

Using Genomics to Improve Reproduction Traits in Sheep

H.D. Daetwyler^{1,2,3}, S. Bolormaa^{2,3}, K.E. Kemper^{2,4}, D. Brown^{3,5}, A.A. Swan^{3,5}, J.H.J. van der Werf^{3,6}, and B.J. Hayes^{1,2,3}

¹La Trobe University, Bundoora, Australia, ²Department of Environment and Primary Industries, Bundoora, Australia, ³CRC for Sheep Industry Innovation, Armidale, Australia, ⁴University of Melbourne, Parkville, Australia, ⁵Animal Breeding and Genetics Unit, Armidale, Australia, ⁶University of New England, Armidale, Australia

ABSTRACT: Reproduction traits are important but difficult to improve in sheep because they are lowly heritable and are recorded later in life. Genomic information can be used for genomic prediction of breeding values to increase the accuracy of selection. Alternatively, genomic information can be used to identify recessive lethal mutations, which cause embryonic losses. We investigated both uses of genomic information using 54k SNP in the Border Leicester, Merino and Polled Dorset breeds. Genomic prediction accuracy of three reproduction traits was greater when compared to pedigree methods, especially in less related animals. Furthermore, making use of both sire and ewe information in the reference set increased accuracies. Ten haplotypes carrying potential recessive lethal mutations were identified. One haplotype spanning 50 SNP alleles was significantly associated with litter size. The findings suggest genomic tools should be used to increase reproductive efficiency in sheep.

Keywords: sheep; genomic selection; reproduction; recessive lethal mutation

Introduction

Reproduction traits, such the number of lambs born and weaned, are important drivers of profitability in sheep. However, these traits are difficult to improve because they have low heritability and are recorded later in life. Genomic technologies, such as single nucleotide polymorphism (SNP) arrays, could improve genetic gain for these traits through genomic selection (Meuwissen *et al.* 2001). Genomic information can also be used to find regions that harbor recessive lethal mutations, which decrease reproductive success (e.g. VanRaden *et al.* 2011).

Genomic prediction equations are estimated in a reference population of genotyped and phenotyped individuals, and then used to predict breeding values for animals with only genotypes. This can enable the prediction of a selection candidate's performance early in life potentially with greater accuracy than estimated breeding values based on pedigree alone, provided the reference population is large enough. One straightforward way to design a reference population is to use genotyped and phenotyped ewes. Additionally, genotyped sires with many daughter reproduction records could be included. This second data type is attractive for genomic prediction, as the sire 'phenotypes' are more accurate than ewe phenotypes, delivering a greater increase in accuracy per sire genotyped. Combining ewe and sire information, thereby making optimal use of all data, in a reference population is expected to yield higher accuracy predictions.

However, combined reference populations require additional considerations during analysis and validation (e.g. Calus *et al.* 2013; Daetwyler *et al.* 2013).

Lethal recessive mutations have been shown to affect reproduction traits and are expected to arise spontaneously in all populations by chance (e.g. Frankham 1996; VanRaden *et al.* 2011). Genetic drift and inbreeding may increase their frequency and these factors are especially pronounced when the effective population size (N_e) is small. Lethal mutations may also occur by chance within a haplotype that increases production and, in such cases, artificial selection may further increase their frequency. Large datasets have been built for genomic selection analyses, further mining of this genomic data for lethal recessives that never occur in their homozygous state can provide additional tools to improve reproduction traits in sheep.

This paper has two main aims. First, we show that a combined reference population composed of genotyped ewes with phenotypes and genotyped sires with daughter phenotypes can increase the accuracy of genomic prediction. Second, we search for lethal recessive mutations inherited within haplotypes in the Australian sheep flock.

