Genome-wide association study identifies two major loci affecting calving ease and growth related traits in cattle

Hubert Pausch*, Krzysztof Flisikowski*, Simone Jung*, Reiner Emmerling§, Christian Edel§, Kay-Uwe Götz§, Ruedi Fries*

* Lehrstuhl für Tierzucht, Technische Universität München, 85354 Freising, Germany
§ Institut für Tierzucht, Bayerische Landesanstalt für Landwirtschaft, 85586 Poing, Germany
Running head: Calving ease and growth QTL in cattle

Key words: cattle, calving ease, fetal growth, QTL, genome-wide association

Corresponding author:

Ruedi Fries
Lehrstuhl fuer Tierzucht
Technische Universitaet Muenchen
Liesel-Beckmann-Strasse 1
85354 Freising
Germany
Tel: 004981613229
Fax. 004981617103
Email: ruedi.fries@tum.de
ABSTRACT

Identifying quantitative trait loci (QTL) underlying complex, low heritability traits is notoriously difficult. Prototypical for such traits, calving ease is an important breeding objective of cattle (*Bos taurus*) improving programs. To identify QTL underlying calving ease, we performed a genome-wide association study using estimated breeding values (EBVs) as highly heritable phenotypes for paternal calving ease (pCE) and related traits. The massively structured study population consisted of 1800 bulls of the German Fleckvieh (FV) breed. Two pCE–associated regions on bovine chromosomes (BTA) 14 and 21 ($P = 5.72 \times 10^{-15}$ and $P = 2.27 \times 10^{-8}$, respectively) were identified using principal components analysis to correct for population stratification. The two most significantly associated SNPs explain 10% of the EBV variation. Since marker alleles with negative effect on pCE have positive effects on growth-related traits, the QTL may exert their effects on the birthing process through fetal growth traits. The QTL region on BTA14 corresponds to a human chromosome (HSA) region that is associated with growth characteristics. The HSA region corresponding to the BTA21 pCE QTL is maternally imprinted and involved in the Prader-Willi and Angelman syndromes. Re-sequencing of positional candidate genes on BTA14 revealed a highly significant ($P = 1.96 \times 10^{-14}$) associated polymorphism ablating a polyadenylation signal of the gene encoding ribosomal protein S20 (*RPS20*). Our study demonstrates the leverage potential of EBVs in unravelling the genetic architecture of lowly heritable traits.
INTRODUCTION

The recent availability of genome-wide SNP panels in cattle and other livestock species enables the mapping of quantitative trait loci (QTL) as well as the prediction of an animal's genetic merit without relying on phenotypic information (Goddard and Hayes 2009). However, the complex genetic architecture of agriculturally important traits renders the systematic identification and characterization of individual QTL a difficult task. The proportion of trait variance explained by an average QTL is very small. First mapping results in cattle seem to validate the classical quantitative genetic model of a large number of loci of small additive effects (Cole et al. 2009), (Barendse et al. 2007), (Daetwyler et al. 2008) and agree with findings from mapping QTL in the human genome (Manolio et al. 2009). Besides the relative contribution of a QTL to the trait variation, the heritability of the trait is a major determinant of the mapping power (Goddard and Hayes 2009).

The heritability of calving traits, i.e. traits that describe the birthing process (dystocia in the case of difficulties) and the perinatal viability (stillbirth) of the calf as affected by the birthing process, are low, ranging from 0.04 to 0.15 (Steinbock et al. 2003), (Lin et al. 1989), (Seidenspinner et al. 2009). Calving traits are of considerable economic importance due to veterinary treatment costs, calf loss and lower production of cows affected by dystocia. Estimated breeding values (EBVs) for calving traits are used as selection criteria in attempts to reduce calving problems both in dairy and beef breeds (e.g. (Van Tassell et al. 2003) (Freer 2008)). Calving traits are complex since they are influenced by a sire-effect through the size of the calf as well as dam-effects consisting mainly of the pelvic dimensions. Routine progeny testing results in highly reliable EBVs for calving traits and thereby boosts the heritability to levels that make them amenable to QTL mapping even with medium-sized samples.

An important prerequisite for unbiased QTL mapping based on linkage disequilibrium (LD) is homogeneity of the mapping population (Devlin and Roeder 1999). The
heavy use of genetically superior bulls, facilitated by artificial insemination, and introgression lead
to massively stratified populations. We attempted to correct for population stratification by principal
components analysis (PCA) - based approaches that have been successfully in human genome-wide
QTL mapping (PRICE ET AL. 2006).

Here we report the mapping of two loci affecting very low heritability calving traits
in a heavily structured dual purpose (dairy, beef) cattle population. The mapping approach was
facilitated by the use of EBVs and consequent correction of population stratification.

