Proc. Assoc. Advmt. Anim. Breed. Genet. 20:372-375

GENOME-WIDE EPISTASIS ASSOCIATION OF ULTRASOUND-SCANNED CARCASS TRAITS IN BEEF CATTLE: TWO-STAGE MODELS

A. A. Ali[1,2], M. S. Khatkar[1] , H. N. Kadarmideen[3] and P. C. Thomson[1]

[1] ReproGen - Animal Bioscience Group, Faculty of Veterinary Science, University of Sydney, Camden NSW 2570, Australia, [2] CSIRO Animal, Food and Health Sciences, PMB PO Aitkenvale, Townsville, QLD 4814, Australia, [3] Faculty of Health and Medical Sciences, University of Copenhagen, Grønnegardsvej 7, 1870 Frederiksberg C, Denmark

SUMMARY

Most genome-wide association studies (GWAS) on complex traits in livestock have focused on identifying single-locus effects ignoring any epistatic interaction. Here we consider analytical methods that explicitly look for statistical interactions between two loci. Two-stage models were used to ease multiple testing problems and computational demand using beef cattle data as an example. The results suggest that fitting epistasis models for GWAS using two-stage models is a useful strategy for detecting significant interactions between genetic loci and may help in searching for candidate genes and polymorphisms influencing phenotypic variation.

INTRODUCTION

Genome-wide association studies (GWAS) are now routinely used to identify genomic regions associated with traits of interest; however, this ignores an important class of genomic associations, that of epistatic interactions (Carlborg and Haley 2004; Hemani *et al*. 2013). Identifying genome-wide interactions among SNPs (single nucleotide polymorphisms) using high-density SNP chip genotypes is a difficult task due to statistical complexity (e.g. multiple testing) and computational burden (Marchini *et al*. 2005; De Lobel *et al*. 2010). For example, consider the current Bovine SNP Chip which comprises of more than 50,000 SNPs, the number of SNP combinations would be 1.25×10^9 for testing two SNPs at a time; this analysis could take several days or even weeks on a standard workstation. The number of possible interactions involving more than two loci will be exponentially higher. For these reasons, epistasis is not yet a standard tool in complex trait studies. Rather than testing all possible pair-wise comparisons, a more practical strategy might be to examine a subset of SNPs which could have influence on a trait of interest. Here we show a two-stage approach for analysing genome-wide epistasis association (GWEA) using real ultrasound scan measures for carcass traits on beef cattle.

MATERILS AND METHODS

Animals and data: Animals used were part of a northern Australian breeding project of the Co-operative Research Centre for Beef Genetic Technologies. A total of 583 heifers of Brahman breed were ultrasound-scanned for eye muscle area (EMA, cm^2), rump fat depth at P8 site (P8, mm) and rib fat depth measured the between 12^{th} and 13^{th} ribs (RIB, mm). The fixed environmental effects recorded were age (in days), month of calving, herd, and cohort (combination of experimental location and heifer's year of birth). The details of the resource population and the data were described previously (Barwick *et al*. 2009; Bolormaa *et al*. 2011). Animals were genotyped using 10,000 ParaAllele/Affymetrix SNP chips (Khatkar *et al.* 2007). After quality control, 565 animals representing 51 sire families and 6,715 SNPs on bovine autosomal chromosomes (BTA) were analysed. The SNP positions were mapped to the genome

assembly UMD 3 (http://www.cbcb.umd.edu/research/bos_taurus_assembly.shtml). Missing SNP genotypes were imputed using Beagle (Browning and Browning 2007). Genotypes were coded as 0 for the homozygote for allele (A), 1 for the heterozygote (AB), and 2 for the homozygote for allele (B).

Additive association model: SNP effects were estimated by a single-trait-single-SNP association analysis. The additive effect of a SNP on each trait was estimated from its associated regression coefficient. The traits and SNP were fitted using the following linear mixed:

$$\text{Trait} = \mu + \text{fixed\_effect} + \beta N_{SNP} + \text{Animal} + \varepsilon$$

where Trait is the phenotypic measurement for EMA, P8 or RIB trait, μ is the overall mean, fixed_effects were significant fixed effects specific to each trait (described in Results), N_{SNP} is the number of copies of the allele (0, 1, and 2) at the SNP as a covariate, Animal is the polygenic effect of animals to account for the effect of relatedness, and ε is the random error. Random animal and residual effects were assumed to be normally distributed with zero mean, and additive genetic variance σ_A^2 and residual variance σ_ε^2.

