*Proc. Assoc. Advmt. Anim. Breed. Genet. 18:125-128*

# GENOME WIDE ASSOCIATION STUDIES FOR NET FEED INTAKE, BODY WEIGHT AND HIP HEIGHT IN BEEF CATTLE

**S. Bolormaa[1] and M.E. Goddard[1,2]**

CRC for Beef Genetic Technologies
[1] Department of Primary Industries, Victoria
[2] University of Melbourne

## SUMMARY

We report on a genome-wide association study (GWAS) from the CRC for Beef Genetic Technologies using the 50k Illumina SNP chip. Here, we present the results for net feed intake (NFI), body weight and hip height. The aims of this analysis are to discover SNPs associated with all traits but especially NFI and to test the consistency of SNP effects across datasets and breed types using the weight and height data. The data were analysed within datasets and within breed type using a mixed model and fitting one SNP at a time. In each case the number of significant SNPs was more than expected by chance alone. However, the SNP effects for weight and height were consistent between datasets only when estimated in the same breed type (*B.inducus* x *B.taurus* composite breeds). While NFI was only measured in one dataset, we found 9 SNPs associated with NFI on BTA 3, 5, 7 and 8 with $P \leq 6.0 \times 10^{-5}$.

## INTRODUCTION

In genomic selection, the estimation of breeding values is based on genetic markers. This would be particularly useful for traits that are very expensive to measure such as net feed intake (NFI). In beef cattle, some studies (Barendse *et al.* 2007; Nkrumah *et al.* 2007; Sherman *et al.* 2009) have reported associations between markers and NFI. For instance, Barendse *et al.* (2007) analysed 8,786 polymorphic SNPs in 189 Australian beef cattle, selected for high and low NFI, and detected 161 SNPs associated with NFI at P<0.01. However, the availability of 50K SNP chip from Illumina provides us with an opportunity to conduct a more powerful genome-wide association study (GWAS) for NFI.

Before genomic selection can be implemented with confidence, it is necessary to confirm in independent populations the associations that have been discovered in one population. Often such attempts at confirmation have been unsuccessful. Failure to confirm associations can be due to breed-specificity. That is, an association is only found in one breed or group of breeds. Here we use data on weight and hip height in cattle to test the consistency of associations across breed types. To do this, we have conducted GWAS in cattle from two datasets including the Beef CRC I dataset comprising *Bos taurus*, *B. indicus* and crosses between the two subspecies, and the Beef CRC II dataset with *B. indicus* and crosses between B. *indicus* and *B. taurus* cattle.

## MATERIALS AND METHODS

**SNP data**. In total, 53,798 SNPs were genotyped. Preliminary edits were carried out at the Animal Genetics and Breeding Unit of the University of New England in Armidale, NSW. In brief, genotypes were discarded if they did not have high quality scores (>95%) and the proportion of missing genotypes did not pass defined criteria (≥20%). Sixteen thousand and eight SNPs had minor allele frequency (MAF) < 0.05 and 8,469 SNPs deviated from Hardy-Weinberg equilibrium (HWE; P<0.0001). However, these were not removed from the further analyses. Out of the initial 53,798 SNPs, 50,650 were polymorphic and included in the GWASs.

**Animals and population structure.** Eight hundred and fifty two steers from the Beef CRC I population had both genotype and phenotype data for NFI, weight and height (Table 1). These steers were from 7 different pure breeds. Four breeds (Angus, Murray Grey, Shorthorn and Hereford) were *Bos taurus* (BT), 1 breed (Brahman) was *Bos indicus* (BI) and two breeds (Santa Gertrudis and Belmont Red) were BT×BI synthetic breeds. Additionally, 1,456 cows from the Beef CRC II population had weight and most had height data. These cows were sourced from two different breed types, BI (Brahman) and BI×BT crosses.

**Traits**. CRC I steers were recorded in the feedlot for four traits: net feed intake (NFI), average daily gain (ADG), daily feed intake (DFI), and mean metabolic weight (mMWT); and prior to the feedlot period, for post weaning hip height (pwHH). The CRC II heifers were recorded for first wet season weight (w1WGT) and hip height (w1HH) (Table 1).

