Proc. Assoc. Advmt. Anim. Breed. Genet. 20:515-518

MULTI-TRAIT QTL MAPPING IN BEEF CATTLE

S. Bolormaa[1], J.E. Pryce[1], A. Reverter [3], Y.D. Zhang[4], W. Barendse[3] and M.E. Goddard[1,2]

[1]- Department of Primary Industries, Victoria; [2]- University of Melbourne, Victoria; [3]- CSIRO Animal, Food and Health Sciences, Queensland; [4]- Animal Genetics and Breeding Unit , NSW

SUMMARY

We report on the preliminary results of a multi-trait QTL mapping experiment using a genome-wide association study (GWAS) for 32 growth, feed efficiency, carcass, meat quality and reproduction traits of beef cattle. The GWAS were performed on 10,181 animals using the 800K Illumina SNP chip. The multi-trait analyses increased power to detect and map QTL. Each QTL appeared to have a pattern of pleiotropic effects across traits that was unique.

INTRODUCTION

Polymorphisms that affect complex traits or quantitative trait loci (QTL) often affect multi-traits, yet genome wide association studies (GWAS) are usually performed one trait at a time. When correlated traits are analysed independently the sampling errors tend to be correlated and this makes the interpretation of the results difficult. Also some account needs to be taken of the multi-trait testing that arises from performing many significance tests. Multi-trait analysis of linkage experiments has been reported to increase the power to detect QTL (Knott and Haley 2000; Korol *et al*. 2001). This paper investigates whether additional power can be extracted from a GWAS by analyzing traits together rather than one at a time.

Even if a QTL is detected for more than one trait in GWAS that is performed on single traits, it is possible that the most likely position for the QTL varies from trait to trait. Therefore we also consider whether multi-trait analysis can provide an increase in the precision of mapping QTL.

The obvious solution to the deficiencies of testing one trait at a time is a multi-trait analysis. However, typically not all animals have been measured for all traits and the individual animal data may not even be available. Therefore an approximate meta-analysis was used the estimates of SNP effects from individual trait GWAS.

The objectives of this study were to test a simple multivariate method to detect SNPs affecting beef traits, to understand the patterns of pleiotropic effects of genes that affect feed efficiency, growth, carcass, meat quality and fertility traits and to examine the ability of multi-trait analysis to increase the precision with which QTL are mapped.

MATERIALS AND METHODS

SNP data. In total, 729,068 SNP were genotyped. The SNP genotype data used in this study was a subset of Beef CRC genomic dataset. Details on genotyping, editing and imputation of the Beef CRC genomic data set has been described by Bolormaa *et al* (2013). A total of 10,181 animals with full genotypes and measured for at least one trait were used in this study.

Animals and population structure. The cattle were sourced from 9 different populations of 3 breed types. They include 4 different *Bos taurus* (Bt) breeds (Angus, Murray Grey, Shorthorn, Hereford), 1 *Bos indicus* (Bi) breed (Brahman cattle), 3 composite (Bt×Bi) breeds (Belmont Red, Santa Gertrudis, Tropical composites), and 1 recent Brahman cross population (F_1 crosses of Brahman with Limousin, Charolais, Angus, Shorthorn, and Hereford) (Bolormaa *et al.* 2013).

Traits. Phenotypes for 32 different traits were collated from 5 different sources including growth, feedlot, carcass, meat quality and reproduction. The trait definitions, number of records for each trait and heritability estimate and mean and its SD of each trait were reported by Bolormaa *et al.* (2013) and Zhang *et al.* (2013).

Statistical analysis. The association between each SNP and each of the traits was assessed by a regression analysis using the ASReml software (Gilmour 2009) and the following mixed model: trait ~ mean + fixed effects + SNP_i + animal + error; with animal and error fitted as random effect. Model details are given in Bolormaa *et al.* (2013) and Zhang *et al.* (2013). The effects of 729,068 SNPs were divided by their corresponding standard errors to calculate signed t values.

A multi-trait test of the effect of SNP i was conducted by storing the signed t-values for the 32 traits for SNP i in the vector $\mathbf{t_i}$. Then $\mathbf{t_i'V^{-1}t_i}$, where **V** is the correlation matrix among the SNP effects, is distributed as a chi-squared with 32 degrees of freedom under the null hypothesis that the SNP does not affect any of the traits. The correlation matrix **V** was approximated by the correlations among the estimated SNP effects across 729,068 SNPs. To avoid identifying a large number of closely linked SNPs whose association with traits is due to the same QTL, only the most significant SNP was retained from each 1Mbp interval. The most significant SNPs from the 2,523 1-Mbp-intervals were retained if it was significant at $P<10^{-4}$ and these SNPs were used to construct a new **V** matrix for use in clustering the SNPs into groups that have a similar pattern of effects on the 32 traits.

RESULTS AND DISCUSSION

In the multi-trait analysis 2,028 SNPs were significant ($P<5\times10^{-7}$), corresponding to a false discovery rate of 0.17%, and this was better than for any individual trait. When traits were analysed individually, for 29 out 32 traits the FDR was less than 2.5%. Therefore the multi-trait test did have greater power to detect QTL than the individual trait analyses. The multi-trait analysis was particularly successful in detecting QTL whose pattern of effects across traits was unusual.

