Proc. Assoc. Advmt. Anim. Breed. Genet: 24: 427-430

GENOMIC ANALYSIS OF GENOTYPE BY ENVIRONMENT INTERACTIONS IN POST-WEANING WEIGHT OF AUSTRALIAN SHEEP

D.L. Waters [1], S.A. Clark[1], N.Moghaddar[1] and J.H.J. van der Werf[1]

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

SUMMARY

Genotype by environment interactions affecting post-weaning body weight in Australian sheep were investigated using a linear reaction norm incorporating genomic information. The definition of the environmental covariable used in the reaction norm was the best linear unbiased estimation of contemporary group effects for post-weaning growth rate. Significant variation in slope was estimated, and genetic correlations between low, medium and high growth environments ranged from 0.61 to 0.94, suggesting the presence of genotype by environment interaction. A putative QTL was detected on chromosome 11, significantly associated with both the intercept and slope of the reaction norm. Overall, SNP effects for the intercept and slope were highly correlated (0.87). The results suggest that selection based on (genomic) breeding values for the intercept and slope could yield animals that are more robust.

INTRODUCTION

Environmental extremes are very costly to agricultural systems. Annual farm-gate gross domestic product declined by approximately $3 billion AUD between 2017 and 2020 in Australia, due to drought (Reserve Bank of Australia, 2020). Climate change is expected to exacerbate this problem, with increasingly extreme and variable environments predicted for the future (Cowan *et al.* 2014). A potential response to this problem could be to breed livestock that are more robust to changes in the environment. An understanding of genotype by environment (GxE) interactions is necessary for this.

Genotype by environment interactions occur when the performance of genotypes is dependent on the environment they are recorded in. GxE acts as a source of variation from which to select robust livestock; robust genotypes maintain their genetic merit across environments, while sensitive genotypes change in merit. Biologically, if significant GxE implies that performance in different environments can be considered as different traits (Falconer 1952), a varying genetic architecture could be expected to determine merit across environments. Therefore, it should be possible to detect changes in the relative contribution of QTL influencing a trait across different environments using SNP and environmental data.

A popular way to model GxE is using reaction norm models (RNM). These allow the breeding values of animals to change across an environmental trajectory, often modelled as a linear function. This results in two breeding values for each animal; one corresponding to its breeding value in the mean environment (the intercept), and the other corresponding to the degree of change in breeding value across environments (the slope). Selection of animals with high breeding values for intercept and small breeding values for the slope could increase the mean of a trait while maintaining robustness of a flock to environmental extremes. Breeding values for the slope could also be used in a genome-wide association study to detect QTLs with environmental-dependent effects that are responsible for causing GxE (Silva *et al*. 2014).

Several QTLs affecting body weight are known to segregate in Australian sheep (Al-Mamun *et al*. 2015). This presents an interesting opportunity to investigate the behaviour of QTL in a trait with significant GxE (Clark *et al*. 2015). The aim of this study was to explore variance due to GxE and

identify genomic regions contributing to GxE and robustness in post-weaning weight using a genomic RNM.

MATERIALS AND METHODS

Data. Phenotypic data consisted of bodyweight records at weaning (64-120 days old) and post-weaning (123-329 days old) on 21,131 lambs in 206 contemporary groups born between 2007 and 2019 in the Information Nucleus Flocks and Resource Flocks. These flocks were located across Australia and were linked through common sires. All lambs were genotyped with an Illumina 50k-Ovine panel. An additional 10k was imputed from a recent Neogen GGP-Ovine panel that contained SNPs not included on the original Illumina chip. In total, the genomic data consisted of 60,345 SNPs after quality control.

The environmental covariable was defined by the best linear unbiased estimation of contemporary group effects using the rate of growth between weaning and post-weaning as the response variable, measured in grams/day. A minimum growth period of 40 days was applied to ensure the growth period was large enough to accurately reflect the environment. A small number of extreme contemporary groups were removed to prevent inflation of the regression coefficients. The environmental covariable ranged between -59.1 and +57.1 g/day when centred around zero, and was standardised between -1 and 1 for analysis.

