

## **VALIDATION OF BREEDING VALUES FOR ROBUSTNESS IN AUSTRALIAN MERINOS**

**D.L. Waters<sup>1</sup>, J.H.J. van der Werf<sup>1</sup>, D.J. Brown<sup>2</sup>, S.F. Walkom<sup>2</sup> and S.A. Clark<sup>1</sup>**

<sup>1</sup>School of Environmental & Rural Science, University of New England, Armidale, NSW, 2351 Australia

<sup>2</sup>Animal Genetics and Breeding Unit\*, University of New England, Armidale, NSW, 2351 Australia

### **SUMMARY**

Livestock production often involves raising animals in environments which can vary substantially between locations and years. It could be beneficial to select animals that have genetic merit which is more robust to environment variation, rather than animals that are more sensitive. This study attempts to validate breeding values for robustness estimated using reaction norm models. Reaction norm models were used to regress breeding values for body weight across different growth environments in the Information Nucleus Flock. The same model was fit to MERINOSELECT data, and the rank-correlation for EBVs of sires with progeny in both datasets was calculated. The pattern of genetic variance and heritability across environments was very similar between datasets. The rank correlation of breeding values for a subset of sires with the best distribution of progeny in both populations was 0.60, 0.22 and 0.17 for the intercept, slope and scale-corrected slope, respectively. The results indicated that the genetic variation in robustness across growth environments was, to some extent, repeatable across the two datasets. Genotypes that re-ranked more in the INF/RF also tended to re-rank more in MERINOSELECT, although the relationship was weak. The analysis could benefit from the inclusion of genomic data to increase linkage across environments and between datasets.

### **INTRODUCTION**

Genotype-by-environment (GxE) interactions occur when the effect of an animal's genotype is dependent on the environment it exists in. This can result in variation between individuals in the robustness of their genetic effect to different environments. In extensive livestock systems where environments can vary substantially between years, genotypes that consistently rank highly for important traits across environments (i.e., robust genotypes) could be more valuable than sensitive genotypes who tend to change in rank.

Reaction norm (RN) models have been used widely to study GxE and rank individuals based on their robustness to environmental variation. Unlike univariate models, RN models allow the estimated breeding value (EBV) of genotypes to change as a function of an environmental covariable (EC), which describes the quality of the environment. When a linear function is used, the change in EBV across the EC is given by the slope. The slope can be directly used as an EBV for how robust the performance of a genotype is across the EC, while the slope can be 'scale-corrected' to yield an EBV for how much a genotype re-ranks across the EC (Waters *et al.* 2022).

Although some research has demonstrated RN models can increase the accuracy of phenotypic predictions (Oliveira *et al.* 2018; Mota *et al.* 2020), it could also be useful to explore whether RN variance components and individual breeding values (e.g., intercept, slope and scale-corrected slope) are repeatable in independent, but genetically linked populations. This would provide some guidance on how reliable EBVs based on linear RN models might be if applied in practice to select for robustness. The Australian Sheep CRC Information Nucleus Flock (INF) and the Meat and

---

\* A joint venture of NSW Department of Primary Industries and the University of New England

Livestock Australia Resource Flock (RF) contain the progeny performance of sires across several locations representative of the Australian sheep environments (Van der Werf *et al.* 2010). Many of these same sires also have progeny recorded across several locations and years in the wider MERINOSELECT population (Brown *et al.* 2007). This data structure presents the opportunity to investigate such a question. The aim of the study was to investigate whether RN breeding values for the robustness of performance in post-weaning weight to different growth environments in the INF/RF can be validated in the MERINOSELECT data.

## MATERIALS AND METHODS

The analysis consisted of two parts. The environmental covariable (EC) was first estimated for each animal in the INF/RF and MERINOSELECT data. Reaction norm models were then fit separately to both datasets to estimate breeding values for robustness across growth environments, which were then compared for sires with the most progeny in both data sets. All models were fit using ASReml 4.2 (Gilmour *et al.* 2021).

**Estimate the EC.** MERINOSELECT and INF/RF animals with a weaning weight (WWT) recorded between 50-120 days of age and a post-weaning weight (PWT) recorded between 120-329 days of age, along with a recorded sire and dam were extracted. Animals were excluded from the analysis if they were born or reared as quadruplets or greater, and if the age of dam was more than 12 years old at the time of recording. Contemporary groups were formed based on a flock × year × management group combination and required at least 15 animals from at least 3 different sires.

The best linear unbiased estimation (BLUE) of the post-weaning growth rate (PWGR) of each contemporary group was used as the EC for each animal. PWGR was calculated as the difference between PWT and WWT measurements, divided by the number of days between the measurements and expressed in grams per day. Animals with less than 40 days between WWT and PWT measurements were removed, along with animals deviating more than 3 standard deviations (SD) from their contemporary group mean for PWGR. This left 12,087 and 277,060 animals in the INF/RF and MERINOSELECT populations respectively. An animal model with PWGR as the response variable was fit jointly to the INF/RF and MERINOSELECT data to obtain a BLUE of PWGR for each contemporary group, forming the EC. The EC was centred to a mean of zero.