**Table 1. Number of records (N), mean, standard deviation (SD) and estimates of heritability ($h^2$) and associated standard errors (SE) for all traits studied**

| Trait | CRC | N | BT | BI | BT×BI | ALL-mean | ALL-SD | ALL-$h^2$ | ALL-$h^2$_SE |
|---|---|---|---|---|---|---|---|---|---|
| NFI | I | 852 | 486 | 78 | 288 | -2.6 | 1.2 | 0.18 | 0.13 |
| DFI | I | 852 | 486 | 78 | 288 | 12.3 | 2.1 | 0.16 | 0.13 |
| ADG | I | 852 | 486 | 78 | 288 | 1.4 | 0.4 | 0.24 | 0.14 |
| mMWT | I | 852 | 486 | 78 | 288 | 93.8 | 11.4 | 0.31 | 0.15 |
| W1WGT | II | 1456 | - | 590 | 866 | 301.3 | 44.3 | 0.61 | 0.11 |
| pwHH | I | 812 | 466 | 65 | 281 | 116.4 | 6.5 | 0.25 | 0.18 |
| W1HH | II | 1224 | - | 360 | 864 | 126.0 | 5.8 | 0.60 | 0.12 |

NFI = net feed intake (kg/day); DFI = daily feed intake (kg/day); ADG = average daily gain (kg/day); mWGT = mean metabolic weight ($kg^{0.75}$); w1WGT ="end of wet season 1" weight (kg); pwWGT = post-weaning hip height (cm); w1HH = ="end of wet season 1" hip height (kg); - = not available

**Statistical analyses**. The association between each SNP and each of the traits was assessed by a regression analysis using the ASReml software (Gilmour *et al.* 2002). The mixed model applied was as follows: trait ~ mean + fixed effects + $SNP_i$ + animal + error; with animal and error fitted as random effect. Fixed effects were different for the CRC I and CRC II datasets. For CRC I data, breed, origin of herds, sex, year of measurement, season, market-weight destination and nutritional treatment were fitted as fixed effects, age deviation from group mean and $SNP_i$ were fitted as covariate effects. Whereas for CRC II data the effects of breed, origin of herds, sire group, cohort, calving month and their first degree interactions were fitted as fixed effects. Using the same model without fitting $SNP_i$, estimates of heritability were calculated based on the genotyped animals (Table 1) and their ancestors.

The two datasets were analysed separately. Additionally, the three breed types within the CRC I data (BT, BI and BTxBI) were also analysed separately as well as in a joint analysis. Similarly, the two breed types (BI and BTxBI) represented in the CRC II dataset were analysed separately as well as jointly. Weight and hip height were measured in both datasets and therefore we could compare the SNP solutions from the two datasets. We calculated the number of SNPs that were significant in both datasets and for these SNPs we calculated two parameters to assess the agreement between the results: the correlation between SNP solutions in the two datasets, and the proportion of SNPs in which the solutions were in the same direction; that is, the proportion in which the same SNP allele increased the trait.

SNPs effects with high standard errors are sometimes large but the effects are poorly estimated. Therefore, the SNP effects were divided by their standard error before correlations of the SNP effects were estimated.

In a GWAS there are many thousands of significance tests performed. Therefore, we compared the number of SNPs that were significant to the number expected by chance using a False Discovery Rate (FDR) as described by Storey (2002).

## RESULTS AND DISCUSSION

Heritability ($h^2$) estimates for NFI based on genotyped animals was a bit lower (0.18) than, but within a SE, of the one found for the complete CRC I dataset (David Johnston, pers. comm.). Overall, $h^2$ estimates were low to moderate for feedlot traits and moderate to high for weight and height (Table 1).

The number of significant SNPs at threshold of P<0.001 (Table 2) was 78 for weight and 75 for height in CRC I, but 156 and 134 in the larger CRC II data. Consequently, the FDRs were lower in CRC II (32% and 37%) compared with CRC I (65% and 67%).

**Table 2. Number of significant SNPs and FDR at different thresholds based on all animals studied**

| | No. of SNPs at P | | | | FDR (%) at P | | | |
|---|---|---|---|---|---|---|---|---|
| Trait | <0.0001 | <0.001 | <0.01 | <0.05 | <0.0001 | <0.001 | <0.01 | <0.05 |
| NFI | 11 | 75 | 615 | 2826 | 46 | 67 | 82 | 89 |
| DFI | 8 | 76 | 624 | 2733 | 63 | 67 | 81 | 92 |
| ADG | 11 | 83 | 698 | 2995 | 46 | 61 | 72 | 84 |
| mMWT | 6 | 78 | 694 | 3141 | 84 | 65 | 73 | 80 |
| w1WGT | 29 | 156 | 935 | 3578 | 17 | 32 | 54 | 69 |
| pwHH | 13 | 75 | 632 | 2907 | 39 | 67 | 80 | 86 |
| w1HH | 26 | 134 | 833 | 3368 | 19 | 38 | 60 | 74 |

Table 3 shows the correlations of SNP effects for weight and height between CRC I and CRC II data sets across breed types, as well as the proportion of SNPs whose effects are in the same direction. Significance analysis was evaluated separately for each breed type. For instance, the number of significant SNPs for weight in CRC I BTxBI and CRC II BTxBI were 176. This is little more than expected by chance given the number tested. However, the correlation between the effects of these 176 SNPs in the two datasets and breeds was 0.43, and in 72% of SNPs the direction of the effect was the same. If these SNPs were significant in both datasets by chance alone we would expect the correlation to be zero and 50% of the SNP effects to be in the same direction. Thus, we can conclude that the two analyses are finding some SNPs with a genuine effect on weight. Similar results were found for height (Table 3).