Statistical analyses. A genomic RNM model was fitted using MTG2 2.18 (Lee & Van Der Werf 2016). The model was of the form: $\boldsymbol{y} = \boldsymbol{Xb} + \boldsymbol{Z_1 a_0} + \boldsymbol{Z_2 a_1} + \boldsymbol{Z_3 Q g} + \boldsymbol{Z_4 c} + \boldsymbol{e}$ where $\boldsymbol{y}$ is a vector of post-weaning weight records for each lamb, $\boldsymbol{b}$ is a vector of fixed effects, $\boldsymbol{X}$ is a design matrix linking fixed effects to records, $\boldsymbol{Z_1}$ and $\boldsymbol{Z_2}$ are the design matrixes linking records to additive genetic effects for the intercept ($\boldsymbol{a_0}$) and slope ($\boldsymbol{a_1}$), $\boldsymbol{Z_3}$ and $\boldsymbol{Z_4}$ are design matrices linking records to animals and dam environmental effects ($\boldsymbol{c}$), $\boldsymbol{Q}$ is a matrix linking animals to genetic groups, $\boldsymbol{g}$ is a vector of genetic group effects and $\boldsymbol{e}$ is the homogenous residual variance. Fixed effects included age at measurement, birth type and rear type interaction, sex, and contemporary group. The variance in intercept and slope was modelled as follows: $\begin{bmatrix} a_0 \\ a_1 \end{bmatrix} \sim N(0, \boldsymbol{G} \otimes \boldsymbol{K})$ where $\boldsymbol{K} = \begin{bmatrix} \sigma^2_{a0} & \sigma_{a1a0} \\ \sigma_{a0a1} & \sigma^2_{a1} \end{bmatrix}$ and $\boldsymbol{G}$ is the genomic relationship matrix (VanRaden, 2008). An environment-specific genetic (co)variance matrix was calculated using: $\mathbf{E} = \boldsymbol{\Lambda K \Lambda'}$ where $\boldsymbol{\Lambda}$ was a 3x2 matrix, with the first column containing a vector of ones for the intercept and the second column containing the standardised coefficient corresponding to the level in the environment. Three environmental levels were used to calculate **E**: -42 g/day (low growth), 0 g/day (average growth) and 42 g/day (high growth). SNP effects for the intercept and slope were estimated through back-solving genomic breeding values for a_0 and a_1 (Strandén & Garrick 2009) . P-values were approximated following Gondro (2015). A threshold value of $p > -\log_{10}(5)$ was chosen for significance.

RESULTS AND DISCUSSION

Variance components for the intercept and slope are reported in Table 1, along with genetic correlations between the three environmental levels. Genetic variance in the slope of the RNM was significantly different from zero. The correlation between intercept and slope was 0.50, indicating that genotypes with high performance in the average environment also tended to have a positive breeding value for slope. The genetic correlation between low and high growth environments was 0.61, while additive genetic variance increased from low to high growth environments. This suggests that both scaling and re-ranking contribute to GxE in this population.

Table 1. Reaction norm variance components for intercept ($\sigma^2 a_0$), slope ($\sigma^2 a_1$), covariance ($\sigma_{a_0 a_1}$) and correlation ($r_{a_0 a_1}$), as well as genetic correlations between low, average and high growth environments and genetic variance (Va) in each

RN variance component			Genetic correlations		
$\sigma^2_{a_0}$	6.63 (0.32)		Low	Average	High
$\sigma^2_{a_1}$	3.91 (0.58)	Average	0.84	-	-
$\sigma_{a_0 a_1}$	2.56 (0.25)	High	0.61	0.94	-
$r_{a_0 a_1}$	0.50	**Va**	4.96	6.63	12.13

Genome-wide SNP associations for the intercept and slope are shown in Manhattan plots (Figure 1). Regions on chromosome 6 and chromosome 11 were significantly associated with the intercept. The region on chromosome 6 was previously associated with body weight in Al-Mamun et al. (2015), while the region on chromosome 11 has not previously been reported in the literature. The same region on chromosome 11 was also significantly associated with the slope of RNM. Overall, the intercept and slope of post-weaning weight appears to be highly polygenic.

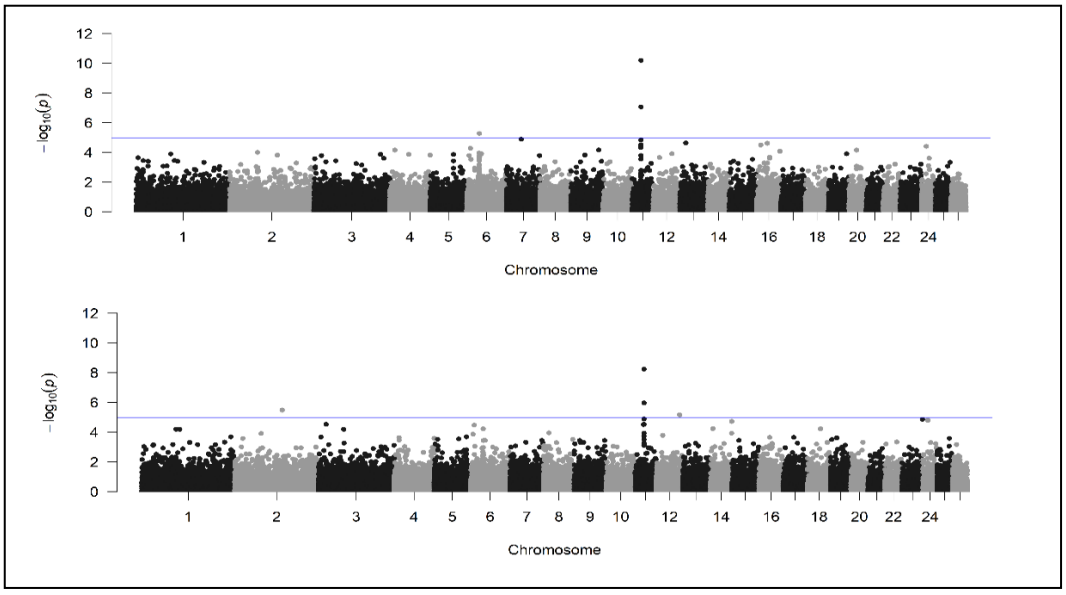

Figure 1. Genome-wide SNP associations for the intercept (top) and slope (bottom) of a reaction norm for post-weaning weight. The blue threshold line corresponds to p = 0.00001