Many highly significant SNPs from the multi-trait analyses were found within narrow regions on *Bos taurus* autosomal chromosomes (BTA) 3, 5, 6, 7, 14, 20 and 29 (Figure 1A). Many of the significant SNPs in both single trait and multi-trait analyses were linked and could be associated with the same QTL. When only the most significant SNPs in each Mb interval were retained, 418 SNPs were significant at $P<10^{-4}$.

A cluster analysis was performed on these 418 SNPs resulting in 12 clusters. Most clusters contained closely linked SNPs indicating that they were associated with the same QTL. Thus the long range LD that exists in cattle caused association between SNP and QTL separated by some Mb. The clustering of all SNPs in a region indicates that they all have the same pattern of effects across traits and therefore all detect the same QTL rather than multiple QTL each affecting an individual trait. However, the cluster analysis did separate the SNPs on BTA 7 into a group near 98Mb and a group near 93 Mb. The group at 98 Mb had a large effect on shear force whereas the group at 93 Mb had effects on muscling and fatness. Thus the analysis points to two separate QTL in this region of BTA 7.

The pattern of pleiotropic effects might be an important clue to the nature of the causative mutation and the function of the gene in which it occurred. Genes that operate in the same pathway might be expected to show the same pattern of pleiotropic effects. Therefore the patterns between QTL were compared to see if they fall into groups that might correspond to pathways. SNPs associated with different QTL seldom clustered together indicating that QTL seldom shared the same pattern of pleiotropic effects. However, there were some consistent patterns. For instance, SNP alleles that decreased shear force nearly always increased marbling. There was also a tendency for SNP alleles that increased hip height to increase weight and decrease fatness. However, this pattern was not consistent across all QTL.

Table 1 shows the effects of some of the significant SNPs that identify different QTL. One might describe these QTL as belonging to 3 groups. The first two QTL had a large effect on shear force and mapped to the positions of known genes affecting this trait (Calpastatin and Calpain 1).

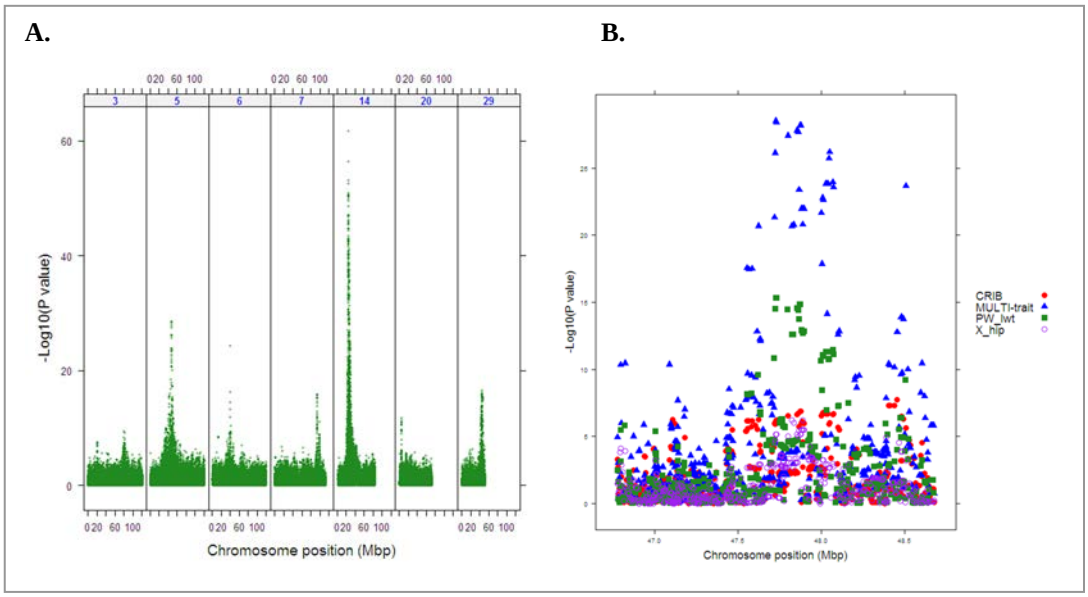

Figure 1. A. The –log$_{10}$(*P*-values) of multi-trait test based on GWAS for 32 traits on chromosomes 3, 5, 6, 7, 14, 20 and 29; B. The –log$_{10}$(*P*-values) of single SNP regressions for 3 traits and the multi-trait test on a region of chromosome 5.

The third SNP appeared to mark an unknown gene affecting shear force on BTA 6. The next 3 QTL could be affecting mature size. The allele of these 3 QTL which increased height also increased weight and decreased fatness. The last 3 SNP primarily affected fatness.

