A NEW SYSTEM FOR COLLECTING AND PROCESSING PHENOTYPIC AND GENETIC INFORMATION FROM SHEEP FOR IMPROVED SELECTION TOOLS

K.G. Geenty[1], J.H.J. van der Werf[1], K.P. Gore[2], A.J. Ball[3] and S. Gill[3]

[1] CRC for Sheep Industry Innovation and the University of New England, Armidale, NSW 2351
[2]Animal Breeding and Genetics Unit, University of New England, Armidale, NSW 2351
[3]Meat and Livestock Australia, University of New England, Armidale, NSW 2351

SUMMARY

Current genetic improvement schemes are well developed in the Australian sheep industry for standard quantitative production traits. With the advent of new molecular genetic technologies, however, a new sheep industry linked system is being developed by the Sheep CRC based on the Information Nucleus (IN) concept. This sytem will deliver new and improved Australian Sheep breeding Values (ASBVs) for difficult and expensive to measure traits for product quality and production efficiency. The IN system is a world first and uses a network of research flocks representing the wider industry breed structure and combines extensive phenotyping and genotyping of up to 5,000 progeny annually. The IN system has developed new phenotyping protocols with modern electronic data capture, a sophisticated database and is linked with contributing stud breeders via Sheep Genetics (SG). Effective delivery of next generation genetic tools to the industry is facilitated by the stud breeder participation including provision of semen from selected sires and involvement in field activities.

INTRODUCTION

The Australian sheep industry has well developed and effective genetic improvement schemes in MERINOSELECT and LAMBPLAN based on delivery of Australian Sheep Breeding Values (ASBVs) by Sheep Genetics (SG) (Brown *et al.* 2007). The current SG system caters for approximately 52 traits routinely measured by commercial stud breeders. However other traits relevant to the breeding objective such as wool and meat quality that are hard or expensive to measure are not included. Such traits can still be improved if they are correlated with measured traits, but progress is likely to be slow unless the traits are highly correlated.

The Information Nucleus (IN) concept was first proposed by Banks *et al.* (2006) and the world's first IN materialized with a new CRC for Sheep Industry Innovation (Sheep CRC) in 2007. The concept involves intensive measurement of many new traits on a large group of animals linked to the sheep industry through selected sires related to other sire families. Sufficient offspring for each sire in different environments allows estimation of heritabilities and genetic correlations for new traits, and assessment of genotype by environment interactions. The system therefore provides genetic parameters and breeding values for important industry sires including important new traits not currently measured in commercial studs.

The recent development of new molecular genetic technologies such as single nucleotide polymorphism (SNP) chips for genome wide genotyping of dense markers gives the IN potential to add value to the current ASBV system. It allows extension of information on ASBVs for new traits to unrelated animals. Such *genomic selection* is already being implemented in young dairy bulls for milk production traits. The combination of intensive phenotyping of animals, potentially accelerated with genomic selection, will provide commercial stud breeders with next generation genetic tools to enhance genetic improvement, particularly for new traits related to efficiency and product quality.

The IN forms a major platform for research programs in the Sheep CRC, outlined previously by Banks *et al.* (2006) and Fogarty *et al.* (2007). This paper describes in more detail how the

Sheep CRC IN has been developed during the initial two years. The system for collection and processing of phenotypic and genetic information for development of next generation genetic tools is outlined.

SYSTEM DEVELOPMENT

The Sheep CRC system is based on data from five annual matings of approximately 100 new sires each year to 5,000 ewes across eight Australian research stations. Sheep breeds and environments represent a cross section of the Australian sheep industry (Fogarty *et al.* 2007). Seventy percent of progeny are slaughtered in the first year and the remaining 30% are females retained for reproductive evaluation.

Components of the system include-

- Development of data capture and management protocols including quality assurance
- Phenotypic data capture and collection of biological samples from some 5,000 progeny each year
- Genotyping and construction of SNP chips on approximately 2,500 progeny each year
- Construction of a database to accommodate the phenotypic and genetic data
- Association analysis with Sheep Genomics Program (SGP) for new new molecular ASBVs
- Establishment of linkages with other industry databases including SG and the SGP
- Delivery of new genetic tools to the sheep industry

IN flocks and data flow. The eight Information Nucleus flocks, representing the University of New England (Armidale), NSW DPI (Trangie and Cowra), Vic DPI (Rutherglen and Hamilton), SARDI (Struan and Turretfield) and DAFWA (Katanning), have 1,000 base ewes per organization (800 Merino and 200 Maternal except for DAFWA with 1,000 Merinos) joined annually to approximately 40 Merino, 40 terminal, and 20 maternal sires (Fogarty 2007). Quality assurance for phenotypic data is through protocols at IN sites, customised validation routines on entry to the IN database, during preliminary statistical analysis by IN data managers for each trait group and by scrutiny of preliminary IN ASBVs from SG. The IN ASBVs are then combined with industry information for standard traits by SG before delivery to the sheep industry (Figure 1). In addition data for new traits will be developed via prototype ASBVs before validation and delivery to industry by SG.

