with a random genotype; therefore, we had a pedigree error for different size of half-sib groups and the critical number of offspring that require to handle one pedigree error was estimated.

The accuracy of the method was tested as the squared correlation (R^2) between true and inferred results using the *lm* function in R (R Core Team 2013) using different numbers of half sibs in each family group.

RESULTS AND DISCUSSION

R^2 between true and detected blocks and haploypes were calculated to evaluate the effect of pedigree and genotype errors on the HSPhase algorithm (Tables 1 and 2).

Pedigree errors had a negative effect on the accuracy of the method when less than 10 half-sibs were available. This was mainly driven by genotypes that did not belong to the half-sib group, incorrectly suggesting a heterozygous site in the sire. And also, when the number of half-sibs in the group is limited, wrong genotypes will have a more significant effect on the detection of recombination. As the number of individuals in the half-sib group increases the effect of an incorrect genotype decreases and with 10 half-sibs R^2 values were generally high. Results also show that generally with more individuals per family the accuracy increased (Table 1).

Table 1. Effect of pedigree errors on accuracy (R^2 +/- standard deviation between inferred and true results) and call rate (percentage of known results) for various HS family sizes

	4	6	8	10	20	40
BP	0.58±0.38	0.77±0.32	0.90±0.19	0.93±0.10	0.96±0.03	0.96±0.02
PB%	0.95±0.13	0.98±0.02	0.99±0.02	0.99±0.02	0.99±0.02	0.99±0.01
BPE	0.09±0.19	0.18±0.20	0.46±0.29	0.65±0.24	0.85±0.04	0.90±0.02
BPE %	0.99±0.02	0.98±0.01	0.99±0.01	0.98±0.02	0.99±0.02	0.99±0.01
SI	0.75±0.25	0.87±0.20	0.95±0.11	0.97±0.06	1.00±0.01	1.00±0.00
SI%	0.50±0.09	0.69±0.06	0.80±0.06	0.88±0.06	0.99±0.01	1.00±0.00
SIPE	0.43±0.16	0.52±0.18	0.71±0.21	0.85±0.17	0.98±0.02	1.00±0.00
SIPE%	0.54±0.04	0.71±0.05	0.82±0.04	0.89±0.03	0.99±0.01	1.00±0.00
HI	0.96±0.13	0.97±0.03	0.98±0.02	0.98±0.02	0.98±0.01	0.99±0.01
HI%	0.67±0.09	0.79±0.04	0.87±0.03	0.92±0.03	0.99±0.01	1.00±0.00
HIPE	0.45±0.06	0.55±0.04	0.65±0.03	0.72±0.02	0.88±0.01	0.94±0.01
HIPE%	0.75±0.03	0.84±0.02	0.90±0.02	0.94±0.02	0.99±0.00	1.00±0.00

BP: Block Partitioning, PB%: Percent of Known Block, BPE: Block Partitioning with 1 pedigree error in the half-sib family, BPE %: Percent of Known Block Partitioning with 1 pedigree error in the half-sib family, SI: Sire Imputation, SI%: Percent of Known Sire, SIPE: Sire Imputation with 1 pedigree error in the half-sib family, SIPE%: Percent of Known Sire with 1 pedigree error in the half-sib family, HI: Haplotype Inference, HI%: Percent of Known Haplotype, HIPE: Haplotype Inference with 1 pedigree error in the half-sib family, HIPE%: Percent of Known Haplotype with 1 pedigree error in the half-sib family.

With sufficient half-sibs, pedigree errors can be easily detected by counting the number of recombination events. Figure 2 shows the image plot for a half-sib group. The second individual does not belong to this family due to excessive recombinations. This is an easy way to identify pedigree errors. Since the algorithm can phase the sire of half-sib groups with high accuracy, by counting the number of opposing homozygotes between this individual and other sires, it is easy to

detect the half-sib family that this individual belongs to, provided that sufficient SNPs are available (Figure 2).

The haplotype accuracy of HSPhase did not markedly decrease with genotype errors of up to 5 percent, provided that more than 10 half-sibs were available. A higher error rate significantly decreased accuracy across all family sizes, but especially when less than 20 half-sibs were available (Table 2). These results show that the algorithm is robust to both genotyping and pedigree errors beyond the levels commonly observed in livestock data. This makes it suitable for routine adoption by applications that require phasing and/or imputation.

Table 2. Effect of genotype errors on accuracy (R^2 +/- standard deviation between inferred and true results) and call rate (percentage of known results) for various HS family sizes

	4	6	8	10	20	40
BPE1%	0.50±0.38	0.70±0.34	0.86±0.23	0.91±0.11	0.95±0.03	0.95±0.02
BPE5%	0.40±0.36	0.52±0.38	0.73±0.33	0.86±0.18	0.92±0.10	0.93±0.03
BPE10%	0.28±0.29	0.36±0.31	0.60±0.36	0.71±0.31	0.88±0.12	0.88±0.08
SIPE1%	0.71±0.23	0.82±0.21	0.93±0.13	0.96±0.06	0.99±0.01	1.00±0.00
SIPE5%	0.61±0.22	0.68±0.22	0.81±0.18	0.90±0.09	0.97±0.04	1.00±0.00
SIPE10%	0.50±0.17	0.55±0.19	0.70±0.20	0.79±0.16	0.95±0.06	0.99±0.05
HIPE1%	0.95±0.38	0.94±0.34	0.94±0.23	0.95±0.11	0.96±0.03	0.96±0.02
HIPE5%	0.85±0.36	0.84±0.38	0.84±0.33	0.85±0.18	0.87±0.10	0.88±0.03
HIPE10%	0.75±0.29	0.73±0.31	0.73±0.36	0.74±0.31	0.77±0.12	0.78±0.08