Epistasis association model: To ease computational burden, a subset of significant SNPs at a lenient threshold of *p*-value ≤ 0.01 was selected from the additive association model, At this threshold, 82 SNPs for EMA, 86 SNPs for P8, and 92 SNPs for RIB were available and analysed for two loci (e.g. $SNP_1 \times SNP_2$) epistasis using the following linear mixed model:

$$\text{Trait} = \mu + \text{fixed\_effect} + SNP_1 + SNP_2 + SNP_1 \times SNP_2 + \text{Animal} + \varepsilon$$

where SNP_1 and SNP_2 are three-level factors for genotypes (e.g. AA, AB, and BB) at SNP_1 and SNP_2, where $SNP_1 \times SNP_2$ in the interaction between SNP_1 and SNP_2 genotypes as an indicator for epistasis effect. Note that while the SNP effect for the additive association model was treated as a covariate (to maximise power of association detection), for estimating interactions it is necessary to treat the effects of SNPs as factors. A separate model was fitted for each pair of SNPs for all pair-wise combination of selected SNPs. Other terms in the model were the same as the additive association model. To account for multiple testing, false discovery rates was estimated for GWEA results using the *q*-value package in R 2.15 version (Storey and Tibshirani 2003). All analyses were performed using a REML procedure in ASReml-R package (Butler *et al*. 2009).

RESULTS AND DISCUSSIONS

Summary statistics and estimates of heritability for ultrasound scan measures of eye muscle area and fat depths traits are presented in Table 1. The effects of age, month of calving, herd and cohort were significantly associated with phenotypes considered.

Table 1. Summary of ultrasound scan measures for eye muscle area (EMA, cm^2), rump fat depth at P8 site (P8, mm) and fat depth between 12^{th} and 13^{th} ribs (RIB, mm) in Brahman cattle

Trait	No. animals	Mean	SD	Min	Max	Heritability
EMA	564	43.8	6.6	28.0	64.0	0.42 ± 0.15
P8	565	3.7	3.7	1.0	12.0	0.48 ± 0.15
RIB	565	1.9	1.0	1.0	6.0	0.65 ± 0.17

SNP association: Results of the GWAS analyses for ultrasound scan carcass related traits are presented in Figure 1. Using a *q*-value threshold of 0.05 which corresponds to *p*-value $\leq 1 \times 10^{-6}$, a total of six SNPs on BTA 5 and 14 were identified as being significantly associated with the ultrasound scan measures of fat depths traits, most of these SNP were located on BTA 14 at 22 to 24 Mb. The most significant SNP ($p = 4.9 \times 10^{-11}$; $q = 3.3 \times 10^{-7}$; MAF = 0.498) was rs29020688 on BTA 14 at 24 Mb for fat depth at P8. This SNP was located in the intronic region of *Bos taurus* XK, Kell blood group complex subunit-related family, member 4 (XKR4) gene. Recently, the XKR4 was identified as a candidate gene affecting rump fat depth in Australian tropical cattle (Bolormaa *et al*. 2011). Other significant SNPs for fat depth traits were rs29010515, rs29010516 on BTA 14 and rs29010471 and rs29026420 on BTA 5.

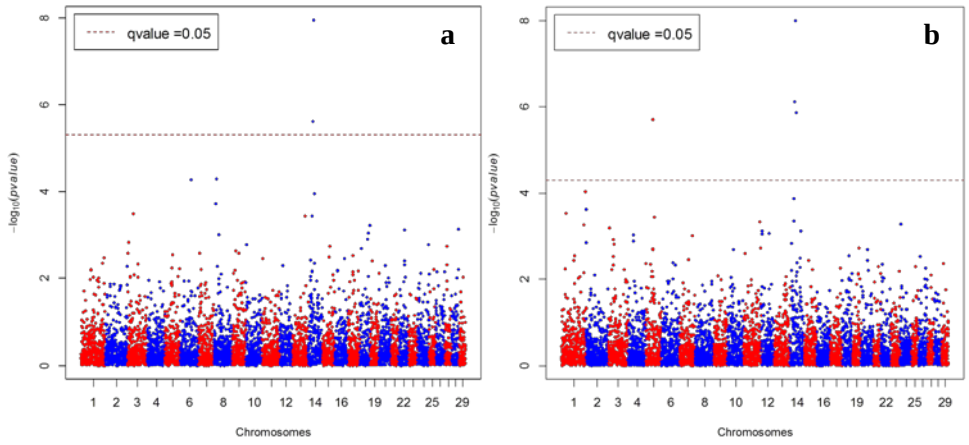

Figure 1. Manhattan plots of GWAS scan, a significant threshold (dashed red line) is drawn corresponding to an FDR of 0.05. a: rump fat depth at P8 site. Note a significant association on chromsome 14. b: rib fat depth. Note the significant associations on chromosome 5 and 14.