SNP effects for the intercept and slope are plotted in Figure 2. The QTL on chromosome 11 appears to be contribute to GxE through a scaling effect, as its effect on slope is proportional to its effect on the intercept. The correlation between SNP effects was higher than anticipated (0.87), given that the genetic correlation between intercept and slope was 0.50. A possible explanation is that the breeding values for slope could be behaving similarly to a low-heritability trait in a multi-trait analysis, drawing on information from the higher heritability intercept. Methods to make breeding values for the intercept and slope more independent such as canonical transformation could remove variation in slope due to the intercept and yield a more useful GWAS for robustness. Further

investigation of this observation is warranted to better understand what the SNP effects for slope actually represent. This study serves as a preliminary investigation of GxE using genomic information. Several improvements for future analysis will be to model heterogenous residual variance, use a higher density SNP panel and explore higher-order polynomials or splines.

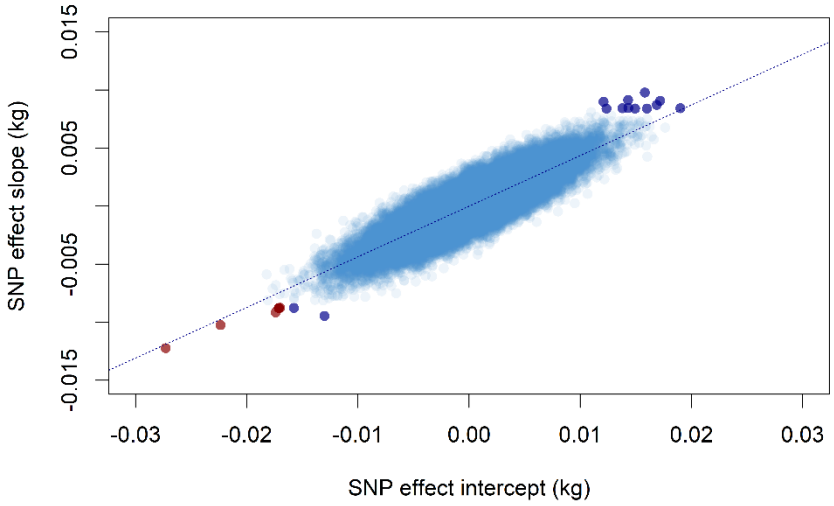

Figure 2. SNP effects for the slope regressed over SNP effects for the intercept. SNPs corresponding to the QTL on chromosome 11 are highlighted in red, while all other SNPs with p > 0.0001 are highlighted in dark blue. The Pearson correlation coefficient was 0.87

CONCLUSIONS

Significant genotype by environment interactions were detected using a linear genomic reaction norm. The genetic variance in intercept and slope indicated that breeding for robustness is feasible based on reaction norm models. The SNP effects for intercept and slope were highly correlated, with a QTL detected on chromosome 11 which affected both intercept and slope. Research into methods that remove variation in the slope due to the intercept could improve GWAS studies for robustness.

REFERENCES

Al-Mamun H.A., Kwan P., Clark S.A., Ferdosi M.H., Tellam R. and Gondro C. (2015) *Genetics Selection Evolution* **47**(1): 1.

Clark S.A., Lee S.H. and van der Werf J.H.J. (2015). *Proc. Assoc. Advmt. Breed. Genet.* **21**: 137.

Cowan T., Purich A., Perkins S., Pezza A., Boschat G. and Sadler K. (2014). *J. Climate* **27**(15): 5851.

Falconer D.S. (1952) *The American Naturalist* **86**(830): 293.

Gondro C. (2015). *Primer to Analysis of Genomic Data Using R*. Springer. Cham.

Lee S.H. and Van Der Werf J.H.J. (2016) *Bioinformatics* ***32***(9): 1420.

Santana M.L., Bignardi A.B., Eler, J.P., Cardoso, F.F. and Ferraz J.B.S. (2013) *J. Anim. Breed. Genet.* **130**(5): 394.

Silva F.F., Mulder H.A., Knol E.F., Lopes M.S., Guimaraes S.E.F., Lopes P.S., Mathur P.K., Viana M.S.and Bastiaansen J.W.M. (2014) *J. Anim. Sci.* ***92***(9): 3825.

Strandén I. and Garrick, D.J. (2009) *J. Dairy Sci.* **92**(6): 2971.

VanRaden P.M. (2008) *J. Dairy Sci.* ***91***(11); 4414.