Fine Mapping QTL And Candidate Genes Discovery For Residual Feed Intake On Chromsomes 5, 15, 16, And 19 In Beef Cattle.

M.K. Abo-Ismail[*], E. J. Squires[*], K. C. Swanson [*], D. Lu[*], Z. Wang[†], J. Mah[†], G. Plastow[†], S. Moore[†] and S. P. Miller[*].

Introduction

Feed efficiency remains a trait of great importance to beef producers and breeding values can be estimated for selection purposes when the trait is measured (Schenkel *et al*. (2004)). However in practice selection is limited due to the cost of measurement making this trait a candidate for selection with molecular genetic information (Dekkers and Hospital (2002)). High density SNP panels are becoming more powerful and providing encouraging results in beef and dairy cattle (Hayes *et al*. (2009)). When prediction across breeds is required as is often the case in beef cattle, Marker Assisted Selection (MAS) based on direct or functional mutations increases the genetic gain (Dekkers (2004)). Different whole genome association studies to detect putative quantitative trait loci (QTL) for feed efficiency have been conducted (e.g. Sherman *et al*. (2010)). A whole genome scan (WGS) with the Illumina BovineSNP50 Beadchip has identified a large number of regions associated with residual feed intake (RFI) (results not published). Information from the WGS can be coupled with an understanding of the biological function underlying genes to further enhance the list of candidate genes for fine mapping.

The objectives of this study were: (1) to map QTL for RFI (2) estimate the proportion of variation in RFI explained by the detected QTL and (3) identify the candidate gene or the flanking genes to elucidate the biological meaning behind the underlying potential QTL.

Material and methods

Animals and phenotypic data. Phenotypes were obtained from detailed studies of growth and feed intake during the post-weaning period on 589 heifers (20), steers (338), and bulls (231). Animals were crossbred, with breed composition formed by Angus (39%), Charolais (11%), Piedmontese (9%), Simmental (25%), and Limousin (2%). Animals originated from one of three University of Guelph cooperating herds in Ontario, Canada. Calves were weaned at approximately 200 days of age, and then transported to the Elora Beef Cattle Research Centre for various post-weaning nutrition trials. Average daily gain (ADG), metabolic midpoint (MWT) and individual feed intake were determined. RFI was calculated

[*] Centre for Genetic Improvement of Livestock, University of Guelph, Canada.
[†] Faculty of Agricultural, Life and Environmental Sciences, University of Alberta, Canada.

from the difference between observed dry matter intake (DMI) and the DMI being modeled on ADG, MWT.

Genotypic data. Genotyping was performed at the Alberta Bovine Genomics Laboratory, University of Alberta. The genotypic data is as described by (Lu *et al*. (2009)). Briefly, 56,947 SNP were genotyped then quality control procedure removed SNP with minor allele frequency less than 10% and SNP that were not in Hardy-Weinberg, resulting in 34,976 SNP across 29 *Bos taurus* autosomal chromosomes (BTA). In this study the statistical analyses were performed on 1,897 significant SNPs ($P<0.05$) from a previous whole genome scan study using single marker regression model on the 34,976 SNPs (results not published).

Statistical analyses. The phenotypes of RFI were adjusted for the fixed effects, trial treatment year (44 levels), heard year (25 levels), sex, heterosis, age at end of test and the proportion of SM CH LM PI and GV breed except breeding values using ASREML (Gilmour *et al*. (2000)). This model was modified to take in the account the polygenic effect to estimate breeding values.

Bayes analysis. Missing genotypes were inferred using fastPHASE (Scheet and Stephens (2006)) then BayesCπ was used to calculate π which is the proportion of SNPs having no effect on the phenotype. BayesB was then used to run the analysis again with the same set of SNPs (1,897) and the estimated π to estimate SNP effects, the SNP effect variance and the error variance. A SNP was declared to be significant based on Posterior Probability of Association (PPA), as follow ($PPA/(1-\pi) > 100$) (Thomas *et al*. (2010)). BayesCπ and BayesB were implemented in R scripts fully described by Fernando *et al*. (2009). The proportion of additive genetic variance explained by detected QTL was estimated using the function of additive genetic variance explained by each locus $\sigma^2_g = 2pq\alpha^2$ (Falconer and Mackay 1996).