**Independent reaction norms.** Contemporary groups more than 3 SD from the population mean EC, and individuals more than 3 SD from their contemporary group mean PWT were excluded from the analysis. To reduce the number of uninformative animals in the MERINOSELECT data, contemporary groups were only included if they contained 1) at least one direct progeny of a sire with progeny or grand-progeny in the INF/RF, or 2) more than 25% of the animals were related to the INF/RF, with the minimum relationship being a grandsire with grand-progeny in the INF/RF. This left 11,638 and 206,733 animals in the INF/RF and MERINOSELECT data, respectively. The linear RN models were of the form:

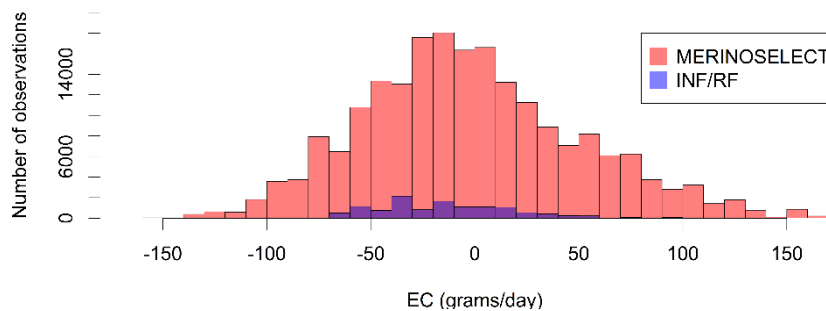
$$\mathbf{y} = \mathbf{X}\mathbf{b} + \mathbf{Z}_1\mathbf{a}_0 + \mathbf{Z}_2\mathbf{a}_1 + \mathbf{Z}_3\mathbf{c} + \mathbf{Q}\mathbf{g} + \mathbf{e} \quad (1)$$

Where  $\mathbf{y}$  is the vector of PWT records,  $\mathbf{X}$  is an incidence matrix for the fixed effects  $\mathbf{b}$ ,  $\mathbf{Z}_1$  and  $\mathbf{Z}_2$  are matrices relating records to the additive genetic effects for the intercept ( $\mathbf{a}_0$ ) and slope ( $\mathbf{a}_1$ ) respectively,  $\mathbf{Z}_3$  is an incidence matrix relating records to the additive maternal effects ( $\mathbf{c}$ ), which were estimated independently of the additive genetic effects,  $\mathbf{Q}$  is a matrix of the proportion of each animal's genome originating from 451 genetic groups,  $\mathbf{g}$  is the vector of random genetic group effects, and  $\mathbf{e}$  is the vector of the residual effects. Fixed effects included age at measurement, birth type and rear type interaction, sex, and contemporary group. The residual variance was estimated independently at four and six intervals along the EC for the INF/RF and MERINOSELECT population respectively. The variance of the intercept and slope was modelled as follows:

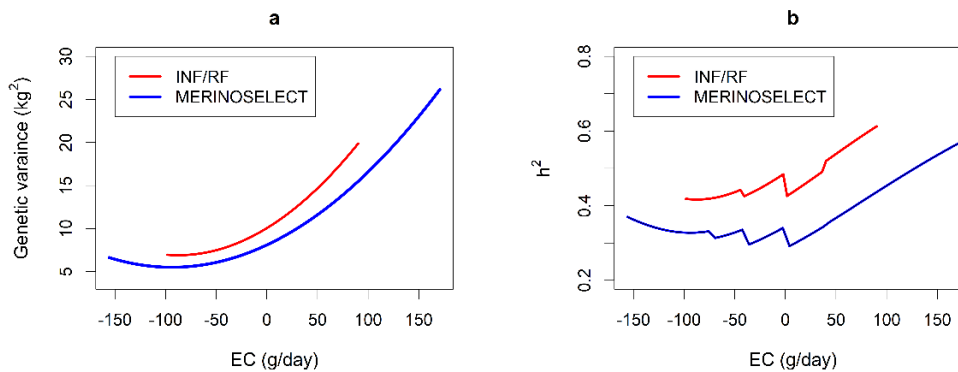
$\begin{bmatrix} \mathbf{a}_0 \\ \mathbf{a}_1 \end{bmatrix} \sim N(0, \mathbf{A} \otimes \mathbf{K})$  where  $\mathbf{K} = \begin{bmatrix} \sigma_{a_0}^2 & \sigma_{a_1 a_0} \\ \sigma_{a_0 a_1} & \sigma_{a_1}^2 \end{bmatrix}$  and  $\mathbf{A}$  is the pedigree relationship matrix. The genetic variance across the EC was obtained using  $\mathbf{G} = \mathbf{A}\mathbf{K}\mathbf{A}'$ , where  $\mathbf{A}$  contained two columns; the first was a vector of 1's, and the second was a vector of EC values. The heritability of PWT at a given EC level was obtained by dividing the genetic variance at the EC by the sum of the genetic, maternal and residual variance. Scale-corrected EBVs for the slope were estimated using a genetic regression (Waters *et al.* 2022), which makes the slope EBVs independent of the intercept EBVs.