Materials and Methods

Genomic Selection

Phenotypes, Genotypes, Heritability and Repeatability. Three traits were investigated: number of lambs weaned per ewe joined (NLW), number of lambs born per ewe joined (NLB), and litter size (number of lambs born per ewe lambing, LSIZE). The initial dataset consisted of 3984 genotyped ewes with reproduction phenotypes from the CRC for Sheep Industry Innovation (CRC) Information Nucleus (van der Werf *et al.* 2010) and SheepGENOMICS (SG) (White *et al.* 2012). This data was augmented with 524 genotyped sires with derived phenotypes based on their daughter phenotypes (daughter trait deviations, DTD). Derived phenotypes were calculated with 367,393 records from 202,898 animals from Sheep Genetics database, which provides the national genetic evaluation for sheep in Australia. The ewes used to derive the phenotypic data used in the analyses (described below) overall consisted of 74% of Merino (MER) and 12% of Border Leicester (BL), and less than ~5% of each of the other breeds. Hence, most were pure MER, but first cross MER-BL were also prevalent within CRC.

All available phenotype data was fitted in single trait BLUP mixed models to estimate heritability (h^2) and repeatability (t) for each trait (Henderson 1984). These parameters were needed to calculate weights to derived phenotypes (trait deviations) when combining sire and ewe records in genomic best linear unbiased prediction (GBLUP), described below. The model applied was: $\mathbf{y} = \mathbf{X}\mathbf{b} + \mathbf{Z}_1\mathbf{Q}\mathbf{q} + \mathbf{Z}_1\mathbf{a} + \mathbf{Z}_2\mathbf{pe} + \mathbf{e}$, where \mathbf{y} is a vector of phenotypes, \mathbf{X} , \mathbf{Z}_1 , and \mathbf{Z}_2 are design matrices, \mathbf{Q} is a matrix that accounts for breeds and strains, \mathbf{b} is vector of fixed effects, \mathbf{q} is a vector of breed effects, \mathbf{a} is a vector of animal genetic effects, \mathbf{pe} is a vector of permanent environmental effects, and \mathbf{e} is the vector of random errors.. The following distributions were assumed: $\mathbf{a} \sim N(0, \sigma_a^2 \mathbf{A})$, $\mathbf{q} \sim N(0, \sigma_q^2 \mathbf{I})$, $\mathbf{pe} \sim N(0, \sigma_{pe}^2 \mathbf{I})$, and $\mathbf{e} \sim N(0, \sigma_e^2 \mathbf{I})$, where \mathbf{A} is the numerator relationship matrix, σ_a^2 is the genetic variance, σ_q^2 is the variance of breed effects, σ_{pe}^2 is the permanent environment variance and σ_e^2 is the residual variance. Fixed effects included the mean, conception site, lambing site, year of lambing, birth year of ewe, age at lambing, and conception method. The conception methods used for mating were natural, artificial insemination, and hand mating. The ewes were born from year 1990 to 2011 and lambed in years 1991 to 2012. There were 3 different classes of age of lambing: <1.5; 1.5-2.5; and >2.5. Only 24,783 ewes lambed at age less than one year old.

The genotypes were provided by the CRC and SG projects. All animals were genotyped using the Ovine50K SNP chip comprising a total of 54,977 single nucleotide polymorphisms (SNP). The genotype quality control and imputation of sporadic missing genotypes is described in Daetwyler *et al.* (2012). The final genotyped dataset for genomic prediction included 4,508 animals (524 sires) genotyped at 48,599 SNP. Results for genomic predictions are based on 4508 animals, and are reported separately for the MER and BL breeds and their crosses, as these two breeds made up the majority of genotyped animals (Table 1).

Table 1. Size of reference population (TD + DTD) within Border Leicester (BL) and Merino (MER) breeds and in total across all breeds (ALL).