MATERIAL AND METHODS

Animals and phenotypes
Bulls of the dual purpose breed FV (n = 1829) were genotyped using the Illumina BovineSNP 50K
Bead chip® comprising 54,001 single nucleotide polymorphisms (SNPs). Phenotypes in the form of
EBVs for beef (daily gain (DG)) and conformation traits (body size (BS)) as well as functionality
traits such as paternal calving ease (pCE) and paternal stillbirth incidence (pSB) were obtained from
the Bavarian State Research Center for Agriculture (www.lfl.bayern.de/bazi-rind) (November 2009
version, Table S1). Breeding value estimation was based on best linear unbiased prediction (BLUP)
animal model. The calving process is described by a score ranging from 1 (unassisted delivery) to 4
(surgical delivery, fetotomy). Stillbirth is recorded as categorical trait (alive or not 48 hours post
partum). Paternal and maternal effects on calving ease and stillbirth incidence are estimated
multivariately for the first versus later parities. Parity specific EBVs are combined to produce
paternal and maternal EBVs, respectively.

Genotypes and quality control
Of 1829 genotyped FV animals, six were excluded from further analyses due to genotype call rates below 90%. The remaining samples exhibited an average genotyping rate of 99.14%. 549 SNPs were omitted because their chromosome position was not known. 728 SNPs were discarded because genotyping failed in more than 10% of animals, 8480 SNPs were excluded due to a minor allele frequency smaller than 1% and 810 SNPs showed a significant (\( P < 1 \times 10^{-3} \)) deviation from the Hardy-Weinberg-equilibrium, indicating genotyping errors, and were thus not considered for further analyses. Pairwise identity by descent (IBD) was calculated based on identity by state (IBS) information derived from the remaining 43,863 SNPs using the method-of-moments approach implemented in PLINK (Purcell et al. 2007). IBD relationship of each pair of animals was compared with the corresponding pedigree relationship calculated using the PyPedal package (Cole 2007). Comparison of the marker with the pedigree relationship revealed several inconsistencies, most likely resulting from mislabeling of DNA samples and false relationships. Unresolved inconsistencies led to the exclusion of 23 animals (Figure S1). The final set consisted of 1800 animals.

The phenotype and genotype data are available from the authors upon request.

**Single marker analysis**

Single marker analysis was first carried out without considering population stratification. The EBVs were regressed on the number of copies of one of the alleles as implemented by the PLINK --assoc option. Quantile – quantile plots of the expected vs. the observed P values were inspected for an inflation of small P values indicating false positive association signals due to a structured population. The genome-wide inflation factor was computed according to Devlin and Roeder (Devlin and Roeder 1999).

We next applied a PCA-based approach, implemented in the EIGENSOFT 3.0 package (Price et al. 2006), for eliminating false positive association signals due to ancestry.
differences and resulting population stratification. One SNP of a pair in LD with $r^2 > 0.25$ was
excluded using the PLINK --indep-pairwise option (500 SNP window, shifted at 50 SNP intervals).
A smartpca version of the EIGENSOFT 3.0 package (compiled from source code with
modifications for the bovine chromosome complement) was ran on the pruned data set consisting of
20,000 autosomal SNPs with the following option: the value of each marker is replaced with the
residuals from a multivariate regression without intercept on the 5 preceding markers to further
reduce redundancies due to LD. Eigenvalues ($\lambda$) and eigenvectors were calculated for all axes of
variation. Correlation of ancestry adjusted EBVs and genotypes was calculated using the previously
obtained eigenvectors with a smarteigenstrat version of EIGENSOFT 3.0, compiled for the bovine
chromosome complement. The resulting test statistic is equal to $(N - K - 1)$ times the squared
correlation and $\chi^2$– distributed, where $N$ is the number of samples and $K$ the number of axes with an
eigenvalue that amounts to at least 70% of the mean eigenvalue (Jolliffe's criterion, (JOLLIFFE 2002))
used to adjust for ancestry (PRICE ET AL. 2006). Quantile – quantile plots were inspected and the
genomic inflation factors were calculated (see above) to judge the extent of false positive signals.
SNPs were considered as significantly associated for $P$ values below the 5% Bonferroni-corrected
type I error threshold for 43,863 independent tests. Allele substitution effects were estimated for
each significant marker in a linear regression model implemented in R (www.r-project.org) with
axes of variation with $\lambda \geq 0.7$ as covariables.

Haplotype analysis
Haplotypes for each chromosome region with significant association signals were reconstructed
using default parameters in fastPHASE (SCHEET AND STEPHENS 2006) and inspected by means of
bifurcation plots obtained with sweep (SABETI ET AL. 2002) to visualize recombination events and to
define the length of haplotypes. The resulting haplotypes were analyzed for association in a
multilinear regression model implemented in R (see above).
Estimating the power of the genome-wide association study

According to Goddard and Hayes (2009) the correlation (r) between marker and trait, \( r_{m,t} \), is equal to \( r_{m,q} r_{q,g} r_{g,t} \), with m representing the marker genotype, q the QTL, g the genotypic value and t the phenotypic value (EBV) of an animal. \( r_{m,q}^2 \) measures the LD between marker and trait, \( r_{q,g}^2 \) the variance explained by the QTL and \( r_{g,t}^2 \) the reliability of the EBV. Using this equation and the formula for the standard error of the correlation coefficient the number of animals (N) required for identifying a QTL can be calculated as follows:

\[
N = \left( \frac{1 - r_{m,t}^2}{r_{m,t}^2 \frac{1}{z^2 (1 - \alpha)}} \right)^2
\]

where z is the normal score and \( \alpha \) the Bonferroni-corrected type I error rate for 43,863 independent tests. Assuming a reliability of the EBV of 0.9, a LD between marker and QTL of \( r^2 = 0.35 \) and the QTL to explain 4% of the genetic variance, the required number of animals amounts to about 1700. Thus the power of our study with \( N = 1800 \) should allow to identify QTL explaining at least 4% of the genetic variance using EBVs of high reliability.