BPE1%, BPE5%, BPE10%: Block Partitioning with 1%, 5% and 10% Genotyping Errors, SIPE1%, SIPE5%, SIPE10%: Sire Imputation with 1%, 5% and 10% Genotyping Errors, HIPE1%, HIPE5%, HIPE10%: Haplotype Inference with 1%, 5% and 10% Genotyping Errors

Figure 2. Block structure in a half-sib family of chromosome 1 using real data from Hanwoo cattle with 11 individuals (dark and light gray are for the first and second haplotype of the sire; markers of unknown origin are shown in white). The second individual is not related to this sire given the large number of recombination events observed.

ACKNOWLEDGEMENTS

This study was funded by a grant from the Next-Generation BioGreen 21 Program (No. PJ008196), Rural Development Administration, Republic of Korea

REFERENCES

Boettcher P.J., Pagnacco G., and Stella A. (2004) *J. of Dairy Sci.* **87:** 4303.
Sargolzaei M. and Schenkel F.S. (2009) *Bioinformatics* **25**: 680.
Browning S.R. and Browning B.L. (2011) *Nature Reviews Genetics* **12**: 703.
Hickey J., Kinghorn B.P., Tier B., Wilson J.F., Dunstan N., and Van der Werf J.H.J (2011) *Genetics Selection Evolution* **43:** 12.
Ferdosi M.H., BP Kinghorn, JHJ van der Werf and C Gondro (2013) (under review)
R Core Team (2013). *R Foundation for Statistical Computing*, Vienna, Austria

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:550-553

ACCURACY OF IMPUTATION IN A POPULATION OF TROPICAL COMPOSITE CATTLE WITH PARTICULAR EMPHASIS ON THE USE OF ALLELIC R^2 AS A QUALITY CONTROL METRIC

M. Kelly, M.R.S. Fortes and S.S. Moore

The University of Queensland, Queensland Alliance for Agriculture and Food Innovation, Centre for Animal Science, Brisbane, QLD 4062

SUMMARY

Imputation of genotypes from low-density single nucleotide polymorphism (SNP) panels to higher density panels is a common approach applied to increase the density of genotypes used in genomic selection and genome wide association studies (GWAS). Accuracy of imputation from Illumina BovineSNP50 to Illumina HD SNP panels was assessed within tropical composite beef cattle using 589 animals. The average imputation accuracy was high according to the percentage of concordant genotype calls (CONCORD) (96%) or the correlation between actual and imputed genotypes ($r^2_{(a,i)}$)(0.94). Considering imputed genotypes for a genome wide association study, we estimated that on average the power of GWAS to be approximately 12% less than genotyping all animals on Illumina HD. The accuracy of imputing individual SNPs was found to vary substantially, depending on multiple factors such as minor allele frequency and chromosome. There was a large number of SNPs for which the $r^2_{(a,i)}$ was less than 0.9. The allelic R^2 statistic reported by BEAGLE was able to identify a large number of such SNP. Placing a threshold on allelic R^2 statistic resulted in a marginal increase in average correlation between actual and imputed genotypes but a large decrease in the percentage of SNP with $r^2_{(a,i)}$ less than 0.81 (from 14% to 2.4%)

INTRODUCTION

Imputation of genotypes across different single nucleotide polymorphism (SNP) panels or from low density panels to high density panels is a routine way of increasing the number of markers for genomic selection (GS) and genome wide association studies in livestock. On average imputation accuracy is high and so genomic breeding values developed on imputed or actual genotypes are highly correlated (Brondum *et al.* 2012, Erbe *et al.* 2012). The impact of imputation accuracy on GWAS is less well understood. Additionally, the impact of using imputation on a diverse multi-breed reference population, such as the Tropical Composite beef cattle from northern Australia merits investigation. Breed diversity may have a negative impact on imputation accuracy and therefore it may affect both GWAS.

The aims of this study were: 1) to test the accuracy of imputation in a population of tropical composite beef cattle, 2) to test the effectiveness of using quality control statistics as a threshold for removing poorly imputed SNPs.

MATERIALS AND METHODS

Animals and genotypes. Genotype data from 589 Tropical Composite animals were used in this analysis. The Tropical Composite cattle consisted of both crossbred cattle and stabilised crosses from a range of founder breeds. Details on management and breeding of this cattle population developed by the Cooperative Research Centre for Beef Genetic Technologies are provided elsewhere (Barwick *et al.* 2009, Burns *et al.* 2013, Corbet *et al.* 2013). The Illumina HD bead chip was used to genotype the samples according to the manufacturer's protocols (Illumina Inc., San Diego, CA). Standard quality control: SNPs with call rate < 0.9 or

minor allele frequency < 0.01 were excluded. Missing genotypes were imputed using BEAGLE 3.2 (Browning and Browning, 2009). Quality control and imputation for missing genotypes resulted in 729,068 SNP with complete genotypes for 589 cattle.

Imputation from low density SNP panel. Imputation from the intersecting SNPs from Illumina BovineSNP50 to Illumina HD (729,068 SNP) was performed using the default settings in BEAGLE (Browning and Browning 2009). A 30 fold cross validation was used to ensure that the reference set of genotypes used to impute new genotypes was representative of the full reference population. The cross validation was performed in 3 steps as follows: 1) Groups of 20 animals were randomly allocated into 30 cross validation sets, 2) One set of 20 animals was imputed from BovineSNP50 to Illumina HD using the remaining groups HD SNP as reference genotypes, 3) this process was performed 30 times so each group had been used as a test set once.