However, even QTL that have a similar pattern of pleiotropic effects, show differences in the detail of this pattern. For example, of the 3 'mature size' QTL, the one on BTA 14, which is presumably PLAG1 (Fortes *et al*. 2013), was the only one of the 3 mature size QTL that decreased shear force. It also had more marked effects on blood IGF concentration, fatness and reproduction than the other two SNPs. On the other hand, the QTL on BTA 5 had an unusual pattern of effects in that it redistributes fat from the P8 site to the rib and intramuscular depots. This QTL map was close to the gene HMGA2, which contains polymorphisms affecting growth, fatness and fat distribution in humans, mice and pigs (Anand and Chada 2000; Kim *et al*. 2006;Voight *et al*. 2010).

Table 1 also shows 3 SNPs associated with effects on marbling or intramuscular fat. There was a tendency for alleles that increase marbling to increase subcutaneous fat depth but this was not consistent. The QTL on BTA 7 had little effect on subcutaneous fat depth but a large effect on retail beef yield; the QTL on BTA 3 increased weight as well as fat; and the QTL on BTA 10 decreased shear force.

Based on these limited results, it would appear that each QTL has its own pattern of effects. Thus we have failed to discover groups of QTL that belong to the same pathway except for calpain and calpastatin. This could be explained if genes exist in a network rather than in pathways. Then each gene has a unique position in the network and therefore a unique pattern of effects.

The precision with which a QTL can be positioned on the genome in a GWAS is limited by two sources of errors. Firstly, the LD between SNP markers and the QTL is highly variable and therefore the nearest SNP is not necessarily the one in greatest LD with the QTL. Secondly, the

LD with the QTL is not observed directly but only via the effect of the QTL on a phenotypic trait.

Table 1. **Effect of some of significant multi-trait SNPs in the individual traits (signed t-values >1 are shown)**

chr	position	shear force	p8 fat depth	rib fat depth	intra-muscular fat	retail beef yield	weaning wt	IGF at weaning	weaning hip height	age at puberty in BB*	age at puberty in TC*
7	98540675	-8.6	1.1	1.4	1.5		1.7	2.1			
29	45778237	-10.5	2.9	2.5	4.1		-1.9				
6	68101121	-6.6	1.7	2.9	2.9	1.4			1.3	2.2	
5	47727773	1.9		-5	-4.4	-2	8.1	-1.9	9.6	1.2	3.3
6	40093712	1.7	-1.9	-2.6	-2.5		8.1	-2.6	9.5	2	1.1
14	25015640	-2.3	-7	-4.1	-1.6		9.8	-7.6	10.9	6.3	3.5
7	93007435	-2.5	-1.4		-3.2	6.3	2	-1.7	-2.4		2.9
3	80105316		1.1	3	2.5	-2	1.6				
10	89027305	-5.8			3.9		1.3	1.3			

* = Age at first detected corpus luteum in BB and TC

Because the QTL typically only explains a small amount of the variance of the trait , the effect of a SNP on the trait is estimated with error and this can also cause a SNP that is not the nearest to the QTL, to have the largest effect. By using more than one independent trait to map the QTL, the second source of error can be reduced but not the first source. Figure 1B shows the significance of SNPs from the multi-trait analysis and for 3 single trait GWAS in a region of BTA 5. The 3 separate traits map the QTL to slightly different positions and the multi-trait analysis may represent a good compromise.

ACKNOWLEDGEMENTS

We gratefully acknowledge the funding from the former Cooperative Research Centre for Beef Genetic Technologies for this study and ARC project DPI093502.

REFERENCES

Anand A. and Chada K. (2000). *Nat. Genet*.**24**:377.

Bolormaa S., Pryce J.E., Kemper K., Savin K., Hayes B.J., Barendse W., Zhang Y., Reich C.M., Mason B.A., Bunch R.J., Harrison B.E., Reverter A., Herd R.M., Tier B., Graser H.-U. and Goddard M.E. (2013) (submitted to *J. Anim. Sc.*).

Fortes M.R.S., Kemper K., Sasazaki S., Reverter A., Pryce J.E., Barendse W., Bunch R., McCulloch R., Harrison B., Bolormaa S., Zhang Y.D., Hawken R.J., Goddard M.E., and Lehnert S.A. (2013) *Anim. Genet.* (accepted)

Gilmour, A. R., B. J. Gogel, B. R. Cullis, and R. Thompson. (2009) ASReml User Guide Release 3.0 VSN International Ltd, Hemel Hempstead. HP1 1ES, UK.

Kim K.S., Lee J.J., Shin H.Y., Choi B.H., Lee C.K., Kim J.J., Cho B.W. and Kim T.H. (2006) *Anim. Genet.* **37**:419.

Knott S. A. and Haley C. S. (2000) *Genetics*. **156**:899.

Korol A.B., Ronin Y.I., Itskovich A.M., Peng J. and Nevo E. (2001). *Genetics*. **157**:1789.

Zhang Y.D., Johnston D.J., Bolormaa S., Hawken R.J. and Tier B. (2013) (submitted to *Anim. Prod. Sc.*).

Voight B.F., Scot L.J., *et al*. and McCarthy M.I. (2010) *Nat. Genet*.**42**:579.