Annotation and polymorphism analysis of candidate genes

The GENOMETHREADER software tool (Remme et al., 2005) was used to predict the genomic structure and localization of the candidate genes based on the University of Maryland UMD3.1 assembly of the bovine genome sequence (Zimin et al., 2009) and the Dana-Farber Cancer Institute bovine gene index release 12.0 (Quackenbush et al., 2001) together with the annotated RNA sequences of the UMD3.1 assembly (Zimin et al., 2009). The GENOMETHREADER output was viewed and edited using the Apollo sequence annotation editor (Lewis et al., 2002). The exons and
the promoter regions of the candidate genes were PCR-amplified (the primers are listed in Table S2) and re-sequenced in 12 FV bulls with specific genotypes for the SNP with the most significant signal for the pCE EBV (BTA14 – ARS-BFGL-NGS-104268), i.e. in one bull with GG and in 11 bulls with AG genotypes.

**Genotyping of candidate gene polymorphisms**

Genotypes of selected SNPs were determined by TaqMan® genotyping assays (Applied Biosystems Applera, Darmstadt, Germany). DNA samples were available for 810 FV animals only. Candidate gene polymorphisms were genotyped in these animals, and the genotypes of the remaining 990 animals of the study population were inferred using the EM algorithm implemented in *fastPHASE*.

**RESULTS**

**Single marker analysis**

In a first attempt to identify QTL for pCE, we applied a linear regression model that did not account for the covariance of related animals. This model yielded 1146 autosomal SNPs exceeding the significance threshold and a genome-wide inflation factor of 4.75. However, an apparent association signal was observed on chromosome 14 (P: $1.64 \times 10^{-55}$; Figure S2). The inflation of significant association signals most likely results from relatedness of animals leading to massively structured population. The 1800 bulls within our study descend from 234 sires and 328 maternal grand sires. The paternal half sib families and the maternal grand sire families encompass up to 81 and 137 members, respectively. This is manifested by an average coefficient of relationship of 0.047 and distinct clusters of related animals (Figure S3A and S3B). Recent introgression of HF into FV can be uncovered by PCA. A 50% HF sire was broadly used within the FV population in the early
1980s to improve milk performance and udder quality of cows. Of 1800 FV bulls within the study population, 1050 exhibit HF ancestry via two of his sons (both 25% HF), as can be visualized by contrasting the top two axes of variation of the PCA (Figure S3C). Thus, HF admixture and the paternal and maternal sire families lead to a massively structured study population and concomitant inflation of significant association signals.

Therefore, the association study was repeated, now correcting for population stratification using a PCA-based approach implemented in the EIGENSOFT 3.0 package. The correction was based on 773 axes of variation that met the Jolliffe's criterion. In addition to the highly significant association with the pCE EBV on chromosome 14, that was already observed in the analysis without correction for stratification, the PCA-based analysis now also revealed significant association on chromosome 21 (Figure 1A). The Q-Q plot (Figure 1B) and an inflation factor of 0.97 document that the PCA-based analysis successfully eliminated association artifacts resulting from population stratification.

Eight SNPs on chromosome 14 and three SNPs on chromosome 21 meet the Bonferroni-corrected significance threshold (Table 1). Of the eight significant SNPs on chromosome 14, six lie within a 1.4 Mb interval (from 24.06 Mb to 25.4 Mb). Two significant SNPs outside this interval are in LD ($r^2 = 0.48$ and 0.68) with the most significantly associated SNP on chromosome 14. Three significantly associated SNPs in high LD define a second pCE QTL region on chromosome 21 (2.15 Mb to 2.39 Mb). While the minor allele of the significant SNPs on chromosome 14 has a negative effect on the pCE EBV, it is the major allele of the significant SNPs on chromosome 21 that lowers the pCE EBV. The most significant SNP on chromosome 14 exhibit an allele substitution effect of -7.01, corresponding to 58% of the standard deviation of the EBV. The substitution effect of the major allele of the most significant marker on chromosome 21 is -2.93, i.e. 24% of the standard deviation of the EBV (Figure 2A).

pCE is highly correlated with the paternal stillbirth incidence (pSB) as well as with
growth related EBVs such as for DG and BS (Table S3). Consequently, association signals can also
be observed for these EBVs, particularly on chromosome 14 (Table 1, Figure S4). The QTL alleles
that lower the pCE and pSB EBVs have a positive effect on the growth related EBVs.

There are several chromosome regions showing suggestive association (\(P < 1 \times 10^{-3}\),
Table S4), most prominently a second region on chromosome 14 with 5 SNPs located between 58.3
and 59.3 Mb.