Imputation accuracy and analysis. The accuracy of imputation calculated across animals within SNP was assessed two ways: 1) the concordance between actual and imputed genotype calls (CONCORD) and 2) the correlation between actual and imputed number of copies of the Allele coded B according to Illumina's A/B coding convention ($r^2_{(t,i)}$). The correlation was used as the primary statistic for assessing imputation as it is less influenced by minor allele frequency (Browning and Browning 2009). When imputing data generally we do not know the true accuracy of imputation for each SNP, BEAGLE provide a statistic called the allelic r^2 (R^2_{est}) which estimates the squared correlation between actual and imputed SNP. The effectiveness of this measure in identifying SNP with low CONCORD and ($r^2_{(a,i)}$) was assessed.

RESULTS AND DISCUSSION

On average imputation was good with a concordance rate of 0.96 and a ($r^2_{(a,i)}$) of 0.88 (Table 1). Thus the power of performing GWAS using imputed genotypes would be approximately 12% lower than using Illumina HD genotypes.

Table 1 Summary of concordance and correlation between actual and imputed genotypes with an increasingly stringent threshold applied using allelic r^2

Threshold on R^2_{est}	CONCORD	$r^2_{(t,i)}$	Markers excluded (%)
0	0.96	0.88	0.0
0.5	0.96	0.89	1.7
0.75	0.96	0.90	6.7
0.95	0.99	0.97	70.7

The measures of imputation accuracy in Table 1 are comparable with other studies performed in cattle with Erbe *et al.* (2012) finding concordance of actual and imputed genotypes of 0.97 in Holsteins and 0.96 in Jersey cattle. Present results were on the lower range of correlations between actual and imputed genotypes of 0.92-0.98, reported by Brondum *et al.* (2012). A slight reduction in imputation accuracy may be expected in the current study due to diverse genetic background of the cattle under investigation. Although the average imputation accuracy was quite high there was substantial variation in imputation accuracy. Imputation accuracy was affected by a number of factors including minor allele frequency and chromosomes, chromosome X in particular was imputed with lower accuracy.

As the threshold on R^2_{est} for excluding SNP became increasingly stringent the mean CONCORD and $r^2_{(a,i)}$ were high their means increased from 0.961 to 0.986 and 0.88 to 0.97 for CONCORD and $r^2_{(t,i)}$ respectively (Table 1). The editing of SNP based on R^2_{est} also decreased the number of SNP with low call rates, this is demonstrated visually in Figure 1 where fewer SNP with low $r^2_{(t,i)}$ appear successively from (a) through to (d). Additionally R^2_{est} was highly correlated with $r^2_{(a,i)}$(0.81).

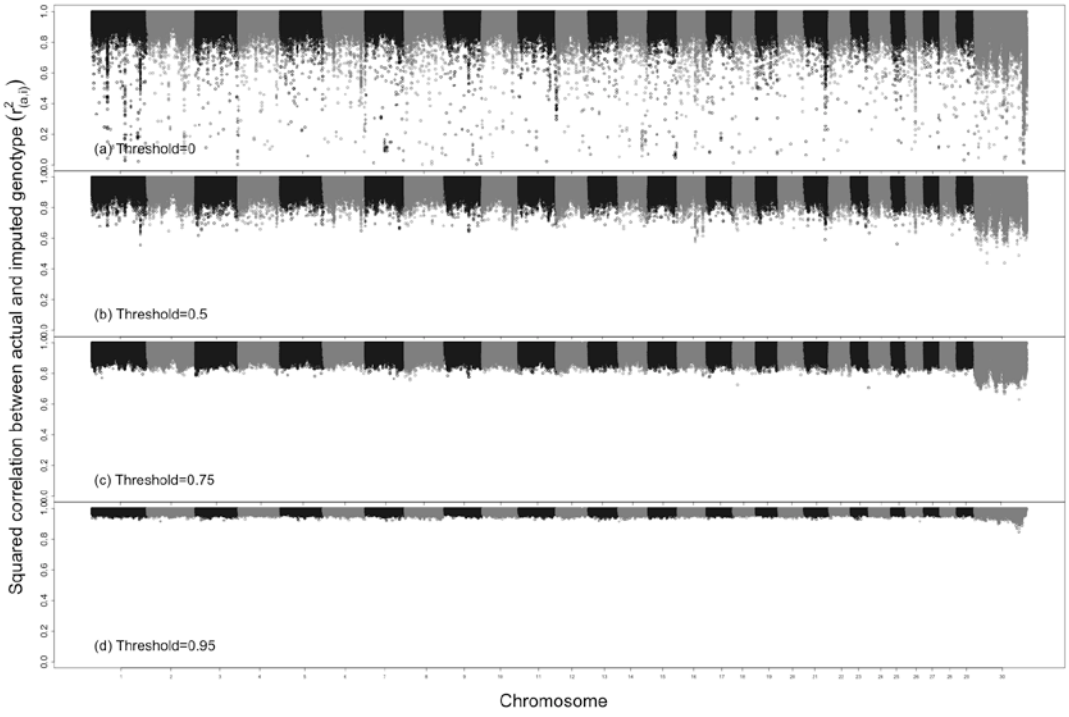


Figure 1. The correlations between imputed and actual genotypes with increasingly stringent thresholds applied using allelic r^2.

