Contributed paper

THE EFFECT OF GDF9 ON LITTER SIZE IN AUSTRALIAN SHEEP

N. Moghaddar[1], K. L. Bunter[2], A. A. Swan[2] L. Li[2], P. M. Gurman[2] and J. H. J. van der Werf[1]

[1] School of Environmental & Rural Science, University of New England, Armidale, NSW 2351, Australia

[2] Animal Genetics Breeding Unit, University of New England, Armidale, NSW, 2351 Australia

SUMMARY

Growth differentiation factor 9 (*GDF9*) is a known autosomal gene which regulates ovulation rate in mammals. In sheep, numerous polymorphisms have been reported in coding regions of *GDF9* with a significant impact on ovulation rate and hence litter size. To study the effect of *GDF9* on litter size in Australian sheep breeds, an association analysis was performed between 1,600,633 imputed sequence single nucleotide polymorphisms (SNPs) on OAR5 and litter size phenotypes in 8,850 Merino and 7,613 maternal sheep breed ewes (predominantly Border Leicester, Coopworth, Corriedale and composite maternal lines) respectively. Results showed a significant association between litter size and SNPs in the *GDF9* region in maternal breeds. After filtering for high linkage disequilibrium, a highly significant SNP (*p_value* = 9.09E-09) was found in an intron of the *GDF9* gene at OAR5:41841588, which accounted for a 0.22 increase in litter size and explained 4.75% of the total genetic variance. This SNP and the surrounding SNPs in the region of *GDF9* were not significantly associated with litter size in Merinos. Information on this SNP genotype could be useful for obtaining a more accurate estimate of genetic merit for reproduction traits in some breeds of sheep.

INTRODUCTION

Sustainable livestock farming requires a constant increase in productivity and profitability of the enterprise. Reproductive performance is among the economically important traits in sheep breeding objectives as it directly affects the profitability of the enterprise. In sheep populations, the rate of genetic improvement in reproduction traits with conventional selection, based on breeding values derived from phenotypes and pedigree information, can be low. This is mainly due to low heritability of reproduction traits, incomplete recording in the industry, phenotypes being sex-limited and only available later in life, in particular for adult ewe performance.

Application of genomic information in breeding programs, including using information about polymorphisms affecting the genetic variation of a trait, can increase the accuracy of estimated breeding values (Moghaddar *et al.* 2019) and potentially can lead to significant improvements in genetic gain for reproduction traits. *GDF9* is a known autosomal gene with a significant impact on fertility traits in different mammals, including some sheep breeds. Literature shows some polymorphisms in *GDF9* are responsible for increased ovulation rate and higher litter size in both heterozygous and homozygous genotypes (e.g. Hanrahan *et al.* 2004; Silvia *et al*. 2011; Våge *et al*. 2013). However, sterility is reported for homozygous genotypes of some other mutation in *GDF9* gene in some sheep breeds, such as Belcare, Cambridge and Icelandic sheep breeds (Hanrahan *et al*. 2004; Davis 2005; Nicol *et al*. 2009; Pérez-Ruiz *et al*. 2020). The objective of this study was to perform a genetic analysis of segregating variants of the *GDF9* gene and estimate impact on litter size of adult ewes in Australian Merino and maternal sheep breeds using recently available whole genome sequence data.

MATERIALS AND METHODS

Phenotypes. Adult litter size (LS) phenotypes for Merinos and maternal breeds, which were respectively derived from the national Sheep Genetics database MERINOSELECT and

LAMBPLAN were used in this study. Both data sets included research (Information Nucleus Flocks and MLA resource flock) and industry animals (Sheep Genetics). Merino population consisted of purebred animals and the maternal population were a multi-breed/admixture of maternal sheep breeds including predominantly Border Leicester, Coopworth, Corriedale and composite maternal lines. Litter size phenotypes reflected the number of lambs counted at birth or were derived from pregnancy scanning records (Bunter *et al*. 2019, 2021). The total number of genotyped animals with LS recorded were 8,850 and 7,613 respectively for Merinos and maternal breeds, and 82% and 43% of these ewes had repeated records for Merino and maternal data set respectively. These data belonged to ewes born between 2007 and 2018.

Genotypes. Imputed sequence data on OAR5 were used in this study. A description of the imputation procedure is provided in Bolormaa *et al.* (2019). Briefly, research and industry data with low-density genotypes (12k, 15k) were imputed to 50k genotypes based on a large 50k reference set, and then all the 50k genotypes were imputed to high-density genotypes (500k: HD) using a 2,266 multi-breed reference set. Subsequently, animals with HD genotypes were imputed to sequence level using 726 multi-breed animals as a reference set (with on average 10x coverage). The final set of sequence data provided 1,600,633 variants on OAR5 after quality control and filtration for variants with low imputation accuracy ($r < 0.63$). SNPs with minor allele frequency of greater than 0.005 and at least 0.95 call rate were used in this study.

