

Genome-Wide Scan For Fertility Traits In Cattle Using Single-Marker Association And A Genealogy Based Method

N.F. Schulman^{*†}, G. Sahana^{*}, J. Vilkki[†], B. Guldbrandtsen^{*} and M.S. Lund^{*}

Introduction

Fertility of high yielding dairy cows has been declining in recent years (e.g. Washburn *et al.* 2002). As fertility traits have low heritability, are difficult to measure, and have negative correlations with production traits (Pösö and Mäntysaari 1996) they are difficult to improve using traditional breeding methods. Genetic markers in linkage disequilibrium with fertility traits of interest could be used in marker assisted selection or could be included in genomic selection approaches (Meuwissen *et al.* 2001) to improve genetic progress of these traits. Single-marker association has been a widely used mapping method for dense marker sets especially in human studies (e.g. The Wellcome Trust Case Control Consortium 2007). However, this method does not take into account the dependencies of close markers. Including the structure of dependencies could increase the amount of information in the analysis and improve power and precision of the mapping study, especially when the minor allele frequency is low. One way of including this information is to use local genealogy around each marker (Mailund *et al.* 2006). The objective of this study was to detect associations between single nucleotide polymorphism (SNP) markers and female fertility traits of dairy cattle using two different mapping methods, single-marker association and a genealogy based method (GENMIX) (Sahana *et al.* 2010a)(submitted).

Material and methods

Animals and genotypes. A total of 1054 Finnish Ayrshire bulls were genotyped using the Illumina BovineSNP BeadChip (Illumina Inc., San Diego). The bulls had a half-sib family structure, 39 families with ten or more sons and 65 families with less than 10 sons. The total number of markers available for analyses after quality control was 37,002. The marker order was based on the Btau_4.0 assembly. FastPHASE (Scheet and Stephens 2006) was used to construct the haplotypes.

Traits and phenotypes. We analyzed seven different female fertility traits, non-return rate for heifers (NRRH) and cows (NRRC), time from first to last insemination in days for heifers (IFLH) and cows (IFLC), number of inseminations for heifers (AISH) and cows (AISC) and time from calving to first insemination in days for cows (ICF). The heritabilities of these traits are low (www.nordicebv.info) and they are moderately to highly

^{*} Aarhus university, Faculty of Agricultural Sciences, Department of Genetics and Biotechnology, 8830 Tjele, DK

[†] MTT Agrifood Research Finland, Biotechnology and Food Research, 31600 Jokioinen, Finland

genetically correlated (Fogh *et al.* 2003). The phenotypes used were de-regressed estimated breeding values (EBV) from a single trait mixed model. The EBVs were provided by Nordic Cattle Genetic Evaluation and de-regressed with the Interbull de-regression program (www.interbull.slu.se).

Single-marker analysis. The linear model tested associations between phenotypes and marker genotypes by successively fitting single SNPs into the following additive model.

$$y_{ij} = \mu + S_i + bg_{ij} + e_{ij}$$

where y_{ij} is the phenotype of bull j , belonging to half-sib family i ; μ is the population mean; S_i is the effect of i -th half-sib family; g_{ij} is number of copies of allele 1 of the SNP (corresponding to 0, 1 or 2 copies) carried by individual j of the i -th half-sib family; b is the allele substitution effect; and e_{ij} is the random residual. The analyses were done in statistical package R (R Development Core Team 2009). The significance of the allele substitution effect was tested with a t-test against a null hypothesis of $b = 0$.

GENMIX. The markers were grouped into haplo-groups and the observed haplotypes were clustered based on local genealogies forming a tree (Mailund *et al.* 2006). The tree was split at three different levels. Successively each clustering of haplotypes was included as a fixed effect in the model for analysis as below.

$$y_{ij} = \mu + S_i + Q_{ij}^1 b^1 + Q_{ij}^2 b^2 + e_{ij}$$

where Q_{ij}^1 and Q_{ij}^2 are row design vectors that identify the two haplotype-clusters carried by individual j of the i -th sire family and b^1 and b^2 are vectors of haplotype-cluster effects. Remaining terms are as described earlier. This model was compared to the null model using the 'ANOVA' function in statistical package R.