Figure 1 shows that many SNPs were imputed with low accuracy. The ability to identify such SNPs was examined by considering the ability of R^2_{est} to identify SNPs with CONCORD or $r_{(a,i)}$ lower than 0.9. False negatives were be defined as SNPs with correlation or concordance lower than 0.9 that were not excluded by quality control. Conversely, false positives would be SNPs with correlation or concordance greater than 0.9 that were excluded. As the R^2_{est} threshold for selecting SNPs becomes more stringent the number of false negatives decreases substantially (Table 2). There is a trade off as the number of false positives also increases, this is especially evident when the threshold is 0.9 or above. A reasonable compromise is to set the threshold to approximately 0.8 where false negatives (for the correlation) are reduced from 14.8% to 2.4% while false positives are 9.4%.

Table 2 Percentage of false negatives and false positives for concordance and correlation with an increasingly stringent threshold applied to BEAGLE r^2

	Percentage false negatives[*1]		Percentage false positives[*2]	
Threshold on allelic r^2	Concordance	Correlation	Concordance	Correlation
0	3.7	14.8	0.0	0.0
0.5	3.6	13.6	1.6	0.4
0.75	2.9	10.5	5.9	2.0
0.95	0.0	0.0	69.6	65.6

[*1] Percentage false negatives: percentage of SNPs with correlation or concordance lower than 0.9 that were not excluded; [*2]Percentage of false positives: percentage of SNPs with correlation or concordance greater than 0.9 that were excluded.

The current study focused on a small part of genotype quality control for use of imputed genotypes in GWAS studies. Attention must be played to quality control at all stages of the analysis. The detection of imputation accuracy per individual animal would also be an important step to improve the overall quality control. It was found that the genotype probability of each genotype call averaged over each animal was not related to overall imputation accuracy (data not shown).. In summary, special consideration of individual SNP imputation accuracy could avoid detection of false QTL, when performing genome wide associations with imputed SNP data. It is possible to use R^2_{est} as a quality control statistic to reduce imputation accuracy issues.

ACKNOWLEDGEMENTS

Northern Pastoral Group, Department of Employment, Economic Development and Innovation (DEEDI) and CSIRO Animal Food and Health Science. Animal Genetics and Breeding Unit in particular Bruce Tier. Meat and Livestock Australia, Australian Centre for International Agricultural Research, Cooperative Research Centre for Beef Genetic Technologies, Northern Pastoral Group, CSIRO – Animal Food and Health Science, DEEDI and The University of Queensland gave financial support.

REFERENCES

Barwick S. A., Johnston D. J., Burrow H. M., Holroyd R. G., Fordyce G., Wolcott M. L., Sim W. D. and Sullivan M. T. (2009) *Animal Production Science* **49**: 367.

Brondum R. F., Ma P., Lund M. S. and Su G. (2012) *J Dairy Sci* **95**: 6795.

Browning B. L. and Browning S. R. (2009) *Am J Hum Genet* **84**: 210.

Burns B. M., Corbet N. J., Corbet D. H., Crisp J. M., Venus B. K., Johnston D. J., Li Y., McGowan M. R. and Holroyd R. G. (2013) *Animal Production Science* **53**: 87.

Corbet N. J., Burns B. M., Johnston D. J., Wolcott M. L., Corbet D. H., Venus B. K., Li Y., McGowan M. R. and Holroyd R. G. (2013) *Animal Production Science* **53**: 101.

Erbe M., Hayes B. J., Matukumalli L. K., Goswami S., Bowman P. J., Reich C. M., Mason B. A. and Goddard M. E. (2012) *J. Dairy Sci.* **95**: 4114.

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:554-557

UTILITY OF IMPUTED SNP GENOTYPES FOR GENOME-WIDE ASSOCIATION STUDIES IN DAIRY CATTLE

M.S. Khatkar, P.C. Thomson and H.W. Raadsma

ReproGen, Faculty of Veterinary Science, University of Sydney, Camden, NSW 2570

SUMMARY

Comparisons of genome-wide association studies (GWAS) based on imputed and actual genotypes were made using a dataset of 2,205 dairy bulls genotyped with a 50K SNP chip. The animals were divided into a reference (25 %) and a test panel (75 %). The genotypes of the test animals specific to two commercial lower density chips (*i.e.* 3K and 7K) were imputed up to the 50K using the IMPUTE2 software. The 'best guess' genotypes and allele dosages (estimated number of copies of an allele) were used as imputed genotypes. The association of SNP genotypes with phenotypes were conducted on five dairy traits (*viz*. milk yield, fat yield, protein yield, survival and daughter fertility) using true and imputed 50K genotypes of the test animals. The accuracy of imputation had a clear impact on the ability to detect the significant associations but varied between the 3K and the 7K, and among the five traits. The allele dosage model was superior to the best-guess model. Filtering the SNPs based on an indirect indicator of accuracy of imputation significantly improved the repeatability of GWAS results obtained from the imputed genotypes. Overall our results show that imputed genotypes can be used effectively to increase the power of GWAS.