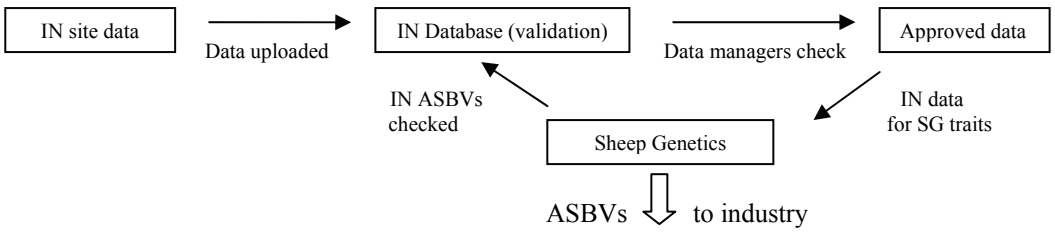

Figure 1. The flow of information from IN sites to SG and then to the sheep industry

Protocols

Detailed data collection protocols, including animal management guidelines and quality assurance, are in the IN Operational Manual (www.sheepcrc.org.au/insite). Also included in the manual is an outline of the IN program, a summary checklist of activities, trait dictionaries,

procedures for collection of biological samples and protocols for nutritional and animal health monitoring and management. Procedures for data transfer with the IN database are included. The Operational Manual is web based with regular updates.

Phenotyping. Phenotypic data for the 4,500-5,000 progeny each year for 163 traits are summarized in Table 1. In the database there are 266 traits including components of those included (eg. multiple values for wool micron profile data, meat colour and fatty acid data). New trait groups are in italics below.

Table 1. Trait groups representing phenotypic data captured from progeny

Program (no. traits)	Main trait groups
Sheep (81)	Lambing including *autopsies*, *visual scores*, live weight/condition score, worms
Meat (44)	Carcase measures, muscling/*yield*/fat, *meat colour/minerals/tenderness/flavo*
Wool (38)	Fleece weight, wool length/strength/micron profile, *colour, UV stability, comfort*

Newly developed e-sheep techniques are used for data capture including RFID tags and readers, auto weighing and drafting and blue tooth technology for wool information and sheep scores.

Genotyping. The first ovine SNP chip was made available in August 2008 by Illumina Inc. This SNP50 BeadChip was developed by an international consortium, with major inputs from Australia and New Zealand allowing genotyping with 55,000 SNPs. The CRC has genotyped 3650 animals using this SNP and along with 4200 genotypes from Sheep Genomics (Oddy *et al.* 2005) an association study is currently being carried out to derive a SNP set and prediction equation for genetic merit of individual traits. A set of about 700 industry sires with accurate ASBVs for commercial traits will be used for validation. Animals genotyped were mainly from the major breeds including Merino, White Suffolk, Poll Dorset and Border Leicester, as well as Texel and Suffolk, to assess genomic diversity and across breed application of genomic selection. The whole genome analysis will provide an insight into the ability to predict ASBVs from genomic data. It may also give further information about the segregation of quantitative trait loci (QTL) in certain regions, potentially providing further information about the genetic regulation of traits.

Database. The INF database (infdata.une.edu.au) is hosted by the Animal Breeding and Genetics Unit (AGBU) at the University of New England and is a relational SQL compliant platform with an ASP.NET front end allowing simultaneous multi-users. Data uploads, extracts and reports are available to acredited IN sites and researchers while technical aspects are controlled by the database manager. Data uploads are verified at submission according to data ranges, allowed values, specified trait names and correct pedigree information. An automated response is sent by email for error reports and/or final confirmation of upload. Real time web based reporting includes data collection statistics, summary statistics and phenotypic animal performance. Data extracts are generic for users or specialized for SG. The many and varied phenotypic traits and large volume of genotyping information are accommodated and configured at the front end.

Industry links and delivery. The major link with the sheep industry is through SG both in sourcing sires for the IN flocks from their database and developing and delivering ASBVs. Involvement by breeders through semen provision and attendance at field events has created significant 'buy in' to the IN program and this helps facilitate effective delivery and uptake of new and improved ASBVs, especially for new traits that are difficult or expensive to measure by stud breeders. In addition the combining of IN data with that from breeders for standard traits provides more data and sire linkage that strengthens the accuracy of ASBVs. The industry 'footprint' of the IN provided by common use of sires and their close relatives is significant. Current estimates are that over 35% of Merino, 30% of Border Leicester and 70 % of terminal breed sheep will have direct links to sires used in the IN. The link with SGP will facilitate verification of individual gene markers and development of enhanced ASBVs with use of molecular technologies (Hynd 2006).

CONCLUSIONS

The IN will impact on genetic gain in the Australian sheep industry by –

- developing an efficient system for collection and processing of phenotypic data and combining this with molecular genetic information for new and enhanced ASBVs
- modeling the new system on the industry breed structure with a significant genetic footprint and effective linkage with Sheep Genetics
- ensuring effective delivery of new genetic tools through stud breeder and wider sheep industry engagement

ACKNOWLEDGMENTS

The IN is supported by the Australian Government's Cooperative Research Centres Program, MLA and AWI. Considerable resources are provided at each site by NSW DPI, UNE, Vic DPI, SARDI and DAFWA. Valuable participation by SG and SGP is integral to the system. Development and maintenance of the IN database is supported by AGBU.

REFERENCES

Banks, R.G., van der Werf, J.H.J., and Gibson, J.P. (2006) *Proc. 8th Wld Congr. Genet. Appld. Livest . Prod.* Belo Horizonte, Brazil, August, CD-ROM Communication 30-12.

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

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

Hynd, P.I. (2006) *Proc. 8th Wld Congr. Genet. Appld. Livest . Prod.* Belo Horizonte, Brazil, August. CD-ROM Communication 05-06.

Oddy, V.H., Longhurst, T.J., Nicholas, F.W., Maddox, J.F. and McDonagh, M.B. (2005) *Proc. Assoc.Advmt. Anim. Breed. Genet*. **16:**209.