Significance. In the single-marker method Bonferroni corrections (5%) for total number of markers or number of markers on a chromosome were calculated to get genome-wide (GW) ($-\log_{10}P=5.87$) and chromosome-wide (CW) significance thresholds ($-\log_{10}P= 4.67$ to 4.13). In GENMIX number of SNP/haplotype effects ($7 * \text{number of markers}$) was used.

Results and discussion

With the single-marker association method we detected a total of 52 GW-significant marker trait associations on five different chromosomes (1, 6, 8, 12, and 14) (Table 1). In addition to these there were 193 CW-significant associations on 21 chromosomes. However, 14 of these chromosomes had only one to three associated SNPs which were not close to each other. These are likely to be false positive results.

On BTA1 we detected four SNPs associated with IFLC at GW-significance using single-marker association. The significant SNP spanned a 9.8 Mb region. Two of the SNPs were also associated with AISC at CW-significance. This may reflect the high genetic correlation between these two traits. There was in addition seven other SNPs within this peak area associated with IFLC at CW-significance. The GENMIX method also detected the same peak position at GW-significance. This could be the same QTL area detected by Sahana *et*

al. (2010b) in Holstein cattle with the strongest association at 136 Mb. Schulman *et al.* (2008) found a QTL in Finnish Ayrshire close to microsatellite BMS4014 which is located 27 Mb from the present peak SNP. As the confidence intervals for linkage analysis are large this could be the same QTL area.

On BTA6 we found eight SNPs showing GW association with NRRH, NRRC, AISH, and AISC using single-marker association. Two of them were associated with both NRRC and AISC. The area with CW-significant SNPs was between 25 and 49 Mb for NRRH, 8 and 52 Mb for NRRC, 32 and 49 Mb for AISH and 32 and 61 Mb for AISC. The GENMIX method also detected the strongest association at 37 Mb for NRRC and another area of GW association at 34 Mb for NRRH, NRRC, and AISH. This could be the same QTL area detected by Sahana *et al.* (2010b) around 55 Mb. The strongest association in the present study was at 37 Mb. A possible candidate gene for this QTL is the osteopontin gene located at 37511674 to 37518672 bp on Btau 4.0. This gene has a role in normal oviduct physiology in cattle (Gabler *et al.* 2003).

On BTA8 we detected two GW-significant SNP associated with NRRC at 17 Mb. The GENMIX method had a peak at the same position but reached only CW-significance as the significance thresholds are higher due to more tests done with this method. On BTA12 we found the strongest association of this study at 24 Mb for IFLC. There was a total of 19 SNP associated with NRRC, IFLH, IFLC, and AISC with GW-significance and many of them were associated with more than one trait. The peak area was broad, for example for IFLC the area of chromosome-wide association spanned from 10 to 84 Mb, but the highest peak was still sharp. The GENMIX method showed the strongest associations at 23 and 29 Mb and the peak area was broad also with this method. The same QTL area was earlier detected in a smaller separate sample of Finnish Ayrshire (manuscript in preparation).

On BTA14 we detected one SNP associated with IFLC at GW-significance. In addition there were four other CW-significant SNP in this area and they were less than 1 Mb apart. GENMIX detected this QTL area at CW-significance. In addition a QTL area associated with NRRC and AISC at 60 Mb was detected at CW-significance with this method. This could be the same QTL close to microsatellite *BM4513* (61 Mb) earlier detected in Finnish Ayrshire using linkage mapping (Schulman *et al.* 2008).

