Polymorphisms within the APOBR gene are highly associated with milk levels of prognostic ketosis biomarkers in dairy cows


1Institute of Animal Breeding and Husbandry, Christian Albrecht University, Kiel, Germany; and 2Institute of Functional Genomics, University of Regensburg, Regensburg, Germany

Submitted 2 December 2014; accepted in final form 4 February 2015

One of the most common metabolic disorders in dairy cattle is ketosis or acetonemia, which is characterized by an increased production of the ketone bodies beta-hydroxybutyrate, acetone, and acetoacetate (1). These arise from incomplete hepatic oxidation of nonesterified fatty acids (NEFAs) mobilized from depot fat, as the acetyl-CoA resulting from β-oxidation of NEFAs cannot completely enter the tricarboxylic acid cycle, because oxaloacetate is redirected to gluconeogenesis (2). Synthesis from noncarbohydrate precursors is the main pathway to ensure glucose supply for lactose synthesis in the mammary gland (28), because enteral glucose availability is low due to ruminal fermentation of carbohydrates (3). Milk levels of ketone bodies are established biomarkers for subclinical ketosis (12, 13), which was estimated to occur at a prevalence of 21.8% across European countries (30). Recently, NMR metabolomic analysis revealed that the milk glycerophosphocholine (GPC)-to-phosphocholine (PC) ratio in early lactation and the GPC level in midlactation were prognostic biomarkers for the risk of ketosis in dairy cattle (20). In that study, healthy animals had significantly higher levels of milk GPC and lower levels of milk PC than animals suffering from ketosis. The authors hypothesized that the raised GPC/PC ratios might be due to higher rates of blood phosphatidylcholine from low-density lipoproteins play an important role in early lactation metabolic stability of dairy cows and indicate APOBR to contain a causative variant.

Keywords: ketosis; metabolomics; GWAS; glycerophosphocholine; phosphocholine; APOBR

OVER THE YEARS, MILK PRODUCTION per cow has increased considerably. But this increase has been accompanied by a greater prevalence of health problems (26). Predominant diseases in highly selected dairy cows (e.g., mastitis, metritis, and metabolic disorders) are most prevalent at the beginning of lactation and have been related to the extent and duration of the postpartum energy deficit (9, 17, 22). Following parturition, the energy requirements for milk production increase tremendously and, initially, are not met by daily feed consumption (5). Essentially all dairy cows experience a period of negative energy balance, which is compensated for by increased body fat mobilization (4, 22). Excessive mobilization due to a severe energy deficit impairs the cow’s immune function (18) and fertility (34) and leads to metabolic stress (9). As a consequence, cows experiencing a prolonged period of negative energy balance are more prone to metabolic problems (9, 14).

One of the most common metabolic disorders in dairy cattle is ketosis or acetonemia, which is characterized by an increased production of the ketone bodies beta-hydroxybutyrate, acetone, and acetoacetate (1). These arise from incomplete hepatic oxidation of nonesterified fatty acids (NEFAs) mobilized from depot fat, as the acetyl-CoA resulting from β-oxidation of NEFAs cannot completely enter the tricarboxylic acid cycle, because oxaloacetate is redirected to gluconeogenesis (2). Synthesis from noncarbohydrate precursors is the main pathway to ensure glucose supply for lactose synthesis in the mammary gland (28), because enteral glucose availability is low due to ruminal fermentation of carbohydrates (3). Milk levels of ketone bodies are established biomarkers for subclinical ketosis (12, 13), which was estimated to occur at a prevalence of 21.8% across European countries (30). Recently, NMR metabolomic analysis revealed that the milk glycerophosphocholine (GPC)-to-phosphocholine (PC) ratio in early lactation and the GPC level in midlactation were prognostic biomarkers for the risk of ketosis in dairy cattle (20). In that study, healthy animals had significantly higher levels of milk GPC and lower levels of milk PC than animals suffering from ketosis. The authors hypothesized that the raised GPC/PC ratios might be due to higher rates of blood phosphatidylcholine from low-density lipoproteins play an important role in early lactation metabolic stability of dairy cows and indicate APOBR to contain a causative variant.

Keywords: ketosis; metabolomics; GWAS; glycerophosphocholine; phosphocholine; APOBR

OVER THE YEARS, MILK PRODUCTION per cow has increased considerably. But this increase has been accompanied by a greater prevalence of health problems (26). Predominant diseases in highly selected dairy cows (e.g., mastitis, metritis, and metabolic disorders) are most prevalent at the beginning of lactation and have been related to the extent and duration of the postpartum energy deficit (9, 17, 22). Following parturition, the energy requirements for milk production increase tremendously and, initially, are not met by daily feed consumption (5). Essentially all dairy cows experience a period of negative energy balance, which is compensated for by increased body fat mobilization (4, 22). Excessive mobilization due to a severe energy deficit impairs the cow’s immune function (18) and fertility (34) and leads to metabolic stress (9). As a consequence, cows experiencing a prolonged period of negative energy balance are more prone to metabolic problems (9, 14).