Statistical analysis. Phenotypes used in the association study were first corrected for environmental effects separately for research and industry data and according to the following equation in ASReml 4.1 (Gilmour *et al*. 2009): $\boldsymbol{y= Xb + Z_1a + Z_2pe + Z_1Qg + e}$. In this equation, y represents the phenotypes, $\boldsymbol{b}$ is a vector of fixed effects, consisting of mean, contemporary group (cohort of flock, birth year, management group) and age at lambing, $\boldsymbol{a}$ is the random direct additive genetic effect of the animal, fitted through the pedigree relationship matrix, $\boldsymbol{pe}$ is random permanent environmental effect of the animal, $\boldsymbol{g}$ is random effect of reed and **e** is random residual effect. $\boldsymbol{X}$, $\boldsymbol{Z_1}$, and $\boldsymbol{Z_2}$ are corresponding incidence matrices and $\boldsymbol{Q}$ is a matrix of contributions of genetic groups for all animals in the pedigree. The pre-corrected phenotype for each individual was the sum of the within group genetic and residual effects ($y^* = Z_1\hat{a} + \hat{e}$).

Association analysis was performed according to the single SNP mixed model regression method based on the following equation $\boldsymbol{y^*=Xb + Zu + e}$ in the Gemma V0.96 program (Zhou *et al*. 2014). In this equation y^* refers to the pre-corrected phenotypes, $\boldsymbol{b}$ refers to mean and allele substitution effect of the investigated SNP, $\boldsymbol{u}$ refers to the random additive genetic effect of the animal fitted by genomic relationship matrix (G), and $\boldsymbol{e}$ is the residual effect. G was calculated using 50k genotypes based on Yang *et al*. 2011, and X and Z are incidence matrix relating fixed and random effects to phenotypes.

RESULTS AND DISCUSSION

The litter size results showed the maternal breeds on average were more prolific (LS = 1.73) than Merinos (1.34) (Table 1). However, the heritability of litter size was higher in Merinos (0.13) compared to maternal breeds (0.08).

Table 1. Summary statistics of phenotypes and pedigree-based heritability of litter size in Merino and maternal sheep populations

Population	No. of Records	No. of Animals	Average	sd	range	h^2(se)
Merinos	547,807	295,748	1.34	0.53	1- 4	0.13 (0.02)
Maternal breeds	703,503	305,916	1.74	0.63	1- 5	0.08 (0.01)

The association results showed that SNPs in the *GDF9* region significantly affected LS in

maternal breeds (41,841,034 to 41,843,517 bp, Oar_V3.1, Ensembl Genome Browser; www.ensembl.org) (Figure 1). However, SNPs in this region did not have a significant effect on LS in Merinos. The significant region in maternal breeds was within the *GDF9* gene as well as both upstream and downstream of the gene. A total of 298 SNPs in this region were significantly associated initially (-log p value ≥ 6). However, the number of significant SNPs retained in this region after pruning for high LD (LD≥0.95) was 34 and spanned from position 40,685,116 to 45,175,518 bp on OAR5. Conditional and joint analysis of these remaining SNPs in stepwise multiple regression (Yang *et al*. 2012) identified that the most significant SNP within the *GDF9* coding region to be located at OAR5:41841588, which is in intron location of the *GDF9* gene. This SNP was associated with a 0.22 increase in LS and explained 4.75% of the total genetic variance. The frequency of this SNP was 1.2% in the maternal breed population. However, this SNP was not segregating in Merinos and other SNPs in the *GDF9* region were also not significantly associated with LS in Merinos.

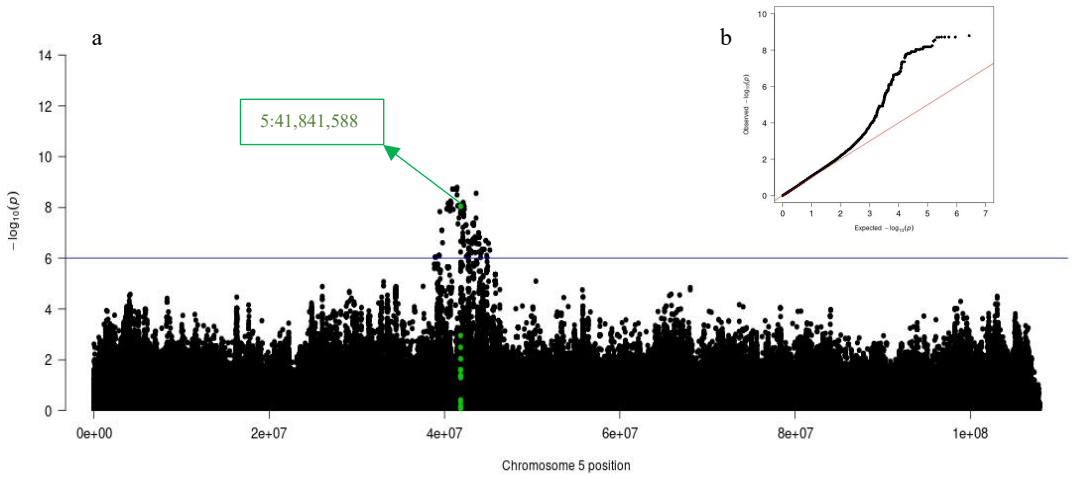

Figure 1. Manhattan plot of p_value of association between genetic markers on OAR5 and litter size in maternal breeds (a) and the associated QQ plot (b). Green dots show SNPs located within *GDF9* coding region