In addition to the chromosomes with genome-wide associations, there were areas showing chromosome-wide association on BTA17, 21, and 27. On BTA17 two areas were detected with both methods at 28 and 63 Mb associated with AISH and IFLC respectively. On BTA21 a QTL area was detected with both methods around 45 to 57 Mb associated with NRRC, IFLH, IFLC, and AISC. On BTA27 both methods detected a QTL area associated with ICF at 40 and 18 Mb associated with NRRH. These could be the same QTL detected by Cobanoglu *et al.* (2005) affecting twinning rate close to BMS2116 (39 Mb) and BMS2650 (16 Mb). The latter one could also be the same area detected in a smaller separate sample of Finnish Ayrshire with the strongest association at 22 Mb (manuscript in preparation).

In this genome scan the two mapping methods gave very similar results. GENMIX detected two QTL areas that were missed with single-marker association. In some cases the strict Bonferroni correction for number of tests was leading to lower power of the GENMIX method and some areas detected by single-marker association did not cross the GW-threshold with this method. It is possible that the relative small sample size of this experiment was the reason why the genealogy based method did not outperform single-marker association (Mailund *et al.* 2006).

Table 1: Genome-wide significant peak markers from single-marker association

| BTA | Rs/ss-number | Position | -log10P | Trait |
|-----|--------------|-----------|--------------|-----------|
| 1 | ss86292428 | 129818221 | 8.31 | IFLC |
| 6 | rs29012386 | 24997938 | 5.93 | NRRH |
| 6 | ss86312899 | 37317020 | 7.99, 6.48 | NRRC,AISC |
| 6 | ss117968366 | 33444497 | 6.81 | AISH |
| 8 | rs43545116 | 16835602 | 5.97 | NRRC |
| 12 | ss105241146 | 28721291 | 7.48 | NRRC |
| 12 | rs41622464 | 22305690 | 6.72 | IFLH |
| 12 | rs41617617 | 24477221 | 10.98, 10.56 | IFLC,AISC |
| 14 | ss86338007 | 72796829 | 5.99 | IFLC |

Conclusion

We detected strong associations between female fertility traits and dense SNP markers on five different chromosomes (1, 6, 8, 12, and 14). The QTL identified by single-marker analyses were supported by GENMIX. Detecting a QTL by two divergent methodologies gives confidence about the QTL presence.

Acknowledgements. The work was funded by the Finnish Ministry of Agriculture and Forestry (grant 2886/502/2008). The genotypes were funded by the Danish Directorate for Food, Fisheries and Agri Business, VikingGenetics, Nordic Genetic Evaluation, and Aarhus University (grant 3412-08-02253).

References

- Cobanoglu O., Berger P.J., and Kirkpatrick B.W. (2005). *Anim. Gen.*, 36:303-308.
- Fogh A., Roth A., Pedersen O.M., Eriksson J-Å., et al. (2003). *Interbull bulletin*, 31:52-55.
- Gabler, C., Chapman, D.A., and Killian, G.J. (2003). *Reprod.*, 126:721-729.
- Mailund, T., Besenbachner, S., and Schierup, M.H. (2006). *BMC Bioinf.* 7:454-475.
- Meuwissen, T.H.E., Hayes, B.J., and Goddard, M.E. (2001). *Genetics* 157:1819-1829.
- Pösö, J., and Mäntysaari, E. (1996). *Livest. Prod. Sci.* 46:41-48.
- R Development Core Team (2009). *Manual, Vienna, Austria*, <http://www.R-project.org>.
- Scheet, P., and Stephens M. (2006). *Am. J. Hum. Genet.* 78:629-644.
- Sahana, G., Mailund, T., Lund, M. S. *et al.* (2010a). (submitted to WCGALP 2010).
- Sahana, G., Guldbrandtsen, B., Bendixen, C. *et al.* (2010b). *Anim. Genet.* (accepted).
- Schulman, N.F., Sahana, G., Lund, M.S. *et al.* (2008). *Genet. Sel. Evol.*, 40:195-214.
- Washburn, S.P., Silva, W.J., Brown, C.H. *et al.* (2002) *J. Dairy Sci.*, 85:244-251.
- Wellcome Trust Case Control Consortium (2007). *Nature* 447:661-683.