Trait	BL		MER		ALL
	TD	DTD	TD	DTD	
NLW	644	93	2,404	177	4,484
NLB	649	93	2,405	189	4,508
LSIZE	629	89	2,394	184	4,458

Trait and Daughter Trait Deviations. Trait deviations (TD) for genotyped ewes and daughter trait deviations (DTD) for genotyped sires were used as the derived phenotypes in GBLUP analyses. Firstly, the phenotype was corrected for fixed effects using the same

model applied to calculate heritability and repeatability, but excluding the animal genetic effect. Then the residuals from this model were used to calculate ewe TD as $\Sigma(\text{residual}_j)/N_j$, where N_j is the number of records for the ewe j , and sire DTD as $\Sigma(\text{TD}_i)/p_i$, where TD_i are the trait deviations of the ungenotyped daughters of sire i , and p_i is the number of daughters. Genotyped daughters of these sires were included in reference population and not included in DTDs to avoid double counting. Sires with less than 3 progeny were excluded from the analysis. The DTD contain only half the genetic merit of the sire, thus DTD was doubled. To account for heterogeneity of variance the TD and DTD were weighted as in Garrick *et al.* (2009). The weighting requires an assumption on the proportion of the genetic variance not captured by the markers (c). We assumed five different values for c (0.25, 0.35, 0.50, 0.60, and 0.75).

Genomic Prediction Analysis. Two methods were used to estimate breeding values: genomic BLUP or GBLUP for Genomic EBVs (GEBV) and BLUP for EBVs. GEBVs were calculated based on the following model: $\mathbf{y} = \mathbf{1}\mu + \mathbf{X}\mathbf{b} + \mathbf{Z}\mathbf{g} + \mathbf{e}$, where \mathbf{y} is a vector of TD and DTD, $\mathbf{1}$ is a vector of ones, μ is the mean, \mathbf{X} and \mathbf{Z} are design matrices, \mathbf{b} is a fixed sex effect, \mathbf{e} is the vector of random errors distributed as $N(0, \mathbf{R}\sigma_e^2)$ where \mathbf{R} is a diagonal matrix with weights as described above and σ_e^2 is the residual variance; \mathbf{g} is a vector of either GEBV or EBV distributed as $N(0, \mathbf{G}\sigma_g^2)$, where σ_g^2 is either genetic variance explained by the markers or the additive genetic variance, and \mathbf{G} is either the genomic relationship matrix (GRM) or the numerator relationship matrix among animals (NRM) (Yang *et al.* 2010). Breed effects were not fitted because they were fitted during calculation of TDs and DTDs. All estimates were performed using ASReml software (Gilmour *et al.*, 2009).

Estimation of Accuracy Through Cross-Validation. Prediction analyses using both GBLUP and BLUP used all genotyped individuals from all breeds that had either their own records or ungenotyped daughters with records in the reference population. 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. The six subsets for cross-validations were chosen either completely at random or by at random stratified by sire family. Sire families based on TD or DTDs were not forced to be equally distributed in subsets. In random sire family cross-validation, sires were randomly allocated to subsets, and all genotyped progeny of a sire were then allocated to the same subset to ensure prediction was across sire families, resulting in a conservative estimate of prediction accuracy. The cross-validation subsets were the same for GBLUP and BLUP. All subsets consisted of multiple breeds. For all GBLUP and BLUP analyses, the validation animals were included in the GRM or NRM but had unknown phenotypes in the calculation of GEBV or EBV.

The correlation between the GEBV and the true breeding values (TBV) is known as accuracy of genomic

prediction. In practice, the TBV are unknown, and the only data available are phenotypes, which are made up of the true breeding value and the environmental effect. The accuracies of genomic prediction were calculated as a Pearson correlation between GEBV and corrected phenotypes (TD + DTD) within the MER and BL breeds, where animals were assigned to breed groups according to the breed of the sire. The upper bound of this correlation is equal to the accuracy of the breeding value. The correlation of GEBV and TBV were approximated by dividing by the accuracy of the TD and DTD within each breed in a subset. Ewe TD accuracies were approximated from BLUP models and sire DTD accuracies were approximated using their BLUP EBV accuracy from only non-genotyped daughters. They were calculated as $r = \sqrt{1 - PEV / \sigma_a^2}$, where r is accuracy of the BLUP solution, PEV is the prediction error variance and σ_a^2 is the additive genetic variance. Correlations from each of the six subsets were averaged within breed and weighted by the number of individuals in the validation set.

Detection of Potential Lethal Recessive Mutations.

