Genomic selection in the Australian sheep industry

Andrew A Swan

Animal Genetics and Breeding Unit (AGBU), University of New England, Armidale, NSW, 2351, AGBU is a joint venture between NSW Department of Primary Industries and the University of New England and Cooperative Research Centre for Sheep Industry Innovation, Armidale, NSW, 2351

What is genomic selection?

Many livestock industries around the world have invested heavily in research and development of genomic selection since it was first proposed by Meuwissen *et al*. in 2001. The basic idea is that if genomic tests with large numbers of markers are available, they can be used to genotype a "reference" population of animals which are also measured for a wide range of economically important traits. Marker effects on these traits can be estimated, establishing a genomic prediction which can then be applied to industry animals outside the reference population. In other words, once genomic prediction equations are developed from an appropriate resource flock, industry animals can be genotyped and a genomic prediction of their genetic merit can be made without the need to measure them.

This use of genomic information allows more accurate selection at an earlier age and for traits which are hard to measure, therefore increasing selection response primarily by shortening generation interval. The first livestock species to capitalise on genomic selection in a significant way has been the dairy industry. In the past dairy sire selection has relied on expensive and lengthy progeny testing programs, with sires not achieving high accuracy estimated breeding values until their daughters have recorded lactations. Genomic predictions are now available on young bulls virtually from birth which are approaching the accuracies of progeny test breeding values (around 75%). With an appropriately designed breeding program to capitalise on this information it is possible to double current rates of genetic gain (Pryce and Daetwyler, 2011).

Successful application of genomic selection requires a number of key ingredients:

- availability of genomic tests with large numbers of markers
- a reference population in which animals are genotyped and measured for all economically important traits
- independent validation populations (also with measurements) in which the accuracy of genomic predictions can be established
- the ability to incorporate genomic predictions into existing genetic evaluation systems together with pedigree and performance data
- new breeding program designs which can capture the benefits of more accurate breeding values early in life.

This paper outlines the research and development needed to establish a platform for genomic selection, and how it has been achieved in the Australian sheep industry via a major program undertaken by the CRC for Sheep Industry Innovation and its partners.

Genomic testing

Genomic selection in the livestock industries has been made possible by the availability of high density SNP marker panels, commonly referred to as “SNP chips” (SNP is an abbreviation of Single Nucleotide Polymorphism). Many markers across the whole genome are needed so that all genes which affect a trait (known as QTL or Quantitative Trait Loci) have markers close by.

The first panel for cattle had 10K (10,000) SNP markers, but it was quickly recognised that this was insufficient to develop genomic predictions of sufficient accuracy, and 50K and 800K SNP panels are now routinely used.

For sheep a 50K SNP panel is currently used, with 10K and 800K SNP panels under development. The current cost of a 50K panel for a single animal is approximately $100AU.

Importantly, the markers on lower density panels are included on high density panels, and a statistical method known as imputation can be used to convert low density genotypes to high density with very high accuracy. Imputation has the potential to reduce the cost of genotyping in a breeding program, for example by genotyping all parents at high density and offspring at low density. Offspring could be imputed to high density using the parental information, and if selected, re-genotyped at high density.

Reference populations

Reference populations have been established in the Australian beef, dairy and sheep industries primarily with industry research funding. The goal of these populations is to develop genomic predictions with the highest possible accuracy, which depends on a number of factors:

- Higher density SNP panels lead to more accurate predictions
- Increasing the number of animals with genotypes and trait measurements increases accuracy
- Accuracy will be higher for traits with higher heritability
- Given the previous factors, accuracy will be higher for populations with smaller effective population size. Species and breeds with a long history of intense selection tend to have smaller effective populations sizes, so dairy cattle and pigs are smaller than beef cattle and sheep for example. Within the Australian sheep population, the Merino breed is more diverse and has a higher effective population size than breeds such as the Poll Dorset.

The bottom line is that large numbers of animals are usually required for reference populations to achieve worthwhile accuracies, and this is the major cost of genomic selection, typically meaning that significant levels of industry funding are required to support reference populations.

A further consideration is that genomic predictions for animals outside the reference population (i.e., selection candidates in breeding herds and flocks) are more accurate if they are more closely related to the reference population. This means that genomic predictions decay over time if the reference population is not maintained. Therefore, funding for an ongoing reference population is a key requirement.

A major achievement of the Sheep CRC has been the establishment of an Information Nucleus Flock (INF) at nine sites across Australia. Sires are used across all sites, with progeny genotyped and measured for a wide range of traits, including wool, meat, disease resistance, reproduction, and lamb survival. The INF encompasses the major breeds, including Merinos, Border Leicesters, and terminal sire breeds (Poll Dorsets, and White Suffolks). To date approximately 15,000 progeny have been genotyped, 45% Merino, 15% Border Leicester, and 40% terminal sire breeds. Measurements have been made on all animals, although no animals have the full set of traits: roughly half of all animals are slaughtered and measured for carcass traits, while only female Merinos and Border Leicester Merino crosses are measured for reproduction and adult wool production, for example. This highlights the difficulty in building an effective reference population: hard to measure traits are still hard to measure. Nevertheless, the INF is arguably one of the best reference populations of any livestock industry internationally.

Validation resources to establish accuracy

Genomic selection will only be worthwhile if the accuracy of genomic predictions is sufficiently high to warrant the increased investment in breeding programs. Consequently, determining accuracy for key traits is critical in a strategic sense for planning future investment and designing breeding programs. In addition, accuracies are required for practical purposes as inputs to genetic evaluation systems which combine genomic information with pedigree and performance information.