INTRODUCTION

A number of SNP chips varying in SNP density and cost are available for genotyping cattle. For the dairy industry, an attractive strategy to increase genotypic information in a population whilst keeping cost of genotyping down is to genotype a large number of animals with a cheaper low-density SNP chip and impute up to high density genotypes using a limited number of reference animals genotyped with a high-density SNP chip (Khatkar *et al.* 2012). In addition to the primary utility of using imputed genotypes for genomic selection, such high-density imputed SNP genotypes on a large number of animals can boost the power of genome-wide association studies (GWAS) and fine-mapping of causal variants (Marchini and Howie 2010). GWAS rely on linkage disequilibrium (LD) between genotyped SNPs and causal mutations and hence benefit from the availability of very high-density SNP panels genotyped on large numbers of animals. In addition, genotype imputation is becoming a popular approach for combining multiple resource populations genotyped using different SNP panels, especially for meta-analysis (de Bakker *et al.* 2008; Jiao *et al.* 2011).

Imputation of genotypes is generally achieved with some uncertainty which may affect the ability to detect SNP associations. A number of studies have examined the accuracy and utility of imputed genotypes for GWAS in human (Marchini and Howie 2010). However, to our knowledge no study has been undertaken in livestock. The population structure, traits and density of the SNP panels in use in livestock are quite different from those in human. Such an analysis would provide useful information for conducting GWAS on imputed genotypes in cattle. Here we compared GWAS based on imputed and actual genotypes using a dataset of dairy cattle genotyped with a 50K SNP chip. We compared two types of imputed genotypes *viz.* 'best guess' and 'allele dosage', and investigated the effect of imputation accuracy on the repeatability of SNP association tests.

MATERIAL AND METHODS

Data. A total of 2,205 bulls genotyped with the Illumina BovineSNP50 chip were used in this

study (Khatkar *et al.* 2012). After filtering the SNP for low minor allele frequency (MAF>1%) and other QC measures, a total of 41,864 SNPs mapped on autosomes on UMD3.0 were used in this study.

Imputation. The 2,205 animals were divided into a reference (25 %) and a test panel (75 %). The 550 animals in the reference panel were selected randomly from the animals born before 2001 and all remaining animals were included in the test panel. For the reference panel, all the 50K genotypes were used. For the test panel a subset of the 50K SNP genotypes specific to two commercial lower density chips, *viz.* 3K (Wiggans *et al.* 2012) and 7K (Boichard *et al.* 2012), were used. Most of the SNPs on the 3K and the 7K chips are present on the 50K chip. The genotypes of the test animals were imputed up to the 50K using the genotypes of the reference animals. IMPUTE2 version 2.1.2 (Howie *et al.* 2009) was used for imputation. The 'best guess' genotype and the allele dosage were used as imputed genotypes. Allele dosage is the expected count of the B-allele.

Accuracy of imputation. Correlations between the actual and imputed genotypes were computed for each SNP by coding the AA, AB, BB genotypes as 0, 1, 2. In addition mean allelic error rates for the imputed genotypes were computed as the percentage of incorrectly predicted alleles *i.e.* mean allelic error rate (%) = number of incorrectly predicted alleles / total number of alleles imputed in the test set × 100.

SNP association. Association of SNP genotypes with five dairy traits (daughter trait deviations, DTD) were computed using the actual 50K and imputed 50K genotypes of the test animals. The five traits analysed were milk yield, fat yield, protein yield, survival and daughter fertility index which reflect a range of heritabilities. The regression of the traits on SNP genotypes were conducted by fitting the SNP allele count or allele dosage as a covariate and animal additive genetic effect as a random effect in a linear mixed model using ASReml (Gilmour, 2009). In addition each observation was weighted with the accuracy of DTD of each bull. The correlation of $-\log_{10}$(*p*-values) obtained by original 50K *vs.* imputed 50K was taken as the accuracy/ repeatability of GWAS on imputed genotypes for each trait.

RESULTS AND DISCUSSION

Overall agreement of SNP genotype association with milk volume as obtained using original and imputed genotypes *i.e.* best guess genotypes and allele dosage is presented in Figure 1. These results are based on imputed genotypes obtained by using the 3K SNP chip on the test animals. The repeatability of the *p*-values obtained using imputed allele dosage (0.92) was higher than the repeatability using best guess genotypes (0.89). Similar results were observed for other traits and when using the 7K SNP chip (results not shown). Higher repeatability using allele dosage could be expected as the probabilities of calling correct genotypes by imputation are included in the computation of allele dosage.

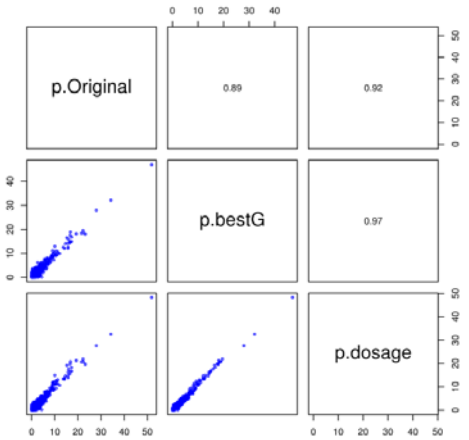


Figure 1. The repeatability of SNP associations with milk volume using imputed genotypes. The values in the upper triangle are Pearson correlation coefficients between -$\log_{10}$(*p*-values) using respective genotypes.

The results presented in Table 1 show further details on the repeatability of SNP association with phenotypes, where SNP genotypes were allele dosages from imputation. The correlation between -log10 of *p*-values varies from 0.84 to 0.92 across five traits. To examine the effect of accuracy of imputation on repeatability of association, the SNPs were classified according to their imputation accuracy. The SNPs with imputation accuracies less than 0.75 have low repeatability (Table 1). These results suggest that imputed genotypes of the SNPs with high error rate may not be useful for GWAS.