Genotypes and Haplotyping. 23907 sheep from many breeds, but mainly Merino, Border Leicester, Polled Dorset, and White Suffolk, were genotyped within the CRC and SG projects (van der Werf *et al.* 2010; White *et al.* 2012). Quality control and imputation was performed, as above, resulting in a 48599 SNP. The genotypes were phased jointly in all breeds using ChromoPhase (Daetwyler *et al.* 2011) to ensure more consistent long-range haplotypes. All unique haplotypes across all breeds were identified by comparing each animal's haplotypes at a position to all other haplotypes. Haplotypes were numbered in order of observation. Phasing was not perfect and included some missing or unphased loci. Haplotypes with missing alleles were considered unique and were not merged with other potentially consistent haplotypes to be conservative. The following haplotype lengths were considered 5, 10, 20, 50, and 100 SNP corresponding to approximately 0.25, 0.5, 1.0, 2.5, and 5.0cM, respectively.

The breed proportion matrix (**Q**) was then used to identify pure individuals for the main four breeds in the dataset. Individuals needed to reach a 95% purity threshold, as determined from pedigree analysis. This resulted in 9517 Merino, 767 Polled Dorset and 331 Border Leicester individuals. There were not enough pure White Suffolk individuals for meaningful analyses. Haplotype frequencies in crosses were ignored as one would not expect to detect deleterious haplotypes due to high haplotype heterozygosity. The expected number of homozygote haplotypes was calculated assuming Hard-Weinberg equilibrium by multiplying the frequency of the homozygote haplotype by the number of individuals in the breed.

Haplotypes that were expected to be observed in a homozygous state at least 6 times based on their frequency in the breed, but were never observed, were further investigated. First, carrier sires and their offspring were

checked for Mendelian inheritance ratios. The expected frequency of haplotype homozygosity in purebred offspring of carriers equals the probability of inheriting the haplotype from the carrier sire (0.25) times the probability of inheriting it from the dam, which is half the carrier frequency in the population. Second, the effect of the carrier status on three female reproduction traits (LSIZE, NLB, and NLW) was investigated using the TD and DTDs of carriers and non-carriers in the following model: $\mathbf{y} = \mathbf{1}\mu + \mathbf{X}\mathbf{b} + \mathbf{e}$, where \mathbf{y} is a vector of TD and DTD, $\mathbf{1}$ is a vector of ones, μ is the mean, \mathbf{X} is a design matrix, \mathbf{b} is a vector of fixed effects including sex and potential deleterious haplotype carrier status (1 for carrier, 0 otherwise). Heterogeneous accuracy of TDs and DTDs was accounted for as in genomic prediction analyses above.

Results and Discussion

Genomic Selection.

Heritability and repeatability. The number of lambs born varied between 0 to 4 (5 categories) per ewe joined, with very few quadruplets. The data for the reproduction traits were approximately normally distributed. Heritabilities and repeatabilities estimated from the full BLUP model were between 0.06 and 0.08 and between 0.08 and 0.10, respectively. Estimated heritabilities were within the range of literature estimates, while repeatabilities were slightly lower (e.g. Safari *et al.* 2005). The permanent environmental effect (pe) explained on average 2% of the total variance. The standard errors of h^2 and the proportion of the permanent environmental effect were small (range 0.0027 to 0.0035).

Genomic Prediction Accuracy. This is one of the first genomic prediction studies for reproduction traits in sheep. The accuracy of GEBVs was generally low (Table 2), reflecting the low heritability of reproduction traits and relatively low number of animals per breed considering the multi-breed nature of our reference population. Overall, the prediction accuracy of cross-validation random splits was much higher than across sire family splits in both BLUP and GBLUP, reflecting the greater relationships between reference and validation animals due to prediction within families (GBLUP accuracy random 0.23, sire family 0.11, Table 2). There was a trend for higher accuracy of GBLUP over BLUP, except in BL for LSIZE. The increase in accuracy seen with GBLUP was more pronounced in sire family splits, where the average GBLUP accuracy was double that of BLUP. In random splits the increase from GBLUP was very modest, but was consistent for NLW and NLB.