The *GDF9* gene, located on OAR5 in sheep, is important for normal folliculogenesis. Some polymorphisms including missense mutations in *GDF9* have been associated with between 0.2 to 0.7 increase in litter size (Davis, 2005) in various sheep breeds. In this study, we observed a highly significant region in *GDF9* and prioritized a SNP located within intron of *GDF9* in the maternal breeds only. This significant SNP was 243 base pairs apart from the causative mutation reported in Norwegian White Sheep (Våge *et al*. 2013), which was introduced to this breed by crossing with Finnish Landrace sheep. It is highly possible that the significant region found in maternal breeds here also originated from Finnish Landrace, due to historical introductions. The increase in litter size (0.22 lambs) in maternal breeds observed herein was within the range of increase in litter size reported for Finnish landrace (Våge *et al*. 2013).

Sterility associated with homozygous genotypes for some mutations in *GDF9* has been reported in some breeds (e.g. Nicol *et al*. 2009). However, in other breeds, such as Finnish Landrace, Norwegian White Sheep and Santa Ines, and for other mutations in *GDF9*, homozygous genotypes

have been reported to be fertile (Våge *et al*. 2013; Silvia *et al*. 2011). Herein, three animals homozygous for the significant SNP in the *GDF9* region were all fertile and showed higher LS than population average. However, due to low frequency of the SNP markers in the significant region and the small number of homozygous genotypes in this study, further investigation is required to confirm the fertility status in homozygous animals.

CONCLUSIONS

This study showed sequence variants located in *GDF9* were significantly associated with litter size in maternal breeds. The allele frequency of the favourable allele was 1.2% in the maternal population and explained 4.75% of the total genetic variance in LS. No such association was observed in Merinos. Further work is required to investigate the relationship between the significant region in *GDF9* with other reproductive traits and also the impact of homozygous genotypes on fertility and litter size. Information about *GDF9* could be useful for more accurate prediction of the genomic merits of selection candidates for LS in some breeds.

ACKNOWLEDGEMENTS

Phenotypic and data and genotypes used in this study were provided by Sheep Genetics from the MERINOSELECT and LAMBPLAN maternal evaluations, and the research was funded by MLA projects L.GEN.1815 and L.GEN.1704.

REFERENCES

Bolormaa S., Chamberlain A.J., Khansefid M., Stothard P., Swan A.A., Mason B., and MacLeod I.M. (2019) *Genet. Sel. Evol.* **51**: 1.

Bunter K.L., Swan A.A., Gurman P.M., and Brown D.J. (2021) *Anim. Prod. Sci.* **61** 333.

Bunter K.L., Swan A.A., Gurman P.M., Boerner V., McMillan A.J., and Brown D.J. (2019) *Proc. Assoc. Advmt Anim. Breed. Genet.***23***:* 560.

Davis G.H. (2005). *Genet. Sel. Evol.* **37:** 11.

Gilmour A.R., Gogel B.J., Cullis B.R., Welham S., and Thompson R. (2015) Hemel hempstead: VSN international ltd.

Hanrahan J.P., Gregan S.M., Mulsant P., Mullen M., Davis H.H., Powell R., and Galloway S.M. (2004) *Biol. Reprod.* **70**: 900.

Moghaddar N., Khansefid M., Van der Werf J.H.J., Bolormaa S., Duijvesteijn N., Clark S.A., Swan A.A., Daetwyler H.D., MacLeod I.M. (2019) *Genet. Sel. Evol*. **51**:72.

Nicol L., Bishop S.C., Pong-Wong R., Bendixen C., Holm L.E., Rhind S.M., McNeilly A.S. (2009). *Reproduction* **138**: 921.

Pérez-Ruiz E., Gallegos-Sánchez J., Cortez-Romero C., Segura-León O.l., Salinas-Ruíz J., and Salazar-Ortiz J (2020). *Anim. Genet*. **51:** 341.

Silva B.D.M., Castro E.A., Souza C.J.H., Paiva S.R., Sartori R., Franco M.M. and Melo E.O. (2011) *Anim. Genet.* **42:** 89.

Våge DI,. Husdal M,. Kent MP,. Klemetsdal G,. Boman IA (2013). *BMC Genet*. **14**: 1.

Yang J,. Ferreira T., Morris AP., Medland S.E. (2012). *Nat. Genet*. **44(4)**: 369.

Zhou X., and Stephens M., (2014). *Nat. Methods* **11:** 407.