Table 1. The agreement of *p*-values for GWAS for five different traits obtained using actual genotypes and imputed genotypes (allele dosages)

Imputation accuracy (range)	n snp	MAER	Imputation accuracy	Repeatability of p-values				
				Milk volume	Fat	Protein	Direct survival	Cow fertility
ALL	39226	3.589	0.902	0.918	0.904	0.879	0.835	0.841
0.0-0.5	87	12.943	0.345	0.120	-0.04	0.104	0.227	0.366
0.5-0.75	1025	5.108	0.688	0.529	0.476	0.586	0.569	0.526
0.75-0.9	12484	4.786	0.857	0.860	0.821	0.808	0.757	0.779
0.9-0.95	19971	3.246	0.927	0.945	0.932	0.912	0.865	0.878
0.95-1.0	5659	1.738	0.963	0.947	0.951	0.946	0.934	0.927

Imputation accuracy is the correlation coefficient between imputed dosage and true genotypes; Repeatability of *p*-values = cor($-\log_{10}$(*p*-values- actual), $-\log_{10}$(*p*-values- imputed)); MAER = mean allelic error rate (%).

The accuracy of the imputation of untyped SNPs cannot be estimated in the absence of any true genotypes for comparison. However, it is possible to have some indication of quality of imputed genotypes. Browning and Browning (2009) suggested using the Pearson correlation between best guess and allele dosage as an indicator of accuracy of imputation. Figure 2 shows the relationship of this indicator with the accuracy of imputation. These results suggest that a large proportion of the SNPs with low accuracy of imputation can be filtered out by using the correlation between best guess and allele dosage as indirect measures. Such a filtering step can significantly improve the results of GWAS obtained from imputed genotypes.

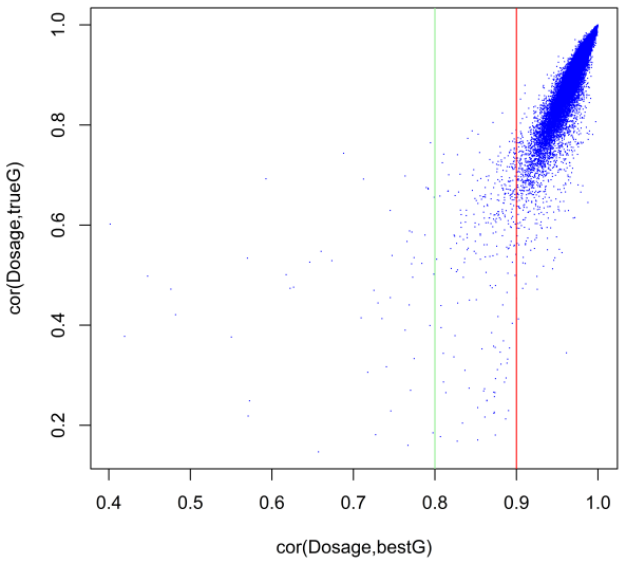


Figure 2. The relationship of correlation between allele dosage and best guess (x-axis) with the accuracy of imputation (y-axis; correlation between dosage and true genotypes).

The main motivation for undertaking a GWAS is usually to identify signals for causal variants or SNP in LD with such variants. Because of high LD between SNPs, especially when high-density SNP chips are used, true signals are generally represented by multiple SNPs in the region. The repeatability of individual SNPs from imputed genotypes is variable as discussed above, however, when the number of SNPs in a sliding window was used to detect the signal, the repeatability of signals using imputed GWAS was higher (results not shown).

Low MAF also affects the accuracy of imputation (Khatkar *et al.* 2012) and hence accuracy of association in GWAS. We excluded all SNPs with a MAF less than 1%. Excluding SNPs with very low MAF and filtering with the indicator of accuracy of imputation (Figure 2) can improve the GWAS results obtained from imputed genotypes.

We only tested the additive genetic effect of the SNP allele. It is possible to use the data on the cows to estimate the dominance effect by contrasting the mean of three genotypes. Such analysis will require using best guess imputed genotypes. With the availability of different SNP panels for bovine, it is becoming common place to genotype the same or different resource populations with different SNP chips. Imputation can help to combine such datasets. Recently we showed that the genotypes of animals can be imputed from 50K to 800K with a very small loss of accuracy of imputation (Khatkar *et al.* 2012). Such high-density imputed datasets will provide resources to conduct very powerful GWAS whilst maintaining the cost of genotyping at a low level.

ACKNOWLEDGEMENT

The data resource was generated by the CRC-IDP, the Department of Primary Industries, Victoria and the University of Sydney, with input from Drs Ben Hayes, Phil Bowman, Amanda Chamberlain, Matthew Hobbs and Mrs Gina Attard. The semen samples were provided by Genetics Australia and the phenotypic data by Australian Dairy Herd Improvement Scheme.

REFERENCES

Boichard D., Chung H., Dassonneville R., David X., Eggen A., Fritz S., Gietzen K.J., Hayes B.J., Lawley C.T., Sonstegard T.S. , et al.(2012) *PLoS ONE* **7**: e34130.

Browning B.L. and Browning S.R. (2009) *Am J Hum Genet* **84**: 210.

de Bakker P.I., Ferreira M.A., Jia X., Neale B.M., Raychaudhuri S. and Voight B.F. (2008) *Hum Mol Genet* **17**(R2): R122.