This study has shown that combining sire and ewe records resulted in higher accuracy in both GBLUP and BLUP (Table 2). The increase in accuracy from this additional information was greater when animals were less related to the reference population (e.g. across sires family cross-validation). More generally, the increase in GBLUP over BLUP, while fairly consistent, was also greater when animals were less related, confirming other studies (Clark *et al.* 2012).

Table 2. Average weighted accuracies of GEBV from cross-validation using BLUP and GBLUP, when using only TDs of ewes or TD plus DTDs of sires.

Trait	Breed	BLUP		GBLUP	
		TD+ DTD	TD	TD+ DTD	TD
<i>sire family</i>					
NLW	MER	0.04	-0.09	0.12	0.09
NLB	MER	0.03	-0.19	0.06	0.11
LSIZE	MER	0.00	-0.16	0.14	0.09
<i>Mean</i>	<i>MER</i>	<i>0.02</i>	<i>-0.15</i>	<i>0.11</i>	<i>0.10</i>
NLW	BL	0.11	-0.04	0.10	0.07
NLB	BL	0.12	-0.01	0.15	0.11
LSIZE	BL	0.11	0.06	0.07	-0.02
<i>Mean</i>	<i>BL</i>	<i>0.11</i>	<i>0.00</i>	<i>0.11</i>	<i>0.05</i>
Mean	All	0.07	-0.07	0.11	0.08
<i>random</i>					
NLW	MER	0.19	0.17	0.21	0.20
NLB	MER	0.31	0.32	0.32	0.33
LSIZE	MER	0.39	0.38	0.40	0.39
<i>Mean</i>	<i>MER</i>	<i>0.30</i>	<i>0.29</i>	<i>0.31</i>	<i>0.31</i>
NLW	BL	0.11	0.06	0.14	0.10
NLB	BL	0.19	0.18	0.20	0.21
LSIZE	BL	0.14	0.14	0.12	0.10
<i>Mean</i>	<i>BL</i>	<i>0.15</i>	<i>0.13</i>	<i>0.15</i>	<i>0.14</i>
Mean	All	0.22	0.21	0.23	0.22

In the sire family split, GBLUP achieved an increase of 0.12 over BLUP across the three reproduction traits in MER, whereas in BL a modest increase was observed only for NLW and NLB. Genomic prediction accuracies can be compared with theoretical predictions (Daetwyler *et al.* 2010b) based on the number of individuals, genome length ($L = 26$ Morgan, effective population size ($N_e = 853$ for MER, $N_e = 243$ for BL, (Kijas *et al.* 2012), and number of independent chromosome segments ($M_e = 2N_eL$), which would give an expected accuracy of approximately 0.14 for MER and 0.11 for BL. Therefore, the achieved accuracies are consistent with expectations.

The accuracy of across breed prediction has been shown to be limited in sheep with the 50k SNP chip (Daetwyler *et al.* 2010a). However, the use of a common reference population or, equivalently a common prediction equation based on SNP genotypes, for all breeds significantly simplifies analyses in multi-breed cross-bred data. An increase in marker density may allow for across breed prediction to play a greater role in the future. In beef cattle using the 800k Bovine chip, Bolormaa *et al.* (2013) found that using a common training population lead to higher accuracies than using breed specific training populations, especially for crossbreeds. A multi-breed

reference population is only advantageous if the LD phase between SNP and QTL are consistent across breeds. This is only expected if the chromosome segments containing the SNP and QTL in different breeds have descended from a common ancestor without recombination. As the common ancestor would have occurred many generations ago the shared segments across breeds are likely to be very short and therefore the SNP and the QTL would need to be closely linked and dense to be effective at tagging QTL

Another way to increase the accuracy of genomic prediction would be to increase the reference population size. Genotyping highly influential sires with many daughters is a more cost effective strategy to increase the sample size, than genotyping all ewes. The approach shown in this study enables the incorporation of sire data into genomic prediction analyses.

