Whole Genome Association Study In Admixed European Braunvieh Cattle

J. Maxa[*], M. Neuditschko[†], M. Förster[*] and I. Medugorac[*]

Introduction

The current European population of Braunvieh cattle (BBV) is mainly spread over the Alpine regions of Austria, Germany, Italy and Switzerland where Original Braunvieh cattle (OBV) were developed. After the first export of OBV from Europe to the USA in 1869 (Yoder and Lush (1937)), a severe population bottleneck and a following stronger selection on dairy characteristics, the American Brown Swiss cattle breed was established in the USA. From the 1970's, the American Brown Swiss were used for crossing with the OBV population in Europe, resulting in the creation of BBV population. This should be taken into account in genome-wide association studies of BBV cattle because population structure can lead to spurious associations between a candidate marker and a phenotype as mentioned in several studies (Pritchard *et al.* (2000); Hirshorn and Daly (2005); Zhang *et al.* (2009)). As well as population structure based on selection, also familial relationships can result in false associations (Yu *et al.* (2006)). Until now, most of association studies in dairy cattle were focusing on Holstein breed and only few have been implemented in the minor but still cosmopolitan breeds such as BBV (Bagnato *et al.* (2008)).

Therefore, the main objective of this study was to set up a whole genome association study in BBV cattle, considering familial relationships within a structured population. Associations were analysed for the main production traits milk yield (MY), milk protein percent (PP), milk protein yield (PY), milk fat percent (FP) and milk fat yield (FY).

Material and methods

Animals and phenotypes. A total of 591 BBV and 42 OBV blood, sperm or hair root samples of progeny tested bulls were included in this study. Sampling of OBV with null proportion of the American Brown Swiss cattle was performed across Swiss and German populations. Sampled BBV originated mainly from Germany and partially from Austria, Italy and Switzerland. All BBV bulls included in this study were tested in Germany and Austria. Estimated breeding values of the BBV bulls for milk production traits MY, PP, PY, FP and FY together with reliability values of EBV were obtained in 2009 from the joint Austria-Germany genetic evaluation of the BBV cattle (Emmerling *et al.* (2002)). For further analyses, EBV were weighted by their reliability values in order to correct for differences in accuracy of EBV.

[*] LMU Munich, Chair of Animal Genetics and Husbandry, 80539 Munich, Germany
[†] Tierzuchtforschung e.V. Munich, 85586 Grub, Germany

Genotypic information. Single Nucleotide Polymorphisms (SNPs) genotyping was carried out by Tierzuchtforschung e.V. München using commercial Bovine BeadChip featuring 54,001 SNPs (BovineSNP50). Furthermore, markers with genotyping errors, less informative in BBV (MAF<0.05) and unknown chromosomal position as well as markers deviating from HWE in sample of unrelated animals (P<0.01) were excluded. This resulted in a total of 35,805 autosomal SNPs that passed quality control and were used for further analyses. Additionally, a smaller set of 5,661 highly informative SNPs was constructed by LD-based SNP pruning (r^2<0.225) and later used for structure analysis.

Statistical analyses. STRUCTURE software version 2.3.1 with the admixture model and correlated allele frequencies (Pritchard *et al.* (2000)) was used to evaluate the extent of substructure of BBV considering the OBV population. An initial 20,000 burn-ins and 100,000 iterations for parameter estimation were sufficient to assure convergence of parameter estimates.

Test of putative associations among genotypes (SNPs) and production traits in BBV cattle was carried out using ASREML package (Gilmour *et al.* (2009)). Traditional mixed model framework integrating genomic tools to uncover population structure and familial relationships was applied. Population structure was incorporated as independent variable and based on previous estimates from STRUCTURE software. Furthermore, we replaced the pedigree based coancestry matrix with marker based genome-wide identity by descent (IBD) matrix as described in Yu *et al.* (2006). The following mixed model was assumed:

$$y = \mathrm{X}\beta + G\alpha + Sq + \mathrm{Z}u + e$$

where: $\boldsymbol{y}$ is a vector of weighted estimates of EBV on the trait of interest; $\boldsymbol{\beta}$ is a vector of fixed effect of birth year; α is a gene content effect (Gengler *et al.* (2007)); $\boldsymbol{q}$ is a vector of population effects; $\boldsymbol{u}$ is a vector of random polygenic effects for each individual; $\boldsymbol{e}$ is a vector of residual effects; **X**, **G**, **S** and **Z** are incidence matrices for the effects in $\boldsymbol{\beta}$, α, $\boldsymbol{q}$ and $\boldsymbol{u}$, respectively. Significance levels of putative associations among markers and phenotypes were determined using the Wald F statistic (Kenward and Roger (1997)).

Results and discussion

Population structure of both BBV and OBV cattle based on STRUCTURE analysis is presented in figure 1. Individuals are represented by a single vertical column divided into two colours where each colour represents one subpopulation. Clearly visible is a highly varying proportion of OBV population in BBV which can affect results of association studies if these are not adjusted.

Application of the mixed model approach, together with adjustment for population structure and unequal relatedness among individuals, in order to test associations between genetic marker and phenotype was described by Yu *et al.* (2006). The method demonstrated improved control of both type I and type II error rates over other methods. The results of the genetic structure as well as pedigree and genome-wide IBD analyses confirmed strong population structure and familial relatedness occurred in our mapping population. Therefore corrections for both were applied in the mixed model for whole genome association study.