Gilmour A.R., Gogel B.J., Cullis B.R., Thompson R. (2009) ASReml User Guide Release 3.0. VSN International Ltd, Hemel Hempstead, HP1 1ES, UK http://www.vsni.co.uk.

Howie B.N., Donnelly P. and Marchini J. (2009) *PLoS Genet* **5**: e1000529.

Jiao S., Hsu L., Hutter C.M. and Peters U. (2011) *Genet Epidemiol* **35**: 597.

Khatkar M.S., Moser G., Hayes B.J. and Raadsma H.W. (2012) *BMC Genomics* **13**: 538.

Marchini J. and Howie B. (2010) *Nat Rev Genet* **11**: 499.

Wiggans G.R., Cooper T.A., Vanraden P.M., Olson K.M. and Tooker M.E. (2012) *J Dairy Sci* **95**: 1552.

Proc. Assoc. Advmt. Anim. Breed. Genet. 20:558-561

HOW ANGUS BREEDERS HAVE REDUCED THE FREQUENCY OF DELETERIOUS RECESSIVE GENETIC CONDITIONS

C.F. Teseling and P.F. Parnell

The Angus Society of Australia, 86 Glen Innes Road, Armidale, NSW, 2350

SUMMARY

Undesirable genetic conditions are present in all species of livestock and could range from causing a reduction in performance or structural problems to being semi-lethal or lethal. Normally the frequency of genetic conditions is low and therefore not considered to be of significant economic importance. However, sometimes the frequency can increase, especially if the progeny of carrier animals perform for economically important traits. Through artificial breeding it is possible for one sire to generate thousands of progeny in a few years, increasing the population's co-ancestry and risk through inbreeding.

The Angus Society of Australia has adopted a policy of supporting the management of undesirable genetic conditions rather than vigorously pursuing their eradication. The development of highly accurate DNA tests and the implementation of GeneProb have made it possible to simultaneously reduce the carrier frequency for three recessive genetic conditions from approximately 7% to approximately 2% in less than four years.

INTRODUCTION

Mutations occur in the cells of living organisms and are a source of new genetic variation that is necessary for selection and genetic improvement. A mutation results in a change in genotype and when inherited by progeny, it can be beneficial or detrimental and can impact on fertility, performance or structural soundness. Some mutations result in a change in phenotype and in most cases this is how the mutation is eventually detected.

Artificial breeding has many advantages which have resulted in a steady increase in its popularity to the point where 40% of the registered Angus calves in Australia are now the product of artificial insemination while 10% are from embryo transplantation. Through artificial breeding thousands of progeny can be generated by a single sire which can spread genetic conditions through a herd or breed very quickly.

Over the past two decades Angus breeders have very effectively used Breedplan to identify genetically superior animals and then utilised artificial breeding to multiply desirable genetics at the breed level. This strategy (effectively low levels of breed-wide line breeding) has resulted in significant genetic gain in the breed and also the emergence of several very popular bulls that had a large influence on the breed's genetic composition.

THE MOST IMPORTANT GENETIC CONDITIONS IN ANGUS CATTLE

α-Mannosidosis. In the late 1970's it was estimated that about 5% (Peter Healy, personal communication) of Angus stud cattle in Australia could be carriers. The Angus Society decided to eradicate this genetic condition from the seedstock population by only allowing the registration of animals that were tested free, or progeny of free animals. It is still a requirement to test imported genetics to ensure they are not carriers.

Arthrogryposis Multiplex (AM). In 2008 this deleterious genetic condition was discovered in the USA with RITO 9J9 OF B156 7T26 (born in 1979) as the most probable progenitor. In late 2008, a diagnostic test was developed by Professor Beever from the University of Illinois in the USA and rapidly applied in Australia.

Neuropathic Hydrocephalus (NH). In 2009 this deleterious genetic condition was discovered in the USA with G A R PRECISION 1680 (born in 1990) as the most probable progenitor. A diagnostic test, also developed by Professor Beever, became available to Australian breeders in mid-2009.

Contractural Arachnodactyly (CA). In 1996 the first case of Fawn Calf Syndrome (the condition's initial name), was reported in Australia. Extensive pedigree analysis by Animal Genetics and Breeding Unit scientists identified FREESTATE BARBARA 871 OF KAF in 1999 as the most probable progenitor. In the absence of a DNA test, knowledge of known implicated bloodlines helped Angus breeders keep the frequency of carrier animals at a relatively low level until a DNA test was developed by Professor Beever in mid-2010.

ANGUS SOCIETY POLICY

Historically breed societies have been accused of ignoring the issue of genetic conditions and not informing and/or educating the broader beef industry about the proper management of genetic conditions. The discovery of three recessive genetic conditions and the realisation that some of the most widely used bloodlines in the breed were carriers posed a serious threat to the Angus breed.

As a result of the discovery of AM, the Board of the Angus Society of Australia focused their annual technical workshop in November 2008 on genetic conditions and how the breed should confront this issue. The breed had two options; the first was to attempt to eradicate any existing and future genetic conditions, or secondly use available DNA tests to reduce the gene frequency of AM and then manage it in the population.

The Angus Society Board agreed to develop a system to identify potential carrier animals, facilitate the testing of these animals, and report the results publicly to ensure seedstock and commercial producers were well informed. The focus was to manage genetic conditions rather than attempting to rapidly eradicate them from the population.

At the end of 2012, the Angus Society Board decided to assign a 12.5% probability of being a carrier to all animals in the Angus database with unknown pedigree. This was to reflect the risk associated with base animals of which the parents are unknown.