Haplotype analysis

Haplotype analysis was carried out for the associated regions on chromosome 14 and 21 in an
try to delineate the chromosomal segment carrying the pCE QTL. On chromosome 14, the
allele that lowers the pCE EBV could be pinpointed to a specific haplotype that spans 1.58 Mb
(starting at 23.82 Mb) and encompasses 23 SNPs (Table 2). This haplotype version occurs in a
frequency of 10\% in the study population. Its negative effect on the pCE EBV (\(P = 1.56 \times 10^{-16}\)) is
more prominent than of any of the associated SNPs (-0.66 \(\sigma_A\) vs. -0.62 \(\sigma_A\), Figure 2B). This is a
strong indication for the causal variant lowering the pCE EBV to exclusively reside on this
haplotype version.

On chromosome 21, the associated SNPs are contained within a haplotype spanning
0.6 Mb (starting at 1.78 Mb). The most frequent haplotype version occurs in a frequency of 66%
and has a negative effect on the pCE EBV (\(P = 3.15 \times 10^{-7}\)). However it explains less of the genetic
variance than the most significant SNP does (-0.18 \(\sigma_A\) vs. -0.24 \(\sigma_A\)).

Identification and analysis of candidate genes

The assessment of the transcriptional content of the pCE EBV associated regions was based on the
UMD3.1 assembly and annotation of the bovine genome (Zimin et al., 2009). The 23.82 – 25.40 Mb
interval on chromosome 14 encompasses 13 transcripts / genes (Figure 3A). The associated region
on bovine chromosome 14 is conserved in human chromosome 8q21 which has been shown to be associated with adult height (Gudbjartsson et al. 2008). Since adult stature is positively correlated with fetal size and fetal size is an important determinant of the birthing process, we considered PLAG1, MOS, CHCHD7, RDHE2 (alias: SDR16C5), RPS20, LYN, TGS1, PENK, as proposed by Gudbjartsson (Gudbjartsson et al. 2008) as positional and functional candidate genes for the pCE QTL in cattle. Of this list, PLAG1, TGS1, RPS20 and LYN together with SOX17, another gene in the critical region that we considered a functional candidate, were re-sequenced in a panel of 12 animals of our study population. Totally, we screened 30.3 kb resulting in the detection of 48 polymorphisms (Table S5). We decided to genotype four putatively functional SNPs, located in SOX17 (ss250608762), RPS20 (ss250608720, ss250608721) and TGS1 (ss250608741), in 810 animals and analyzed the association with the pCE EBV in the complete study population using genotype imputation (Figure 3B & 3C). Only ss250608721 produced a highly significant signal (P = 1.96 x 10^{-14}). The polymorphism affects a polyadenylation signal of a cistron encompassing the genes for the ribosomal protein S20 (RPS20, a ribosomal component) and the small nucleolar RNA U54 (SNORD54, a ribosomal RNA modifying RNA) (Figure 4).

The association signals on chromosome 21 result from the most proximal region on the chromosome (Figure S5). The region contains, among other transcripts, those encoding SNURF-SNRPN and UBE3A. These two transcripts are encoded in the human chromosome interval 15q11-15q13 that is subject to imprinting. The lack of a functional paternal copy of 15q11-15q13 causes the Prader-Willi syndrome, while the lack of a functional maternal copy of UBE3A is implicated in the Angelman syndrome (Horsthemke and Wagstaff 2008). The SNURF-SNRPN mRNA is derived from a single large transcriptional unit of which more than 70 snoRNAs of the C/D box type are processed (Bachellerie et al. 2002). Preliminary BLAST-analyses indicate the presence of a snoRNA cluster in the proximal region of bovine chromosome 21. However, a systematic annotation has not been attempted. The lack of detailed knowledge of the genomic organization, the
imprinting status and transcriptional content of the associated region on chromosome 21 precluded
the analysis of candidate genes, although a functional implication of the region in fetal growth and
thus pCE seems obvious when considering that fetal growth retardation is symptomatic for the
Prader-Willi syndrome.

DISCUSSION

Our genome-wide association study based on a dense SNP marker map provides strong evidence for
two QTL on chromosome 14 and 21, respectively, that together explain at least 10% of the variation
of the pCE EBV in the German FV breed. The two QTL also explain a substantial fraction of the
pSB EBV as well as of EBVs of postnatal growth such as DG and BS. Stillbirth can be considered
as the dichotomic manifestation of the calving ease score, as dystocia is a major cause of perinatal
mortality. The correlation of pCE with growth-related traits and the coincident QTL point to fetal
growth and the resulting birth weight as major determinant for the ease of delivery (Meiering 1984)
(Johanson and Berger 2003). Thus, the two QTL might primarily affect fetal growth. One could
expect that they would explain a larger fraction of the genetic variation of birth weight, a trait that is
not routinely measured in dairy cattle. Improving postnatal growth along with lactation traits is a
major breeding objective of the FV breed. This dual purpose selection is likely to act on the two
QTL identified in our study. Animals known to carry favorable alleles for the chromosome 14 and
21 QTL could now be more stringently selected with regard to beef traits. However, the
identification of QTL that either affect prenatal or postnatal growth but not both would facilitate to
efficiently improve postnatal beef performance without antagonistically compromising calving ease.
In any case, conventional selection schemes seem to allow favorable selection responses for calving
ease and postnatal growth despite the genetic antagonism (MacNeil 2003) (Bennett 2008) (Bennett
A key factor for successfully mapping of a QTL for a complex trait with very low heritability such as pCE was the use of reliably estimated breeding values for calving traits. If one assumes a heritability of 0.08, a LD between marker SNPs and QTL of \( r^2 = 0.35 \) and 4% of the genetic variation explained by the QTL, one would require approximately 20,000 individuals for the successful identification of a QTL (see Material and Methods). Using EBVs with a reliability of 90%, \( i.e. \) a quasi heritability of 0.9, requires less than 1800 animals to detect association. Breeding values are routinely estimated for many traits and are thus indispensable for dissecting complex trait variation in livestock species.