Using the current approach, the GEBV would have to be combined with BLUP breeding values to be distributed to sheep breeders. In Australian sheep, this has been accomplished with a selection index approach for traits which have a BLUP evaluation (the so called “blending” method) (Harris & Johnson 2010; Swan *et al.* 2012). The use of sire DTDs for genomic predictions further complicates this process and additional steps are necessary to reduce double counting. GEBVs for traits that do not have a BLUP EBV are currently delivered using a one-step approach which augments the NRM inverse matrix used in the BLUP equations with the GRM inverse for genotyped animals (Misztal *et al.* 2009; Aguilar *et al.* 2010; Swan *et al.* 2012). This would eliminate the post processing steps of the blending approaches, as all breeding values would be on the same scale and genomic information would be accounted for implicitly. However, one-step approaches do not avoid all issues related to blending, because they also require an assumption on a scale parameter (λ) describing how much of the genetic variance is captured by the GRM. The value of λ , just like c , is uncertain and would have to be verified by cross-validation.

The accuracies achieved using GBLUP, while low, are encouraging and more genetic gain would be achieved for reproduction traits through genomic prediction than with BLUP. This is especially true for animals that are less related to the reference population. The accuracies shown are conservative estimates of accuracies that could be achieved in commercial breeding programs as they are calculated for animals without direct relationships to the reference population. Making use of all data on both ewes and sires increased the accuracy of prediction.

Deleterious Recessive Haplotypes.

Ten haplotypes carrying potentially lethal mutations were detected across the MER, PD and BL breeds (Table 3). All of these haplotypes were expected as homozygotes in the population at least six times but were never observed as homozygotes. The p-value for not observing a homozygote when expecting 6 is 0.0025, assuming a poisson distribution. The number, length and

frequency of these haplotypes reflected the effective population size (N_e) of the respective breeds. The greatest number of deleterious haplotypes (5) were found in the BL breed across all haplotype lengths. This was expected as it also has the smallest N_e and therefore the highest inbreeding of the three breeds in the study. The deleterious haplotype carrier frequencies in BL were high at approximately 0.13. In three instances, haplotypes of different lengths identified the same genomic location to harbour a recessive lethal haplotype in BL. Haplotypes expected five times but never observed were only reported if they fell in the same regions as haplotype meeting the more stringent threshold (Table 3). This provides further confidence that signals are true.

Table 3. Potential lethal recessive haplotype in Border Leicester (BL), Polled Dorset (PD) and Merino (MER) breeds, where Name contains breed.chromosome.length.location.haplotype, Loc is the location on chromosome in megabases, Freq is the frequency of the haplotype in the population, Obs is the number of observed homozygote haplotypes, Exp is the expected number of homozygote haplotypes and HapAlleles are the alleles within the haplotype.

Name	Loc	Freq	Obs	Exp	HapAlleles
BL.2.20.210.18	215	0.15	0	7	2020202000 2222002020
BL.2.50.84.19	214	0.14	0	6	0002022222 2200222200 2022200000 2020202000 2222002020
BL.4.5.345.13	87	0.14	0	6	20222
BL.4.10.172.29	87	0.14	0	6	0020020222
BL.4.20.86.37	87	0.13	0	5	0200000002 0020020222
BL.8.5.289.3	74	0.14	0	6	00202
BL.8.10.144.14	74	0.13	0	5	0002000202
BL.15.5.143.3	39	0.14	0	6	02222
BL.24.50.5.14	15	0.14	0	6	2220200200 2022000002 2000022022 2222202020 00002002202
PD.13.50.14.10	38	0.11	0	8	0202200020 2002202200 2202222000 0200220200 20220022202
PD.18.5.29.23	8	0.09	0	6	02202
MER.2.5.222.3	59	0.03	0	6	22202
MER.4.10.240.9	122	0.03	0	8	0202220222
MER.9.5.359.2	91	0.03	0	6	02222

Two recessive lethal haplotypes that reached carrier frequencies of 0.10 were detected in the PD breed. In MER, 3 haplotypes were found to possess recessive lethal patterns. As these breeds have a higher N_e than BL, it was expected that fewer deleterious haplotypes would be detected.