MANAGEMENT AT THE BREED AND HERD LEVEL

Compiling lists of carrier animals. The first step in managing genetic disorders at the herd level was to compile lists of potential carrier animals and make these lists available to seedstock breeders to help them ascertain their herd's exposure as well as identify animals that may need to be tested.

No matter when a new genetic condition is discovered and a DNA test becomes available, it will always be inconvenient to at least some breeders as it will be too close to their bull sale to allow sufficient time to test sale bulls prior to sale.

To minimise the lag time between when the genetic condition was initially identified and when the test became available in Australia, the ability to send urgent samples to Professor Beever's lab was negotiated. Several batches of urgent samples were couriered to the USA with some results being available 30 days after sample collection.

Development and implementation of GeneProb. GeneProb is a software program developed by Kinghorn (2000) for the analysis of large datasets to calculate the probability of each animal being a carrier of a specified recessive genetic condition.

The Angus Society worked with their database service provider, the Agricultural Business Research Institute, to implement GeneProb to enable weekly analysis of new results and their publication for each animal on the Society's website.

DNA Testing. With the availability of GeneProb results, Angus members were able to efficiently identify potential carrier animals and focus on testing those animals. Immediately after a diagnostic test became available for each newly defined recessive genetic condition, a large amount of testing followed to determine potential carrier animals. After initial testing and identification of potential carrier animals, testing has become cyclic with a drop-off in the first quarter of each year. Interestingly, this is also the time of the year when the least number of bull sales occur.

Table 1 illustrates the number of animals tested for one, two or all three genetic conditions for each of the Angus breed registers. The percentage of animals downgraded (or culled from the breeding herd) was also investigated, and it was found that almost 50% of all animals downgraded were carriers of at least one genetic condition. This indicated the carrier status of an animal was not the only deciding factor determining whether an animal gets downgraded or not.

Table 1. Number of animals in each Angus register tested and found to be carriers of one, two or all three genetic conditions.

Register	No. animals tested				No. carriers			
	1 Cond.	2 Cond.	3 Cond.	Total	1 Cond.	2 Cond.	3 Cond.	Total
HBR/RAR	22,630	7,313	1,986	31,929	10,956	700	8	11,664
APR	6,905	2,504	599	10,008	3,209	230	1	3,440
MBR	182	35	15	232	34	0	0	34
ACR	36	53	13	102	24	5	0	29

Reduction in carrier frequencies. As soon as a new genetic condition was discovered and bloodlines involved were announced, members adjusted their breeding decisions by changing sires selected for the next year's calves. Figure 1 illustrates that the initial reduction in carrier frequency of calves (based on GeneProb results) was closely associated with the timing of the first announcements concerning the discovery of the genetic condition and subsequent development and release of the DNA test.

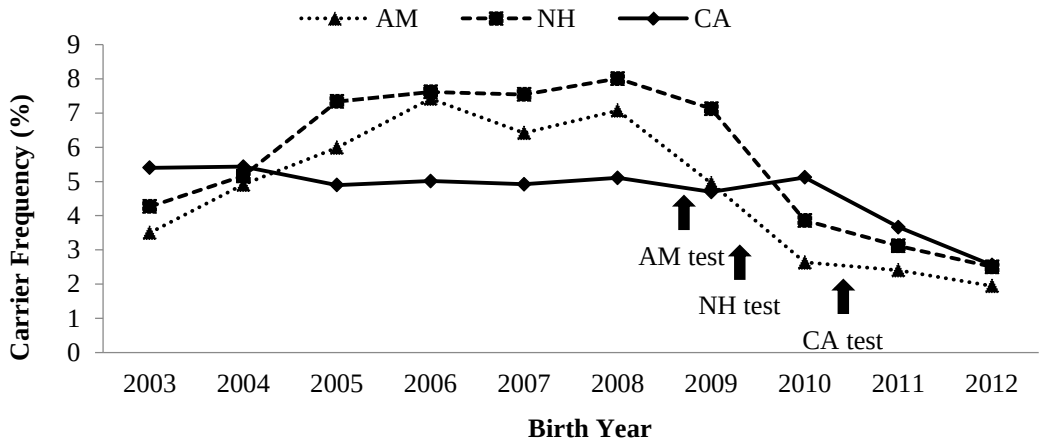


Figure 1. Frequency of carrier calves relative to birth years and the reductions in carrier frequency when DNA tests were made commercially available.

CONCLUSIONS

The initial reduction in the carrier frequency of calves was mainly achieved by using sires that were tested and found to be free of the genetic condition or were expected to be free through pedigree analysis.

Many breeders could not afford to cull all carrier cows and therefore had no option but to continue to breed with some carrier cows. It is expected that the carrier frequencies of calves will continue to decrease as carrier cows are being replaced with cows tested or expected to be free.

The decision of the Angus Society Board to manage, rather than eradicate, recessive genetic conditions has allowed members to respond to this challenge in a financially responsible way. Simultaneously reducing the carrier frequency for three genetic conditions from approximately 7% to approximately 2% in less than four years is a very significant achievement.

ACKNOWLEDGEMENTS

The input the Agricultural Business Research Institute in the application of GeneProb, and the financial support of Meat and Livestock Australia is gratefully acknowledged.

REFERENCES

Kinghorn B.P. (2000) In ‘Animal Breeding – Use of New Technologies’, pp. 112-117, editors B.P. Kinghorn, J.H.J. Van der Werf and M. Ryan, The Post Graduate Foundation in Veterinarian Science of the University of Sydney, Sydney.