Another key factor for successfully mapping the two QTL was careful correction for extensive relationship among the study animals. The adjustment along 773 axes of variation allowed to account for major as well as for more subtle relationships that can possibly not be revealed by pedigree analyses. The association signal on chromosome 21 became apparent only when population structure was corrected for. Thus, PCA based elimination of false positive association signals might enable the detection of QTL with a smaller impact on the trait variation that would otherwise be “buried” in the false positive signals. Suggestive signals (\( P < 1 \times 10^{-3} \), Table S4) are thus more likely to represent real QTL.

Our findings about two highly significant QTL for pCE as well as about additional suggestive QTL are supported by several previous studies on calving ease and growth trait QTL, based on microsatellite marker analyses. Kneeland et al. (KNEELAND ET AL. 2004) identified three regions on chromosome 14 to affect birth weight in a composite breed. The proximal region from 26.0 to 26.7 cM most likely corresponds to the highly significant QTL region identified in our study, the more distal region between 36.2 and 46.2 cM may corroborate a suggestive QTL region resulting from our study. Davis et al. (DAVIS ET AL. 1998) also identified a QTL affecting birth weight at 42 cM. Koshkoih et al. (KOSHKOIH ET AL. 2006) provide additional evidence for two birth weight
QTL on chromosome 14 at 26 and 50 cM, respectively, in a cross of Limousin and Jersey animals. Maltecca et al. (Maltecca et al. 2009) recently identified a birth weight QTL at 19 cM on chromosome 14 in a Jersey – Holstein cross. There are also reports on QTL for postnatally measured growth traits in Wagyu (Mizoshita et al. 2004) (Takasuga et al. 2007) and a Jersey – Limousin cross (Morris et al. 2002), indirectly supporting our suggestive evidence for a secondary pCE QTL on chromosome 14. Casas et al. (Casas et al. 2003) and Davis et al. (Davis et al. 1998) identified a QTL for birth weight in the very proximal region of chromosome 21 in crosses of Brahman with Hereford and Charolais, respectively, providing supportive evidence for the pCE QTL identified in the present study.

There is also support in the literature for suggestive QTL on other chromosomes: Olsen et al. (Olsen et al. 2009) and Holmberg and Andersson-Ecklund (Holmberg and Andersson-Ecklund 2006) identified in a Swedish and Norwegian dairy cattle population, respectively, a dystocia / stillbirth QTL at 36 – 37 cM on chromosome 6. We observe a suggestive pCE QTL at about 40 Mb on chromosome 6. Gutierrez-Gil et al. (Gutierrez-Gil et al. 2009) identified a fetal growth / birth weight QTL in the same region based on a Charolais – Holstein cross. Eberlein et al. (Eberlein et al. 2009) provide evidence for the gene (NCAPG) encoding the Non-SMC Condensin I Complex, Subunit G to encompass this QTL, also based on a Charolais – Holstein cross. However, a prominent calving ease QTL in the Holstein breed on chromosome 18 (Cole et al. 2009) could not be detected in the present study or is not segregating in the Fleckvieh breed.

A preliminary candidate gene analysis identified a highly significantly pCE-associated SNP in a cistron encoding a ribosomal protein (RPS20) and a internally nested small nucleolar RNA (SNORD54). The SNP affects a polyadenylation site. Alternative polyadenylation at tandem poly(A) sites yield transcripts with different 3’ UTR sequences providing the potential of differential regulation of mRNA expression by RNA binding proteins and / or miRNAs (Sandberg et al. 2008) (Licatalosi and Darnell 2010). The marker allele causing the gain of an upstream
polyadenylation signal is associated with a lower pCE EBV, *i.e.* a higher incidence of calving difficulties. This is hypothetically compatible with a shorter and more highly expressed mRNA encoding ribosomal components, leading to a higher ribosome assembly rate and concomitantly stronger fetal growth. Thus we consider the polymorphism as a candidate quantitative trait nucleotide (QTN) position. Interestingly, the pCE QTL on BTA21 is also in a chromosome region encoding factors involved in ribosomal assembly, specifically small nucleolar RNAs (snoRNAs). It is therefore possible that both QTL affect ribosomal biogenesis. Mutation disturbing ribosome assembly are often associated with abnormal fetal growth (Freed et al. 2010) (Lempiäinen and Shore 2009).