## RESULTS AND DISCUSSION

Animals were normally distributed across the EC in both datasets (Figure 1), although there was a larger range in MERINOSELECT. The pattern of genetic variance and heritability across the EC was very similar between the two datasets (Figure 1), although the genetic variance and heritability were slightly lower in MERINOSELECT. Overall, the RN models estimated very similar levels of GxE in both populations.



**Figure 1. Distribution of animals across the EC in the INF/RF and MERINOSELECT data sets**



**Figure 2. Genetic variance (a) and heritability (b) of PWT across the EC in INF/RF and MERINOSELECT data estimated using the independent reaction norm models**

The rank-correlation of EBVs for sires with progeny in both populations was small but positive for the slope and scale-corrected slope (Table 1), and higher for the intercept. The difference in correlation between the intercept and slope is likely a function of the accuracy of the EBVs, as the intercept is generally easier to estimate accurately than the slope. The correlations were considerably higher when considering a subset of 56 sires with the best distribution of progeny in both populations, highlighting the importance of data structure when estimating RN parameters. Unlike

the slope EBVs, the scale-corrected slope EBVs were uncorrelated with the intercept, so they represented the slope variation available for selection independent of the overall performance.

**Table 1. Rank correlation of RN EBVs for sires with direct progeny in both populations (a), and a subset of 56 sires with progeny in at least 4 contemporary groups ranging by at least 50 g/day (b)**

| EBV                   | All sires (a) | Subset Sires (b) |
|-----------------------|---------------|------------------|
| Intercept             | 0.43          | 0.60             |
| Slope                 | 0.15          | 0.22             |
| Scale-corrected slope | 0.02          | 0.17             |

While these results imply that selection based on RN EBVs could yield a response in the robustness of performance across growth environments while simultaneously increasing the mean (intercept), the relationship between datasets was weak to moderate. This was most likely influenced by the distribution of progeny across the EC. To accurately estimate EBVs for the slope (robustness), sires require progeny across a wide range of EC values (Calus *et al.* 2004). Because only a relatively small number of sires had progeny widely distributed across the EC in both datasets, the power to detect a relationship between robustness in the two datasets was probably limited. Utilising genomic data to increase genetic linkage across the EC could help address this issue.

Overall, it appears that the success of breeding for robustness will be dependent on the structure of data available to estimate it accurately. If robustness is to be considered in genetic evaluations, breeders should be encouraged to ensure even stronger genetic linkage across years and locations. Other traits and environmental descriptors should also be explored to better understand the total variation available for selection of robustness.

## CONCLUSION

The analysis demonstrated that the RN models estimated very similar levels of GxE across the two populations. The rank-correlation of EBVs for sires with the best distribution of progeny in both populations was low but positive for the slope EBVs. The results indicated that genetic variation in the RN slope was repeatable the two datasets, so selection based on these EBVs should lead to a response for robustness across growth environments. However, the relationship was not strong. The analysis could be improved by using more accurate EBVs for the slope, which could be achieved by increasing the linkage between environments with genomic data.

## REFERENCES

- Brown D.J., Huismanm A.E., Swan A.A., Graser H.U., Woolaston R.R., Ball A.J., Atkins K.D. and Banks, R.G. (2007) *Proc. Assoc. Advmt. Anim. Breed. Genet.* **17**: 187.
- Calus M.P., Bijma P. and Veerkamp R.F. (2004) *Genet. Sel. Evol.* **36**: 5.
- Gilmour A.R., Gogel B.J., Cullis B.R., Welham S.J. and Thompson R. (2021) ASReml User Guide Release 4.2 Functional Specification, VSN International Ltd, Hemel Hempstead, HP2 4TP, UK
- Mota L.F.M., Fernandes G.A. Jr, Herrera A.C., Scalez D.C.B., Espigolan R., Magalhães A.F.B., Carvalheiro R., Baldi F. and Albuquerque L.G. (2020) *Anim. Genet.* **51**: 2.
- Oliveira D.P., Lourenco D.A.L., Tsuruta S., Misztal I., Santos D.J.A., de Araújo Neto F.R., Aspilcueta-Borquis R.R., Baldi F., Carvalheiro R., de Camargo G.M.F., Albuquerque L.G., Tonhati H. (2018) *J. Anim. Sci.* **96**: 1.
- van der Werf J.H.J., Kinghorn B.P. and Banks R.G. (2007) *Anim. Prod. Sci.* **50**: 998.
- Waters D.L., Clark S.A., Moghaddar N. and van der Werf J.H.J. (2022) *Genet. Sel. Evol.* **54**: 4.