GENOME-WIDE ASSOCIATION STUDIES FOR BODY WEIGHT AND AVERAGE DAILY FEED INTAKE DURING THE FEEDLOT TEST PERIOD

J.A. Torres-Vázquez, J.H.J. van der Werf and S.A. Clark

School of Environmental and Rural Science, University of New England, Armidale, 2351 Australia

SUMMARY

Feed efficiency component traits such as body weight (BW) and daily feed intake are economically relevant traits in beef cattle breeding programs. The objectives of this work were to identify genomic regions associated with BW and average daily feed intake (ADFI) during the feedlot period, and to evaluate whether these genetic variants for each trait were consistent over the 70-day test period. Data on 2070 Angus steers were used to estimate (co)variance components using the genomic relationship matrix (gREML) fitted in ASReml. For the studied traits, a two-trait repeatability (TT-REPM) and a two-trait random regression (TT-RRM) models were performed. SNP-effects for the TT-REPM and TT-RRM were estimated using a post analysis back-solving approach using the genomic estimated breeding values from each model respectively. For each trait, results were validated with single-trait animal models (ST-ANIMs) at the beginning and at the end of the test period using single-SNP regression in the GCTA software. Results from the genome-wide association studies (GWAS) using TT-REPM and TT-RRM were similar to the conventional approach using the ST-ANIMs. For all models, the variants rs43350564 and rs109326204 presented the strongest association with BW and ADFI, respectively. The identified SNP effects remained constant throughout the feedlot test period and could be useful for understanding the biology of feed efficiency. Further studies with more data and possibly with longer feed lot test periods are needed to investigate the effect of genomic regions for feed efficiency traits over the feedlot trajectory.

INTRODUCTION

Feed efficiency, commonly referred to as the conversion of feed into useable animal products, and their component traits such as body weight (BW) and average daily feed intake (ADFI) are economically relevant traits in beef cattle breeding programs. Identifying single nucleotide polymorphisms (SNPs) as genetic markers linked to quantitative trait loci (QTL) associated with BW and ADFI may aid in unravelling the biology underlying feed efficiency. These genetic markers for QTL can be identified through genome-wide association studies (GWAS). In beef cattle, GWAS for BW and ADFI have been addressed using the average measurements during the feedlot test period (Bolormaa *et al.* 2011).

For both traits, previous studies have documented that pedigree-based genetic parameters change during the trajectory of the feedlot period (Torres-Vázquez *et al*. 2018), and therefore it is expected that QTL-effects may also change over time. The objective of this study was to identify genomic regions associated with body weight and average daily feed intake during the feedlot period, and to evaluate whether these genetic variants for both traits were consistent over the 70-day test period.

MATERIALS AND METHODS

The phenotypic data included BW and feed intake measures from 2220 Angus steers collected from 2013 to 2017 at Tullimba Research Feedlot (30°20′S, 151°10′E, altitude 560 m), NSW, Australia. On entry to the feedlot, steers ranged from 500-600 days of age with an average weight of 578 kg. Initially steers were conditioned for 21 days and fed for an additional 70 days over which time all data was collected. Steers were weighed 6 times over the 70-day test period (fortnightly). Daily feed

intake measurements were averaged over 14-day periods to align them with the BW measurements to create average daily feed intake (ADFI). Duplicated and incomplete records were discarded (see Torres-Vázquez *et al.* 2018). The final data file consisted of 2,070 Angus steers. The pedigree file included an historical file with 14,662 animals with 1,454 sires and 7,835 dams; with 191 sires and 1,782 dams having progeny with phenotypic records. Contemporary groups, as defined by BREED-PLAN, included the concatenation of herd, year of birth, birth type (single or twin), breeder-defined management group, observation date and age (Graser *et al.* 2005).

Animals with phenotypes were genotyped with a range of low-density marker chips. These genotypes were imputed to higher density based on a reference of 7626 animals genotyped using the Illumina Bovine 50K v2 (54609 SNP). Quality control of the SNP markers was performed to eliminate SNP with a call rate less than 90% and minor allele frequency less than 1%. The remaining 39,136 SNPs passed the quality control measures and were acceptable for the analyses. Low density genotypes were then imputed to 39,136 SNP using FImpute (Sargolzaei *et al*. 2011). The genomic relationship matrix (GRM) was subsequently created using the GCTA software (Yang *et al*. 2011) from the imputed genotypes.

To analyze the genomic associations between traits a two-trait repeatability (TT-REPM) and a two-trait random regression (TT-RRM) model were undertaken. (Co)variance components for each analysis were estimated using ASReml incorporating the genomic relationship matrix (GRM (Gilmour *et al*. 2009). The most suitable fit of the models were assessed based on the log likelihood (LogL), Akaike's information criterion (AIC), and the Bayesian information criterion (BIC).