**ACKNOWLEDGEMENTS**

This study is part of the project FUGATO-plus GenoTrack and was financially supported by the German Ministry of Education and Research, BMBF (grants 0315134A and 0315134D), the Förderverein Biotechnologieforschung e.V. (FBF), Bonn, and Lohmann Tierzucht GmbH, Cuxhaven.

**REFERENCES**


Barendse, W., A. Reverter, R. J. Bunch, B. E. Harrison, W. Barris et al., 2007 A validated whole-genome association study of efficient food conversion in cattle. Genetics 176: 1893-1905.


DAVIS, G. P., D. J. S. HEITZEL, N. J. CORBET, S. SCACHERI, S. LOWDEN ET AL., 1998 The mapping of quantitative trait loci for birth weight in tropical beef herd. 6th World Congress on Genetics


LEWIS, S. E., S. M. J. SEARLE, N. HARRIS, M. GIBSON, V. LYER ET AL., 2002  Apollo: a sequence


**TABLE 1:**

SNPs showing significant association with paternal calving ease (pCE), paternal stillbirth incidence (pSB), daily gain (DG) and body size (BS) EBVs in 1800 Fleckvieh animals

<table>
<thead>
<tr>
<th>SNP</th>
<th>Chromosome</th>
<th>Minor allele (minor allele frequency)</th>
<th>Physical position (bp)</th>
<th>EBV</th>
<th>Eigenstrat statistic</th>
<th>P value</th>
<th>α</th>
</tr>
</thead>
<tbody>
<tr>
<td>ARS-BFGL-NGS-104268</td>
<td>14</td>
<td>A (0.12)</td>
<td>24,057,354</td>
<td>pCE</td>
<td>60.99</td>
<td>$5.72 \times 10^{-15}$</td>
<td>-0.58</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>pSB</td>
<td>47.12</td>
<td>$6.69 \times 10^{-12}$</td>
<td>-0.54</td>
</tr>
<tr>
<td>BTA-91250-no-rs</td>
<td>14</td>
<td>A (0.10)</td>
<td>24,145,838</td>
<td>pCE</td>
<td>59.32</td>
<td>$1.34 \times 10^{-14}$</td>
<td>-0.62</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>pSB</td>
<td>47.51</td>
<td>$5.47 \times 10^{-12}$</td>
<td>0.58</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>BS</td>
<td>26.89</td>
<td>$2.15 \times 10^{-7}$</td>
<td>0.45</td>
</tr>
<tr>
<td>BTB-01417924</td>
<td>14</td>
<td>G (0.13)</td>
<td>24,182,406</td>
<td>pCE</td>
<td>43.54</td>
<td>$4.15 \times 10^{-11}$</td>
<td>-0.46</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>pSB</td>
<td>33.79</td>
<td>$6.13 \times 10^{-9}$</td>
<td>-0.42</td>
</tr>
<tr>
<td>Hapmap59686-rs29020689</td>
<td>14</td>
<td>A (0.14)</td>
<td>24,365,162</td>
<td>pCE</td>
<td>39.07</td>
<td>$4.08 \times 10^{-10}$</td>
<td>-0.40</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>pSB</td>
<td>38.27</td>
<td>$6.18 \times 10^{-10}$</td>
<td>-0.42</td>
</tr>
<tr>
<td>ARS-BFGL-NGS-28867</td>
<td>14</td>
<td>G (0.10)</td>
<td>20,323,857</td>
<td>pCE</td>
<td>38.03</td>
<td>$6.96 \times 10^{-10}$</td>
<td>-0.50</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>pSB</td>
<td>33.78</td>
<td>$6.17 \times 10^{-9}$</td>
<td>-0.49</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>DG</td>
<td>25.50</td>
<td>$4.42 \times 10^{-7}$</td>
<td>0.41</td>
</tr>
<tr>
<td>UA-IFASA-7112</td>
<td>14</td>
<td>G (0.09)</td>
<td>16,109,986</td>
<td>pCE</td>
<td>36.94</td>
<td>$1.22 \times 10^{-9}$</td>
<td>-0.51</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>pSB</td>
<td>32.09</td>
<td>$1.47 \times 10^{-8}$</td>
<td>-0.49</td>
</tr>
<tr>
<td>Hapmap46735-BTA-86653</td>
<td>14</td>
<td>G (0.20)</td>
<td>25,401,722</td>
<td>pCE</td>
<td>34.40</td>
<td>$4.48 \times 10^{-9}$</td>
<td>-0.36</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>pSB</td>
<td>28.89</td>
<td>$7.65 \times 10^{-8}$</td>
<td>-0.34</td>
</tr>
<tr>
<td>ARS-BFGL-NGS-53975</td>
<td>21</td>
<td>G (0.27)</td>
<td>2,151,256</td>
<td>pCE</td>
<td>31.25</td>
<td>$2.27 \times 10^{-8}$</td>
<td>0.24</td>
</tr>
<tr>
<td>BTB-01532239</td>
<td>14</td>
<td>A (0.28)</td>
<td>24,437,778</td>
<td>pCE</td>
<td>28.05</td>
<td>$1.19 \times 10^{-7}$</td>
<td>-0.26</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>pSB</td>
<td>26.94</td>
<td>$2.10 \times 10^{-7}$</td>
<td>-0.27</td>
</tr>
<tr>
<td>ARS-BFGL-NGS-114372</td>
<td>21</td>
<td>C (0.22)</td>
<td>2,381,941</td>
<td>pCE</td>
<td>25.65</td>
<td>$4.09 \times 10^{-7}$</td>
<td>0.24</td>
</tr>
<tr>
<td>Hapmap52072-rs29018920</td>
<td>21</td>
<td>A (0.22)</td>
<td>2,333,804</td>
<td>pCE</td>
<td>24.12</td>
<td>$9.01 \times 10^{-7}$</td>
<td>0.23</td>
</tr>
</tbody>
</table>

