Sheep CRC 2010 Conference Proceedings

Document ID:	SheepCRC_31_22
Title:	How robust are genomic selection methods?
Author:	S. A. Clark, J. M. Hickey and J. H. J. van der Werf
Key words:	sheep; genomic selection;

This Postgraduate paper was presented at the Sheep CRC Conference held in 2010, as part of the Sheep CRC presentations. The paper should be cited as:

S. A. Clark, J. M. Hickey and J. H. J. van der Werf (2010) - *How robust are genomic selection methods?*

How robust are genomic selection methods?

S. A. Clark[A,B,C], *J. M. Hickey*[B] *and J. H. J. van der Werf*[A,B]

[A]Australian Cooperative Research Centre for Sheep Industry Innovation, Homestead Building, University of New England, Armidale, NSW 2351, Australia.
[B]School of Environmental and Rural Science, University of New England, Armidale, NSW 2351, Australia.
[C]Corresponding author. Email: sclark9@une.edu.au

SUMMARY

Genomic information from many single nucleotide polymorphism (SNP) markers can be used to increase the accuracy of estimated breeding values of young animals. This is termed genomic selection (GS) and is based on prediction of the effects of quantitative trait loci (QTL) in linkage disequilibrium (LD) with markers (Meuwissen *et al*. 2001). However, Habier *et al.* (2007) proposed that GS also relies on "relationships" between individuals to accurately predict genetic value. A better understanding of what GS actually predicts is needed to develop marker panels, training populations and methods for accurately estimating breeding values. The efficacy of methods used to predict genomic breeding value may depend on the underlying model of genetic variation, which is not well known. The aim of this study was to determine the accuracy and robustness of various methods used for genomic selection using a range of underlying genetic models and to compare the accuracy of GS when predicting one generation ahead (training set 1), several generations ahead (training set 2) or across different populations (training set 3).

Three models of variation were used to simulate the genetic value of animals: (i) a QTL model in which few QTLs have a relatively large effect, (ii) a QTL model in which many QTLs have moderate effects and (iii) an infinitesimal model in which very many QTLs each have a very small effect. Genotype information from 60,000 markers was used to estimate the genetic value of animals using the following methods: (a) Bayes B, based on estimation of marker effects, (b) gBLUP, based on genomic relationships between animals and (c) traditional BLUP, based on pedigree relationships.

Table 1. Average correlation between estimated and true breeding values using 60,000 SNP markers and training sets one generation away (1), eight generations away (2) or from another sub-population (3)

Genetic Model of Variation	Training Set	Bayes B	gBLUP	BLUP
100 QTLs	1	0.83	0.56	0.46
	2	0.77	0.37	
	3	0.77	0.33	
1000 QTLs	1	0.65	0.59	0.47
	2	0.49	0.38	
	3	0.47	0.34	
Infinitesimal	1	0.39	0.40	0.45
	2	–0.01	0.00	
	3	0.00	–0.01	

The Bayes B method was the most accurate method for predicting breeding value when there were 100 QTLs and resulted in accuracies similar to those obtained with gBLUP when the infinitesimal model was used. Both genomic methods relied on both relationships and QTL effects to estimate breeding values. Large QTL effects enabled prediction of breeding value regardless of whether animals were related. However, with the infinitesimal model, prediction was based on relationships and none of the methods were able to predict breeding values when the animals were unrelated. We conclude that the Bayes B method is the superior method as it utilizes large QTL effects if they exist, resulting in accurate genomic EBVs, whereas it relies on relationships in the absence of large QTL effects, becoming equivalent to the BLUP method.

REFERENCES

Habier D, Fernando R and Dekkers J (2007) The Impact of Genetic Relationship Information on Genome-Assisted Breeding Values. *Genetics* **177**, 2389–2397.

Meuwissen T, Hayes B and Goddard M (2001) Prediction of total genetic value using genome-wide dense marker maps. *Genetics* **157**, 1819–1829.