Several researchers have documented the equivalences between snpBLUP and gBLUP for genomic selection (Strandén and Garrick, 2009; Gondro, 2015). Therefore, SNP effects for the two-trait models were obtained following the methodology described in Gondro (2015), where:

$$\hat{u}_i = \left(\frac{1}{d} W\right) GRM^{-1} GEBV_i$$

where $\hat{\boldsymbol{u}}_i$ is a vector of the predicting SNP marker effects for the *i*th individual; d represent a scalar of the deviation effects calculated as $2 * \sum(p * q)$; **W** represent the SNP marker matrix corrected for the allele frequency differences (**M** – 2 * (p – 0.5)). **M** is the matrix of marker genotypes coded as: 1 for the heterozygous genotype, 0 and 2 for the genotype which is homozygous for the first and second allele, respectively; GRM^{-1} represent the inverse of the GRM; and $\mathbf{GEBV}_i$ is a vector of genomic estimated breeding values (GEBVs) obtained from the gREML model. For this approach, p-values of 0.05 were estimated based on a t-distribution calculated as the probability value of the 95th percentile of the GEBV distribution.

To validate the SNP-effects calculated for the TT-REPM and for the TT-RRM at days 5 and 70, GWAS were conducted with single-trait animal models (ST-ANIMs) based on data evaluated at the beginning and at the end of the test period (days 1 and 70, respectively) using single-SNP regression in GCTA (Yang *et al.* 2011).

RESULTS AND DISCUSSION

The TT-RRM had the highest LogL, and smallest value for AIC and BIC, showing the best fit for this model (Table 1). In general, high genomic heritability estimates were observed for BW compared to ADFI. The TT-RRM yielded the highest range for genomic heritability estimates with higher genomic heritability estimates for BW. As expected, repeatability estimates for both traits increased across the feed lot test period, and these estimates were higher for BW compared to ADFI, suggesting that measurement errors are more relevant for the accuracy of ADFI. Our genomic heritability estimates for BW and ADFI followed the same pattern as the pedigree-based estimates reported by Torres-Vázquez *et al*. (2018). Using the TT-REPM, the genetic correlation was of 0.69 ± 0.05. How-

ever, with the TT-RRM, this correlation increased from 0.63 at day 1 up to 0.75 at day 82. This was a slight reduction in the difference observed in the pedigree analysis (0.56 to 0.82) by Torres-Vázquez *et al*. (2018). Nevertheless, the increase in correlation potentially indicates that the genes that cause variation in the two traits are more similar over time.

Table 1. Measures of goodness of fit, genomic heritability (h^2) and repeatability (rep) for the two-trait models

Model / Trait	n	Log L	AIC	BIC	h^2	rep
TT-REPM, BW	9	-26.5	100,071.0	100,143.5	0.46 ± 0.04	0.88 ± 0.01
TT-REPM, ADFI	9	-26.5	100,071.0	100,143.5	0.26 ± 0.03	0.59 ± 0.01
TT-RRM, BW	23	7,187.8	94,421.5	94,606.3	From 0.42 to 0.53	From 0.92 to 0.94
TT-RRM, ADFI	23	7,187.8	94,421.5	94,606.3	From 0.36 to 0.23	From 0.68 to 0.70

SNP-effects obtained by back-solving the TT-REPM and the TT-RRM at days 5 and 70 followed the same pattern as those yielded by the ST-ANIM using GCTA (Figure 1). For each trait, only one SNP exceeded the significance level.

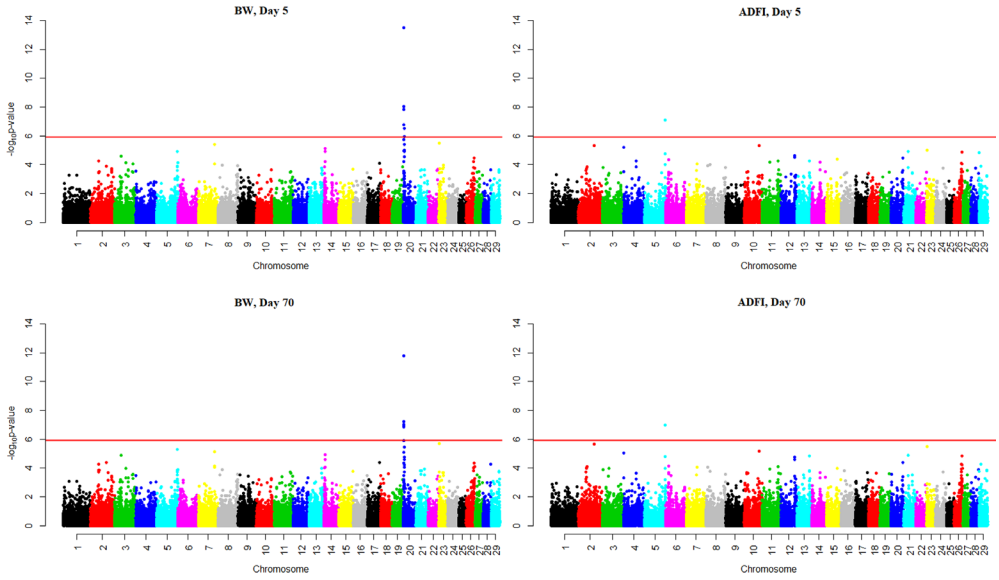

Figure 1. Manhattan plot for the genomic associations of BW and ADFI at days 5 and 70, using the back-solving approach with the two-trait random regression model