Epistasis association: Image plots in Figure 2 depict graphical representation of GWEA results for ultrasound scanned fat depths at P8 site and RIB traits. These heat maps show epistatic signals or so called 'hot spot' (red colour spots) where significant epistatic association were detected at genome-wide level. The colour gradient in the panel on the right side of the plot represents $-\log_{10}$ (*p*-values). Evidence of epistatic association was detected on several chromosomes; however, the strongest epistatic signals were for pairs of SNPs on BTA 8 and 12, BTA 8 and 14, and BTA 8 and 15 for fat depth at P8 site trait. Using a *q*-value of 0.10 which corresponds to a *p*-value of 10^{-4}, significant epistatic interactions between SNPs in DDX56 (BTA 4) and EFHD2 (BTA 16), MAP3K5 (BTA 9) and TMEM132D (BTA 12) genes were detected to be associated with scanned RIB fat depth, whereas an interaction between SNPs in ROBO2 (BTA 1) and DZANK1 (BTA 13) genes was associated with P8 site fat depth. These epistatic genes could be detected only after fitting epistasis models. Once significant gene-gene epistatic interactions are identified, this will facilitate to define networks of interacting genes that can be incorporated into existing functional annotation and molecular pathways, and hence provide the genomic basis of improving carcass traits efficiently in beef cattle.

CONCLUSIONS

This study demonstrated the usefulness of modelling epistasis in the analysis of complex traits, and identified potential candidate genes affecting carcass traits using field data from an Australian Brahman cattle population. Information about epistasis can add to our understanding of the complex genomic networks that form the fundamental basis of biological systems.

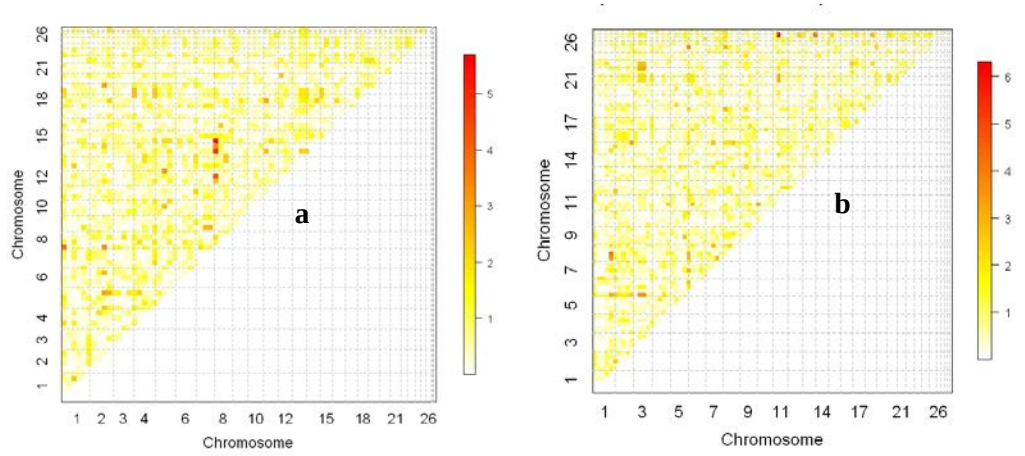

Figure 2. Heatmap image of genome-wide epistasis association. The heatmap legend scale (right side) is on $-\log_{10}$(*p*-value) scale. a: fat depth at P8 site. b: fat depth at RIB. Note red spots indicate epistatic signals.

ACKNOWLEDGEMTNS

We acknowledge the CRC for Beef Genetic Technologies for providing the data. A. A. Ali was the recipient of a CSIRO Animal, Food and Health Sciences PhD scholarship.

REFERENCES

Barwick S.A., Johnston D.J., Burrow H.M., Holroyd R.G., Fordyce G., Wolcott M.L., Sim W.D. and Sullivan M.T. (2009) *Anim. Produc. Sci.* **49**: 367.

Bolormaa S., Neto L.R.P., Zhang Y.D., Bunch R.J., Harrison B.E., Goddard M.E. and Barendse W. (2011) *J. Anim. Sci.* **89**: 2297.

Browning S.R. and Browning B.L. (2007) *Am. J. Hum. Genet.* **81**: 1084.

Butler D.G., Cullis B.R., Gilmour A.R. and Gogel B.J. (2009) Tech. Report QLD DPI.

Carlborg O. and Haley C.S. (2004) *Nat. Rev. Genet.* **5**: 618.

De Lobel L., Geurts P., Baele G., Castro-Giner F., Kogevinas M. and Van Steen K. (2010) *Eur. J. Hum. Genet.* **18**: 1127 .

Hemani G, Knott S, Haley C. (2013) *Plos Genet.* **9**: e1003295.

Khatkar M.S., Zenger K.R., Hobbs M., Hawken R.J., Cavanagh J.A., Barris W., McClintock A.E., McClintock S., Thomson P.C. , Tier B., Nicholas F.W., Raadsma H.W. (2007). *Genet.* **176**:763

Marchini J., Donnelly P. and Cardon L.R. (2005) *Nat. Genet.* **37**: 413.

Storey J.D. and Tibshirani R. (2003) *Proc. Nat. Acad. Sci.* **100**: 9440 .