Functional analysis. The significant SNPs (807 out of 1,879 SNPs) from BayesB were mapped to the bovine genome to characterize the gene or nearest gene, and then submitted to DAVID 6.7 Beta software (Huang *et al*. (2009)) for functional analysis. Only 30% of the gene list was annotated, therefore *Homo sapiens* orthologs were used to obtain 100% annotation. The identified list of bovine genes was also mapped to biological pathways using web based software in KEGG (Kanehisa, et al. (2010)).

Results and discussion

The estimated π was 16.54%. Using this prior in BayesB produced the genomic breeding values highly correlated to the estimated breeding values from ASREML (0.96) indicating the set of SNPs (1879) from WGS described most of the additive genetic effect in these data. The BayesB analyses identified 807 SNPs in significant association with RFI. This is supported by previous studies where many QTLs were found for RFI (Sherman *et al*. (2010)). One hundred SNPs with the highest effect were selected from the 807 significant SNPs, and mapped on the bovine genome to identify corresponding genes or flanking genes. One hundred thirty-three genes were found and submitted to KEGG tool which identified pathways for 8 genes, which were located on BTA5, 15, 16, and 19. Twenty-one out of the 100 SNPs mentioned earlier were located on these 4 chromosomes, and corresponded to 25

genes (Table 1). These genes were then entered to DAVID to determine their biological processes. Estimated allele substitution effects, PPA, SNP information, and identified genes for the 21 SNPs on chromosome 5, 15, 16, and 19 are shown in Table 1.

Table 1: Effect of significant SNPs influencing RFI on chromosomes 5, 15, 16, and 19 and their corresponding gene

BTA	SNP BPPos[1]	Estimate	MAF	PPA/ (1-π)	Gene Location Relative to SNP	Gene BankID[2]
5	7126361[a]	0.0319	0.274	1.117	280459 bp at 5' side	526129
5	7571767[a]	0.017	0.24	1.063	725865 bp at 5' side	526129
5	91627311	0.0217	0.421	1.068		537804
5	100844909	-0.0199	0.135	1.074	70480 bp at 3' side	533844[d]
5	100844909	-0.0199	0.135	1.074	5727 bp at 5' side	514797[d]
5	101180226	0.0188	0.43	1.088	55763 bp at 5' side	786521[e]
5	101180226	0.0188	0.43	1.088	24592 bp at 3' side	512775[e]
5	111005281	0.0204	0.371	1.066	4363 bp at 3' side	616055[f]
5	111005281	0.0204	0.371	1.066	10304 bp at 5' side	513297[f]
5	111719827	0.0169	0.344	1.06	9866 bp at 5' side	282211[g]
5	111719827	0.0169	0.344	1.06	4262 bp at 3' side	523151[g]
15	21302682[b]	-0.0328	0.409	1.128	80709 bp at 5' side	529873
15	21644938[b]	-0.0337	0.275	1.133	205517 bp at 5' side	529873
15	35764005	0.0172	0.309	1.066	40847 bp at 3' side	539898[h]
15	35764005	0.0172	0.309	1.066	10219 bp at 5' side	616568[h]
15	36198436	0.0203	0.389	1.1		519803
15	66610045[c]	-0.0179	0.378	1.075	2983 bp at 5' side	514579[k]
15	66610045	-0.0179	0.378	1.075	74360 bp at 3' side	522802[k]
15	66638381[c]	-0.0233	0.376	1.091		514579
16	35848283	0.0731	0.204	1.166		523479
16	57431452	-0.031	0.259	1.121	13353 bp at 3' side	100048947[l]
16	57431452	-0.031	0.259	1.121	987 bp at 5' side	509065[l]
16	60070864	-0.0224	0.383	1.088		521326
19	5192954	-0.0256	0.46	1.11		538147
19	34420867	0.0185	0.327	1.081		526652
19	49345756	0.0291	0.422	1.125		617397
19	51346200	-0.0162	0.423	1.071		507056
19	63407392	0.0278	0.369	1.1137		539501