Eleven SNPs meet the genome-wide significance level of $P < 1.14 \times 10^{-6}$. SNPs are arranged in the order of increasing $P$ values for the association with the paternal calving ease EBV. The $P$ value for each trait x genotype combination is obtained by a principal components analysis - based approach to account for population stratification. The allelic substitution effect ($\alpha$) is given for the minor
allele in additive genetic standard deviations of the EBV. Physical positions are based on the UMD3.1 assembly of the bovine genome sequence.
<table>
<thead>
<tr>
<th>SNP</th>
<th>Physical position (bp)</th>
<th>Haplotype allele</th>
<th>Minor allele (allele frequency)</th>
<th>Eigenstrat statistic</th>
<th>P value</th>
<th>α</th>
</tr>
</thead>
<tbody>
<tr>
<td>BTB-01953819</td>
<td>23,817,572</td>
<td>A</td>
<td>G (0.26)</td>
<td>0.37</td>
<td>0.54</td>
<td>0.03</td>
</tr>
<tr>
<td>Hapmap45796-BTA-25271</td>
<td>23,853,811</td>
<td>T</td>
<td>A (0.07)</td>
<td>5.39</td>
<td>0.02</td>
<td>-0.18</td>
</tr>
<tr>
<td>ss250608741 *</td>
<td>23,884,989</td>
<td>G</td>
<td>A (0.09)</td>
<td>1.06</td>
<td>0.3</td>
<td>0.06</td>
</tr>
<tr>
<td>ARS-BFGL-BAC-8052</td>
<td>23,893,220</td>
<td>G</td>
<td>A (0.01)</td>
<td>6.72</td>
<td>9.55 x 10^{-3}</td>
<td>-0.45</td>
</tr>
<tr>
<td>ARS-BFGL-NGS-97821</td>
<td>23,946,436</td>
<td>G</td>
<td>A (0.1)</td>
<td>0.98</td>
<td>0.32</td>
<td>0.07</td>
</tr>
<tr>
<td>ARS-BFGL-NGS-104268</td>
<td>24,057,354</td>
<td>A</td>
<td>A (0.12)</td>
<td>61.00</td>
<td>5.71 x 10^{-15}</td>
<td>-0.58</td>
</tr>
<tr>
<td>BTA-91250-no-rs</td>
<td>24,145,838</td>
<td>A</td>
<td>A (0.1)</td>
<td>59.32</td>
<td>1.34 x 10^{-14}</td>
<td>-0.62</td>
</tr>
<tr>
<td>BTB-01417924</td>
<td>24,182,406</td>
<td>G</td>
<td>G (0.13)</td>
<td>43.54</td>
<td>4.15 x 10^{-11}</td>
<td>-0.46</td>
</tr>
<tr>
<td>ARS-BFGL-NGS-110427</td>
<td>24,326,513</td>
<td>A</td>
<td>G (0.11)</td>
<td>0.02</td>
<td>0.89</td>
<td>-0.01</td>
</tr>
<tr>
<td>Hapmap59686-rs29020689</td>
<td>24,365,162</td>
<td>A</td>
<td>A (0.14)</td>
<td>36.94</td>
<td>1.22 x 10^{-9}</td>
<td>-0.40</td>
</tr>
<tr>
<td>ARS-BFGL-NGS-102351</td>
<td>24,401,125</td>
<td>G</td>
<td>G (0.25)</td>
<td>18.34</td>
<td>1.85 x 10^{-4}</td>
<td>-0.21</td>
</tr>
<tr>
<td>BTB-01532239</td>
<td>24,437,778</td>
<td>A</td>
<td>A (0.28)</td>
<td>28.04</td>
<td>1.19 x 10^{-7}</td>
<td>-0.26</td>
</tr>
<tr>
<td>BTB-01530788</td>
<td>24,524,205</td>
<td>A</td>
<td>G (0.34)</td>
<td>8.65</td>
<td>3.27 x 10^{-1}</td>
<td>0.12</td>
</tr>
<tr>
<td>BTB-01530836</td>
<td>24,573,257</td>
<td>G</td>
<td>A (0.35)</td>
<td>4.30</td>
<td>0.04</td>
<td>0.07</td>
</tr>
<tr>
<td>BTB-00557585</td>
<td>24,607,527</td>
<td>A</td>
<td>G (0.35)</td>
<td>4.75</td>
<td>0.04</td>
<td>0.08</td>
</tr>
<tr>
<td>BTB-00557532</td>
<td>24,643,266</td>
<td>A</td>
<td>G (0.35)</td>
<td>4.53</td>
<td>0.03</td>
<td>0.07</td>
</tr>
<tr>
<td>ss250608762 *</td>
<td>24,759,177</td>
<td>G</td>
<td>T (0.01)</td>
<td>1.00</td>
<td>0.32</td>
<td>-0.14</td>
</tr>
<tr>
<td>Hapmap40120-BTA-34288</td>
<td>24,787,245</td>
<td>C</td>
<td>A (0.09)</td>
<td>0.28</td>
<td>0.6</td>
<td>-0.05</td>
</tr>
<tr>
<td>ss250608721 *</td>
<td>24,954,981</td>
<td>A</td>
<td>A (0.16)</td>
<td>58.57</td>
<td>1.96 x 10^{-14}</td>
<td>-0.47</td>
</tr>
<tr>
<td>ss250608720 *</td>
<td>24,955,318</td>
<td>T</td>
<td>C (0.32)</td>
<td>3.56</td>
<td>0.06</td>
<td>0.06</td>
</tr>
<tr>
<td>Hapmap41234-BTA-34285</td>
<td>25,107,556</td>
<td>G</td>
<td>A (0.04)</td>
<td>13.89</td>
<td>1.94 x 10^{-4}</td>