However, when SNPs that were significant in different breed types were compared, the correlation of SNP effects dropped to near zero and the proportion in the same direction was ~50% (Table 3). The negative correlations involving CRC I Brahman cattle could be due to the very small number of animals (78) in this breed. We conclude that the direction of SNP effects is not consistent across breed types.

A total of 75 SNPs were significant (P<0.001) for NFI giving a FDR of 68%, which means that 32% of total findings (or ~24 SNPs) are expected to be true positives. The 9 most significant SNPs were detected on BTA 3, 5, 7 and 8 with $P \leq 6.0 \times 10^{-5}$. In previous studies (Barendse *et al.* 2007; Nkrumah *et al.* 2007; Sherman *et al.* 2009), NFI QTL were identified on BTA 1, 7, 18 and 19

across the 3 studies. Barendse *et al.* (2007) also detected NFI QTL (P=0.0006) on BTA 8. Sherman and *et al.* (2009) in Canadian cattle (Angus, Charolais & composites) mapped QTLs for NFI ($P=7.6\times10^{-5}$) and DMI (P<0.001) on BTA 3 and 7, respectively. However, they found no QTL on BTA 8. This may be due to several reasons including a small number of markers, differences in breed composition across studies, G×E interaction, and even a lack of power in all studies.

**Table 3**. Validation of SNPs for body weight and hip height (P<0.05) between CRC I and CRC II datasets across breed types

| P<0.05 | $ALL_{CRC1}$ : $ALL_{CRC2}$ | $BT_{CRC1}$ : $BI_{CRC2}$ | $BT_{CRC1}$ : $BTxBI_{CRC2}$ | $BI_{CRC1}$ : $BI_{CRC2}$ | $BI_{CRC1}$ : $BTxBI_{CRC2}$ | $BTxBI_{CRC1}$ : $BI_{CRC2}$ | $BTxBI_{CRC1}$ : $BTxBI_{CRC2}$ |
|---|---|---|---|---|---|---|---|
| ***Body weight*** | | | | | | | |
| No. of SNPs | 242 | 206 | 209 | 245 | 203 | 159 | 176 |
| correlation | 0.26 | -0.10 | -0.04 | -0.21 | 0.03 | 0.07 | 0.43 |
| % of same dir | 61 | 46 | 48 | 39 | 50 | 53 | 72 |
| ***Hip height*** | | | | | | | |
| No. of SNPs | 208 | 161 | 149 | 213 | 226 | 197 | 198 |
| correlation | 0.12 | 0.00 | 0.02 | 0.09 | -0.17 | 0.10 | 0.36 |
| % of same dir | 53 | 47 | 51 | 55 | 43 | 54 | 66 |

No. = number of SNPs; correlation = correlation of corrected SNPs; and % of same dir = proportion of the direction of the effect, which was the same; ALL = all CRC I or CRCI I animals; BT = *Bos taurus*; BI = *Bos indicus*; BTxBI = crosses of *Bos taurus* and *Bos indicus*

## CONCLUSIONS

For all traits, we found more significant SNPs than expected by chance. FDR was lower in the larger population and when using a more stringent significance level. Comparing CRC I and CRC II results for body weights and hip height, we found agreement between significant SNPs effects only when the same breed type was used in both studies. This implies that the power to detect SNPs when all breed types are analysed together is reduced because the effects in different breed types are different. The SNPs that were significant for NFI could not be tested in CRC II cattle because these heifers were not measured for NFI. However, we will test them in other cattle populations. To achieve greater power in GWAS we have agreed to collaborate with scientists in North America so that a larger sample can be included in one analysis.

## ACKNOWLEDGEMENTS

We are grateful for comments from Toni Reverter, Cynthia Bottema and Bill Barendse.

## REFERENCES

Barendse, W., Reverter, A., Bunch, R.J., Harrison, B.E., Barris, W. and Thomas, M.B., (2007) *Genetics* **176**:1893.

Gilmour, A.R., Gogel, B.J., Gullis, B.R., Welham, S.J. and Thompson, R. (2002) *VSN International Ltd*. Hemel Hempstead, U.K.

Nkrumah, J.D., Sherman, E.L., LI, C., Marques, E., Crews, D.H., Bartusiak, J.R., Murdoch, B., Wang, Z., Basarab, J.A. and Moore, S.S. (2007) *Journal of Animal Science* **85**:3170.

Sherman, E.L., Nkrumah, J.D., Bartusiak, C.L.R., Murdoch, B., Moore, S.S. (2009) *Journal of Animal Science* **87**:37.

Storey, J.D. (2002) *Journal of the Royal Statistical Society, Series B* **64**:479.