Accuracy can be calculated by internal cross validation within reference populations, but in the sheep industry independent validation resources have been established. These are comprised of sires with highly accurate estimated breeding values (EBV) from industry flocks (excluding the INF resource flocks). The sires are genotyped and genomic estimated breeding values (GEBV) are independently derived using the genomic prediction from the INF reference population. The accuracy can then be estimated as the correlation between EBV and GEBV.

Current accuracies for some of the key economic traits are shown in Table 1. Accuracies are low to moderate for Border Leicesters and terminal sires, and moderate to high in Merinos. For fleece weight and fibre diameter, these accuracies are similar to the accuracy of an animal’s own performance measurement. However, it is likely that some of this high accuracy is due to the influence of strain differences within the Merino breed. The accuracy relevant to within flock selection is likely to be lower.

Table 1. : Accuracies (%) of genomic selection in key economic traits for Merinos, Border Leicesters and terminal sire breeds

Trait	Merino	Border Leicester	Terminal sires
Body weight	37-66	39-45	20-34
Eye muscle depth	44	26	45
Fat depth	45	29	30
Worm resistance	27	6	35
Fleece weight	58-69		
Fibre diameter	77		
Staple strength	35		

Incorporating genomic information in genetic evaluation systems

In order to maximise the accuracy of EBVs genomic information should be integrated with existing pedigree and performance information. Currently this is done by combining GEBVs and existing EBVs derived from separate analyses using a selection index like approach. Ideally, all information should be combined in a single analysis, and this is an area being very actively researched at the moment.

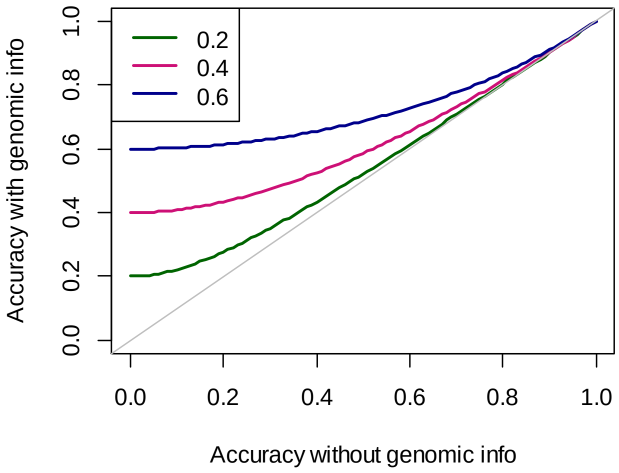

Figure 1 Accuracy of genomically enhanced EBVs (blended EBV) versus existing EBV for different levels of accuracy of genomic prediction (0.2, 0.4 and 0.6)

The additional accuracy due to inclusion of genomic information in EBVs can be seen in Figure 1, showing that the major increase in accuracy occurs at low levels of accuracy of the existing EBV. As more pedigree and performance data increases the accuracy of the existing EBV there is less gain from adding genomic information, and consequently there is little benefit in a breeder doing a genomic test on a progeny-tested sire. This reinforces the fact that genomic selection is most effective when combined with a change in breeding program design to select animals earlier.

Incorporation of genomic information into EBVs has been developed for the sheep industry in two pilot projects coordinated by the Sheep CRC. As part of these projects, young rams have been genotyped in industry flocks, and a pipeline has been developed encompassing DNA sampling and extraction, genotyping in a commercial lab, quality control of genotypes, and data analysis and reporting. Genomically enhanced EBVs are now being provided for these pilot rams as part of fortnightly routine genetic evaluations, and include up to 36 traits and selection indexes depending on breed.

Capturing the benefits of genomic selection in breeding programs

As has been discussed above, the main benefits of genomic selection come from selecting animals at younger ages with higher accuracy, increasing genetic progress by reducing generation interval. Van der Werf (2009) estimated increased annual responses of 30% in a terminal sire index and 40% in a Merino index. These benefits arose in the case of terminal sires from improved accuracy of carcass traits which are not normally measured, and in the case of Merinos from improved accuracy of adult wool traits which are normally measured only on breeding ewes after their first selection, if at all.

Further improvements in response are possible for Merinos by first mating at 6 months of age rather than the industry standard age, which is 18 months, but there is less scope to do this in terminal and maternal sire breeds where early mating is already common. The largest increases in response are likely when combining genomic selection with advanced reproductive technologies such as juvenile in vitro embryo transfer (JIVET).

Modelling to predict the benefits of genomic selection under different scenarios is an important area of active research for the sheep industry.

Conclusions

The Australian sheep industry has made substantial progress towards implementing genomic selection: an effective reference flock has been established which is producing genomic predictions with beneficial accuracies, and a pipeline has been developed which is providing breeders with genomically enhanced EBVS on a routine basis. There is also considerable interest from breeders in using genomic information within their breeding programs, and much research work remains to identify optimal strategies.

If the past is any indication, it is likely that future advances in DNA technology will increase the benefits of genomic selection. Lower cost genotyping will spread usage across the industry more widely, and higher density information (even to full DNA sequence on key ancestors) will increase the accuracy of genomic predictions, potentially including breeds which are currently not part of the reference population.

References

Meuwissen THE, Hayes BJ, Goddard ME (2001) “Prediction of total genetic value using genome-wide dense marker maps.” Genetics 157: 1819–1829.

Pryce, JE and Daetwyler, HD (2011) “Designing dairy cattle breeding schemes under genomic selection: a review of international research.” Animal Production Science 52: 107-114.

Van der Werf JHJ (2009) “Potential benefit of genomic selection in sheep.” Proceedings of the Association for the Advancement of Animal Breeding and Genetics. 18: 38–41.