<td>-0.42</td>
</tr>
<tr>
<td>BTB-02056709</td>
<td>25,175,950</td>
<td>A</td>
<td>G (0.18)</td>
<td>2.55</td>
<td>0.11</td>
<td>-0.08</td>
</tr>
<tr>
<td>BTB-00559128</td>
<td>25,215,027</td>
<td>A</td>
<td>G (0.21)</td>
<td>0.01</td>
<td>0.92</td>
<td>0.00</td>
</tr>
<tr>
<td>BTB-00557354</td>
<td>25,254,540</td>
<td>G</td>
<td>A (0.12)</td>
<td>1.63</td>
<td>0.2</td>
<td>0.09</td>
</tr>
<tr>
<td>Hapmap46986-BTA-34282</td>
<td>25,307,116</td>
<td>A</td>
<td>G (0.46)</td>
<td>9.62</td>
<td>1.93 x 10^{-3}</td>
<td>0.13</td>
</tr>
<tr>
<td>BTB-01779799</td>
<td>25,351,733</td>
<td>G</td>
<td>A (0.44)</td>
<td>19.00</td>
<td>1.30 x 10^{-5}</td>
<td>0.19</td>
</tr>
<tr>
<td>Hapmap46735-BTA-86653</td>
<td>25,401,722</td>
<td>G</td>
<td>G (0.2)</td>
<td>34.40</td>
<td>4.48 x 10^{-9}</td>
<td>-0.36</td>
</tr>
</tbody>
</table>
Twenty-three SNPs belong to the BovineSNP50 Bead chip collection and 4 additional SNPs designated by * result from re-sequencing. The P values were obtained by using a principal components analysis-based approach to account for population stratification. Genotypes for SNPs resulting from re-sequencing were determined in 810 animals and imputed for the remaining 990 animals of the study population. The allelic substitution effect ($\alpha$) is given for the minor allele in additive genetic standard deviations of the pCE EBV. SNPs are arranged according to their physical position, based on the UMD3.1 assembly of the bovine genome sequence.
**FIGURE 1**: Association of 43,863 SNPs with the estimated breeding value (EBV) for paternal calving ease (pCE) in the Fleckvieh breed. (A) Manhattan plot. Red triangles represent SNPs with $P < 1.14 \times 10^{-6}$ (Bonferroni corrected significance level). (B) Quantile-quantile plot. The shaded area represents the 95% concentration band under the null hypothesis of no association. The open black dots represent the $P$ values of the entire study, open triangles represent SNPs with $P < 1 \times 10^{-8}$. The filled blue dots indicate the $P$ values excluding those from the associated regions on chromosome 14 and 21.
FIGURE 2: Effect of the most significantly associated markers on the estimated breeding value (EBV) for paternal calving ease (pCE) in the Fleckvieh breed. The boxplots show the effects of the most significantly associated SNPs (A) and haplotypes (B) on chromosome 14 and 21 separately and combined. The solid line represents the population mean, the dotted lines indicate one standard deviation of the EBV.
FIGURE 3: Detailed view of the region on chromosome 14 delineated by the haplotype associated with the estimated breeding value (EBV) for paternal calving ease (pCE). (A) Map of genes contained in this region. Red symbols indicate genes re-sequenced in the present study. (B) P values of 27 SNPs from analysis of association with the pCE EBV. The open black dots indicate results from genotyping of the entire study population, the blue triangles represent P values resulting from imputation based on 810 genotyped animals. (C) Heatmap of the pairwise linkage disequilibrium ($r^2$). The triangle delineates a linkage disequilibrium block containing the most significantly associated SNPs, including the potentially functional $ss250608721$ variant in $RPS20$. 
FIGURE 4: Predicted 3' UTR of cattle RPS20. The grey-shaded sequence designates the predicted exon 4, while the predicted polyadenylation (poly(A)) sites are denoted by underscore. The star locates the candidate quantitative trait nucleotide position, ablating a poly(A) site.