The length of the haplotype that is found to be potentially lethal may reveal information about the age of the deleterious mutation. Mutations are expected to get

older as haplotypes shorten. Applying the approximation of Hayes et al. (2003) ($1/2(\text{length of segment in Morgans})$), we can calculate that a 100 SNP haplotype could harbour new recessive mutations that arose 10 generations ago, whereas a 5 SNP haplotype would be much older at 200 generations ago. The length of the lethal haplotypes varied within and across breeds (Table 3). No lethal 100 SNP haplotypes were detected. There was a trend for lethal haplotypes in the BL and PD breeds to be longer than in the MER breed.

Table 4. Potential lethal recessive haplotype in Border Leicester (BL), Polled Dorset (PD) and Merino (MER) breeds, and their observed (ObsFreq) and expected (ExpFreq) frequency in offspring of carrier sires, as well as most significant P-value in three litter size, number of lambs born and number of lambs weaned. Name contains breed.chromosome.length.location.haplotype.

Name	ObsFreq	ExpFreq	Pvalue
BL.2.20.210.18	0.25	0.32	0.023*
BL.2.50.84.19	0.21	0.32	0.009*
BL.4.5.345.13	0.30	0.32	0.425#
BL.4.10.172.29	0.29	0.32	0.62#
BL.4.20.86.37	-	0.31	-
BL.8.5.289.3	0.21	0.32	0.359*
BL.8.10.144.14	-	0.31	-
BL.15.5.143.3	0.24	0.32	0.842#
BL.24.50.5.14	0.19	0.32	0.079*
PD.13.50.14.10	0.21	0.30	0.096&
PD.18.5.29.23	0.26	0.30	0.578*
MER.2.5.222.3	0.21	0.26	0.151#
MER.4.10.240.9	0.15	0.27	0.123&
MER.9.5.359.2	0.26	0.26	0.087#

#litter size, &number of lambs born, *number of lambs weaned

All identified haplotype were investigated to see whether they follow Mendelian inheritance in carrier sire families. The frequencies of potential lethal haplotypes in carrier families was always lower than expected (Table 4). This is likely due the fact that the sires' offspring were often crosses and the lethal haplotype was not segregating in their dams. Futhermore, the haplotype carrier status was fitted as a fixed effect in a model that used TD and DTD as reproduction phenotypes. A subset of all purebred individuals also had reproduction information, leaving 2516 MER, 89 PD and 114 BL individuals for this analysis. A significant effect ($Pvalue < 0.01$) of carrier status was only observed for BL.2.50.84.19 in LSIZE of the BL breed (p value 0.009, Table 4). This indicates that there was likely not enough power to detect these effects in PD and BL. In MER, where power should have been less of an issue, the non-significant effects indicate that further investigation may be needed to adequately confirm these potential

recessive lethal haplotypes. A potential complication in MER is the presence of three different strains based on wool fibre diameter. While the groups are somewhat loosely defined, there is only limited genetic connectedness across the strains. A deleterious mutation may only occur in one strain, but it may be frequent enough to be chosen based on observed-expected ratios. Subsequent testing of haplotype effects on carrier reproduction would not be expected to yield significant results, as the deleterious mutation is not present consistently in all strains.

Haplotypes were not found to be deleterious in multiple breeds. The *a priori* expectation is that deleterious lethal mutations in a population are lost quickly because carriers have a large reproductive disadvantage. While possible, it is unlikely that the same deleterious mutations that predate breed divergence would still be present in several breeds. The cause of the rather high frequency of these haplotypes within breeds warrants further investigation. It may be that the same haplotypes also have positive effects on currently desired production traits and that artificial selection has increased their frequency. Such cases have been confirmed in Holstein cattle (VanRaden *et al.* 2011). The frequency of inheriting the favourable and lethal mutations in the same haplotype is closely related to genomic distance between the two mutations. If they are distant, the relationship of the two mutations will soon be broken by recombination. However, if they are close genomically, they may be inherited in same haplotype for many generations. This is a further argument that one would not expect longer haplotype to consistently carry the same lethal mutation across breeds. The effect of the lethal haplotypes on production traits will be investigated.