The most strongly associated marker (rs43350564) for BW was located on chromosome 20 at the position 4,618,689. This SNP was located 269 bp downstream from *ERGIC1* (*endoplasmic reticulum-golgi intermediate compartment 1).* This gene is potentially involved in increased protein turnover and has been previously associated with increases in multiple liveweight measures in American Simmental, Red Angus and Gelbvieh (Saatchi *et al*. 2014). The most significant SNP (rs109326204) associated with ADFI was located on chromosome 5 at the position 120,378,417. This SNP is located in the CELSR1 gene, which has been associated with decreases in body mass in mice (Zerbino *et al*. 2018).

The methodology implemented in this work to obtain the SNP effects, is easy to apply without transforming phenotypes but it has some limitations. Given the increasing genetic correlation together with the low accuracies, the GEBVs obtained with the TT-RRM were highly correlated between days (>0.987). This yielded very similar SNP effects in each trait with low probabilities of identifying accurately other genetic markers along the test period. Besides, the implemented methodology was sensitive to several factors. For example, the number of animals with phenotypic data collected with the greatest precision, imputation quality, and the genomic accuracy of the trait. In addition, in the presence of small SNP effects for a trait, further samples may be necessary to detect them.

CONCLUSIONS

The TT-RRM showed that the genetic parameters tended to change over the feedlot test period. With this model, the genomic correlation estimates increased over the whole trajectory from 0.63 to day 1 to 0.75 at the end of the period. In this work, SNP effects obtained by the TT-REPM and TT-RRM yielded similar results to conventional GWAS approaches. Two SNPs were identified by GWAS that may be useful for understanding the biology of feed efficiency in beef cattle. Further studies are necessary to investigate the change of genomic regions including more samples in longer feedlot test period for feed efficiency traits.

ACKNOWLEDGEMENTS

The authors thank the Angus Society of Australia for their support of this research and access to the data for this study through the Angus Sire Benchmarking Project (ASBP). Meat and Livestock Australia (MLA) have provided co-funding support for the ASBP (PSH.0528). Generous support has also been provided to the ASBP by numerous Angus bull breeders, co-operator herds, supply chain partners, technicians and research organizations.

REFERENCES

Bolormaa S., Hayes B.J., Savin K., Hawken R., Barendse W., Arthur P.F., Herd R.M. and Goddard M. E. (2011) *J. Anim. Sci.* **89**: 1684.

Gilmour A.R., Gogel B., Cullis B., Thompson R., and Butler D. (2009) VSN International Ltd, Hemel Hempstead, UK.

Gondro C. (2015) 'Primer to analysis of genomic data using R'. Springer. Cham.

Graser H.-U., Tier B., Johnston D.J. and Barwick S.A. (2005) *Aust. J. Agric. Res.* **45**: 913.

Sargolzaei M., Chesnais J. and Schenkel F. (2011) *J. Dairy Sci.* **94**: 421.

Saatchi M., Schnabel R.D., Taylor J.F. and Garrick D.J. (2014) *BMC Genomics* **15**: 1004.

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

Torres-Vázquez J.A., van der Werf J. and Clark S.A. (2018) *Proc. World Congress on Genetics Applied to Livestock Production*. Feed Intake and Efficiency **2**:372.

Yang J., Lee S.H., Goddard M.E. and Visscher P.M. (2011) *Am. J. Hum. Genet.* **88**: 76.

Zerbino D.R, Achuthan P., Akanni W., Amode M.R., Barrell D., Bhai J., Billis K., Cummins C., Gall A., García Giron C., Gil L., Gordon L., Haggerty L., Haskell E., Hourlier T., Izuogu O.G., Janacek S.H., Juettemann T., Kiang To J., Laird M.R., Lavidas I., Liu Z., Loveland J.E., Maurel T., McLaren W., Moore B., Mudge J., Murphy D.N., Newman V., Nuhn M., Ogeh D., Ong C.K., Parker A., Patricio M., Riat H.S., Schuilenburg H., Sheppard D., Sparrow H., Taylor K., Thormann A., Vullo A., Walts B., Zadissa A., Frankish A., Hunt S.E., Kostadima M., Langridge N., Martin F.J., Muffato M., Perry E., Ruffier M., Staines D.M., Trevanion S.J., Aken B.L., Cunningham F., Yates A. and Flicek P.F. (2018) *Nucleic Acids Research*. **46**: D754.