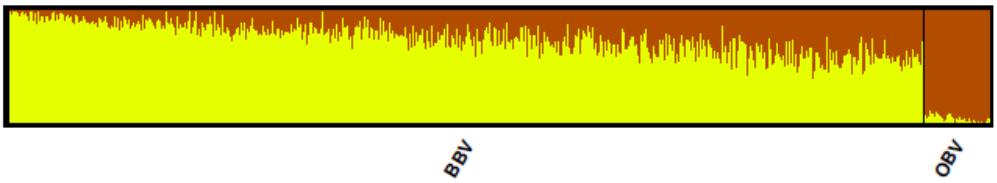

Figure 1: Graphical representation of the estimated membership fractions of individuals of European Braunvieh (BBV) and Original Braunvieh (OBV) populations

In figure 2 results of FP on *Bos Taurus* autosome 14 (BTA14) are presented as an example. Highly significant associations among markers and FP were found in the very proximal region of BTA14. As already reported in other studies (Grisart *et al.* (2002); Winter *et al.* (2002)) the *DGAT1* gene cause a major effect on milk fat content and other milk characteristics in Holstein and other cattle. To verify highly significant associations found on BTA14 in BBV, we performed combined linkage disequilibrium and linkage (LDL) mapping in this chromosome region, resulting in QTL detection and confirmation of the correctness of the association study.

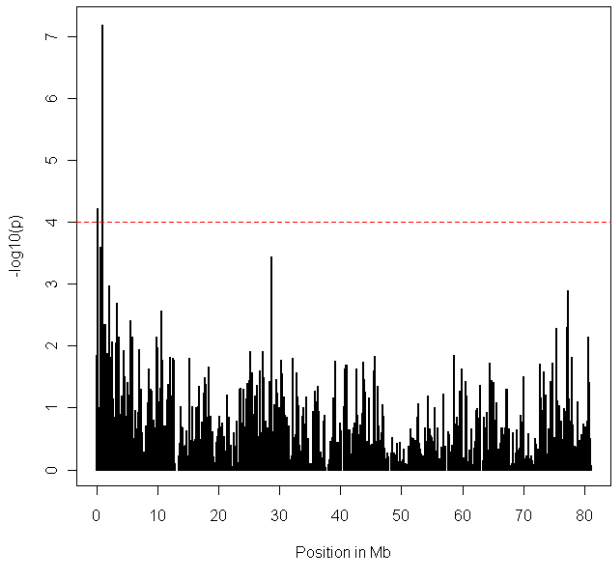

Figure 2: Profile of -log10(p)-values for milk fat percent (FP) on BTA14 with threshold significance on the horizontal dashed line

We used association and LDL mapping on BTA14 only as confirmation example. In addition to this and taking into account a *log*P≥4 we compared models with and without adjustment for population structure. The results showed that model omitting population structure effect

detected 11% significant associations more than the model accounting for stratification. There is cumulating information that the use of all SNPs instead of subset of the most significant can improve accuracy of genomic selection. In this study the largest differences in *log*P-values (0.5 to 1.1) between both models were observed at moderate 1.85 *log*P-value (i.e. P~0.014). These substantial differences at loci with moderate association could have an effect on genomic selection which needs further investigations. Over all 34, 26, 15, 5 and 4 SNPs were associated (*log*P≥4) with FP, PP, MY, PY and FY, respectively. Especially BTA05 followed by BTA06 and BTA14 showed the highest appearance of significant markers.

Conclusion

In this study we applied the mixed model approach for genome wide association study, simultaneously adjusting for population structure and familial relationships – both of these being detected in BBV cattle. Results showed significant associations with several milk production traits especially on BTA05, BTA06 and BTA14, indicating a potential for QTL fine-mapping in the associated autosomal regions.

Acknowledgements

The authors are grateful for funding from Tierzuchforschung e.V. Munich and Deutsche Forschungsgemeinschaft (DFG), project ME3404/2-1. Furthermore, we thank the breeders and breeding associations who sent us samples free of charge to support this study: Arbeitsgemeinschaft Süddeutscher Rinderzucht- und Besamungsorganisationen e. V. and Swissgenetics.

References

Bagnato, A., Schiavini, F., Rossoni, A. *et al.* (2008). *J. Dairy Sci.*, 91:767-783.

Emmerling, R., Lidauer, M., and Mäntysaari, E. A. (2002). *Interbull Bull.*, 29:111-117.

Gengler, N., Mayeres, P., and Szydlowski, M. (2007). *Animal*, 1:21-28.

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

Grisart, B., Coppieters, W., Farnir, F. *et al.* (2002). *Genome Res.*, 12:222-231.

Hirschhorn, J. N., and Daly, M. J. (2005). *Nat. Rev. Genet.*, 6:95-108.

Kenward, M. G., and Roger, J. H. (1997). *Biometrics*, 53: 983-997.

Pritchard, J. K., Stephens, M., and Donnelly, P. (2000). *Genetics*, 155:945-959.

Winter, A., Krämer, W., Werner, F. A. *et al.* (2002). *Proc. Natl. Acad. Sci. (U.S.A.)*, 99:9300-9305.

Yoder, D. M., and Lush, J. L. (1937). *J. Hered.*, 28:154-160.

Yu, J., Pressoir, G., Briggs, W. H. *et al.* (2006). *Nat. Genet.*, 38:203-208.

Zhang, Z., Buckler, E. S., Casstevens, T. M. *et al.* (2009). *Brief. Bioinform.*, 10:664-675.