[1] SNPs with same superscript correspond to the same gene.
[2] Genes with same superscript correspond to the same SNP.

Listed in Table 2 are the biological processes found through functional analysis, and related to feed efficiency. These results are in agreement with the proposed major biological processes contributing to the variation in RFI by Richardson and Herd (2004).

Table 3: Identified biological processes associated with detected QTL

Biological process	No. genes	P	Biological process	No. genes	P
Cell structure	8	0.020	Purine metabolism	2	0.41
Cation transport	8	0.060	Proteolysis	6	0.45
mRNA transcription initiation	3	0.077	MHCI-mediated	6	0.47
mRNA end-processing and	2	0.096	immunity		
stability			Pre-mRNA processing	6	0.51
Lipid, fatty acid and steroid	3	0.178	Fatty acid metabolism	2	0.52
metabolism	4	0.185	Cholesterol metabolism	2	0.55
Protein glycosylation			Protein biosynthesis	3	0.58
Mitosis	3	0.193	mRNA transcription	6	0.66
Protein modification	5	0.225			
Protein targeting	2	0.352			

The additive genetic variance explained by identified QTLs on Chr. 5, 15, 16, and 19 represented a genomic heritability estimate of 0.05 explaining 13.4% of the total additive genetic variance as estimated from the 1,879 SNP using BayesB.

Conclusion

From this study, we report 21 SNPs are significantly associated with RFI where they are located in or close to genes involved in metabolic process and other some biological functions. These genes carry biological meaning that might cause the genetic variation in RFI. Further research is required to investigate these potential candidate genes.

References

Dekkers, J. C. M. (2004). *J. Anim. Sci*., 82: 313-328.

Dekkers, J., and Hospital, F. (2002). *Nat. Rev. Genet,*. 3(1): 22-32.

Falconer, D.S., and Mackay, T.F.C. (1996). 4th ed. Longman Group, Essex, UK.

Fernando,R.,Garrick,D. and Dekkers, J. (2009)http://www.ans.iastate.edu/stud/courses/short/

Gilmour, A.R., Gogel, B.J., Cullis, B.R. *et al.* (2009). *ASReml User Guide Release 3.0.*

Hayes, B. J., Bowman, P. J., Chamberlain, A. J. *et al.* (2009). *J. Dairy Sci.,* 92:433-443.

Huang, D.W., Sherman, B.T. and Lempicki, R.A. (2009). *Nat. Protoc.,* 4(1):44-57.

Kanehisa, M., Goto, S., Furumichi, M. *et al.* (2010). *Nucl. A. Res.,* 38: 355-360.

Lu, D., Sargolzaei, M., Kelly, M. *et al*. (2009). *J. Anim. Sci.,* 87, E-Suppl. 2: 129.

Richardson, E. C., and Herd. R. M. (2004). *Aust. J. Exp. Agric.* 44 (4-5): 431-440

Scheet, P. and Stephens M. (2006). *Am. J. Hum. Genet.,* 78:629-644.

Schenkel, F.S., Miller, S.P., and Wilton, J.W. (2004). Can. J. Anim. Sc., 84:177-185.

Sherman, E. L., Nkrumah, J. D., and Moore, S. S. (2010). *J. Anim. Sci.* 88:16–22.

Thomas, M.G., Peters, S.O., Kazilkaya, K. et al. (2010). P & A Genome XVIII Conf. P537.