The existence of rather frequent lethal haplotypes in the BL and PD breeds is expected to require management of mating programs to minimise the risk of carrier-carrier matings. The risk is greatest for purebred breeding stock, which is used to produce rams for the multiplier and commercial sectors. The identification of carrier animals with genomics would allow breeders to avoid carrier-carrier matings. For untested animals, the probability that they are carriers could be calculated based on their genomically tested relatives. The risk to the commercial sheep sector which utilises crosses is small, as the lethal alleles will not be expected to occur in the homozygous state.

Conclusions

Genomics can be applied to increase reproduction traits in sheep in several ways. We have outlined two approaches: genomic prediction and identification of potential lethal recessive mutations. Genomic prediction, while its accuracy currently is still low for reproduction traits, does offer an advantage over pedigree selection methods, especially in less related animals. We have shown that making use of both sire and ewe information increases the accuracy of genomic prediction. The identification of recessive lethal mutations, which are inherited within longer haplotypes, is an additional tool to increase reproductive efficiency of purebred sheep.

Together, these genomic technologies are effective tools to make Australian sheep more productive.

Literature Cited

- Aguilar I., Misztal I., Johnson D.L., et al. (2010) *J. Dairy Sci.* **93**: 743-52.
- Bolormaa S., Pryce JE, Kemper K, et al. (2013) *J. Anim. Sci.* **91**: 3088-3104.
- Calus M.P.L., de Haas Y. & Veerkamp R.F. (2013) *J. Dairy Sci.* **96**: 6703-15.
- Clark S.A., Hickey J.M., Daetwyler H.D. et al. (2012) *Genet. Sel. Evol.* **44**: 4.
- Daetwyler H.D., Bolormaa S., Brown D.J., et al. (2013) *Proc. of AAABG, Napier, NZ.*
- Daetwyler H.D., Hickey J.M., Henshall J.M., et al. (2010a). *Anim. Prod. Sci.* **50**: 1004-10.
- Daetwyler H.D., Pong-Wong R., Villanueva B. et al. (2010b) *Genetics* **185**: 1021-31.
- Daetwyler H.D., Swan A.A., van der Werf J.H.J. et al. (2012) *Genet. Sel. Evol.* **44**: 33.
- Daetwyler H.D., Wiggans G.R., Hayes B.J., et al. (2011) *Genetics* **189**: 317-27.
- Frankham R. (1996) *Conser. Biol.* **10**: 1500-8.
- Garrick D., Taylor J. & Fernando R. (2009) *Genet. Sel. Evol.* **41**: 55.
- Harris B.L. & Johnson D.L. (2010) *J. Dairy Sci.* **93**: 1243-52.
- Hayes B.J., Visscher P.M., McPartlan H.C. et al. (2003) *Genome Res.* **13**: 635-43.
- Henderson C.R. (1984) *Applications of linear model in animal breeding.* University of Guelph, Guelph.
- Kijas J.W., Lenstra J.A., Hayes B., et al. (2012) *PLoS Biol* **10**: e1001258.
- Meuwissen T.H.E., Hayes B.J. & Goddard M.E. (2001) *Genetics* **157**: 1819-29.
- Misztal I., Legarra A. & Aguilar I. (2009) *J. Dairy Sci.* **92**: 4648-55.
- Safari E., Fogarty N.M. & Gilmour A.R. (2005) *Livest. Prod. Sci.* **92**: 271-89.
- Swan A.A., Johnston D.J., Brown D.J., et al. (2012). *Anim. Prod. Sci.* **52**: 126-32.
- van der Werf J.H.J., Kinghorn B.P. & Banks R.G. (2010) *Anim. Prod. Sci.* **50**: 998-1003.
- VanRaden P.M., Olson K.M., Null D.J. et al. (2011) *J. Dairy Sci.* **94**: 6153-61.
- White J.D., Allingham P.G., Gorman C.M., et al. (2012) *Anim. Prod. Sci.* **52**: 157-71.
- Yang J., Benyamin B., McEvoy B.P., et al. (2010) *Nat Genet* **42**: 565-9.