One of the most common metabolic disorders in dairy cattle is ketosis or acetonemia, which is characterized by an increased production of the ketone bodies beta-hydroxybutyrate, acetone, and acetoacetate (1). These arise from incomplete hepatic oxidation of nonesterified fatty acids (NEFAs) mobilized from depot fat, as the acetyl-CoA resulting from β-oxidation of NEFAs cannot completely enter the tricarboxylic acid cycle, because oxaloacetate is redirected to gluconeogenesis (2). Synthesis from noncarbohydrate precursors is the main pathway to ensure glucose supply for lactose synthesis in the mammary gland (28), because enteral glucose availability is low due to ruminal fermentation of carbohydrates (3). Milk levels of ketone bodies are established biomarkers for subclinical ketosis (12, 13), which was estimated to occur at a prevalence of 21.8% across European countries (30). Recently, NMR metabolomic analysis revealed that the milk glycerophosphocholine (GPC)-to-phosphocholine (PC) ratio in early lactation and the GPC level in midlactation were prognostic biomarkers for the risk of ketosis in dairy cattle (20). In that study, healthy animals had significantly higher levels of milk GPC and lower levels of milk PC than animals suffering from ketosis. The authors hypothesized that the raised GPC/PC ratios might be due to higher rates of blood phosphatidylcholine from low-density lipoproteins play an important role in early lactation metabolic stability of dairy cows and indicate APOBR to contain a causative variant.
quantitative trait loci (QTL) affecting energy metabolism in early lactation. A genome wide-significant QTL on cattle chromosome 25 (BTA25) identified the APOBR gene, which encodes the apolipoprotein B receptor, as a plausible candidate. Further analyses indicated that genetic variation within APOBR was likely to underlie the QTL.

MATERIAL AND METHODS

Samples, metabolite analysis, and genotyping. Milk samples were collected repeatedly from Holstein-Friesian cows at the Karkendamm dairy research farm in Northern Germany. During the first 5 wk of lactation, specimens were obtained weekly. Thereafter, specimens were collected in midlactation (month 6) and at the end of lactation. Some cows left the herd early, and, thus, time series were not complete for all animals. Sample preparation and NMR measurements were performed at the Institute of Functional Genomics, University of Regensburg, Germany, according to established protocols (15, 19). A detailed description of the experimental settings is given by Klein et al. (2012) (20). The original dataset contained information on cows in lactation one to seven, but cows with lactation numbers ≥ 2 were not genotyped. Therefore, the phenotypic dataset was reduced to 248 first-lactation cows. All samples were obtained between October 2008 and June 2010. Table 1 provides a more detailed overview of the phenotypic data used in this study.

Cows were genotyped using the Illumina BovineSNP50 BeadChip (Illumina, San Diego, CA) comprising 54,001 single nucleotide polymorphisms (SNPs). Based on GC-score filtering of the genotyping results, with a threshold of 0.15 applied as recommended by the manufacturer, all animals had <10% missing genotypes. SNPs without a chromosomal assignment with respect to genome build UMD3.1 were excluded from the investigation. To increase the power of the GWAS, we used minimac (16) to impute sporadically missing SNP genotypes (~1% of genotypes) after prephasing the data with MaCH (23). Thereafter, SNPs with a minor allele frequency ≤ 0.05 were removed, resulting in a final set of 41,409 SNPs.

Statistical analyses. We carried out the GWAS by fitting the following linear mixed model to the data for one SNP at a time: $y = X_b + W_v + Z_{da} + Z_{dp} + e$, where $y$ is a vector of all phenotypic observations across individuals for GPC, PC, or GPC/PC, respectively. $W$ is the design matrix for the fixed allele substitution effects ($v$) allocating the SNP genotypes (coded as 0 for the heterozygous and −1 or 1 for the alternative homozygote genotypes, respectively) to the individuals. The marker effect was omitted from the model for the estimation of the variance components of each trait. $X$ is the design matrix for the fixed effects. The respective vector $b$ includes the intercept and systematic effects on the milk metabolites, e.g., the fixed effects of lactation number and test day (119 levels for GPC, 97 levels for PC and GPC/PC, respectively) and, to account for differences between lactation stages, a linear regression on lactation day. Design matrices $Z_v$ for the additive genetic effects ($u_v$) and $Z_b$ for the permanent environmental effects ($u_p$) relate the elements of $u_v$ and $u_p$ to their corresponding element in $y$. The additive genetic effects $u_v$ are assumed to follow a normal distribution with mean zero and variance $\sigma^2_v$, where $A$ represents a numerator relationship matrix obtained from pedigree information consisting of 4,577 animals in total including 184 unique sires of phenotyped cows. The permanent environmental effect $u_p \sim N(0, \sigma^2_p)$ was included because the dataset contained repeated observations within cows. $I$ is an identity matrix with an order equal to the number of cows with records (248 for GPC, 247 for PC and GPC/PC, respectively). The residuals were assumed to be independently and normally distributed.

Significance of the allele substitution effect was tested by a Wald $\chi^2$-test with $H_0: v = 0$ and $H_a: v \neq 0$. Genome- and chromosome-wise significance thresholds at the 5% level were derived by a Bonferroni correction. Linear mixed models for GWAS and pedigree processing were done using the pedigreemm (32) and lme4 package (4a) within the R statistical environment (27). Gibbs sampling was used to obtain estimates and standard deviations of the variance components and ratios from their marginal posterior distributions using 20,000 iterations and a burn-in period of 10,000. Additionally, bivariate models were fitted to estimate genetic correlations between the traits.

Candidate gene analyses. Within 1 Mb to each side of genome-wide significantly associated markers, the current version of the cattle genome (UMD3.1, annotation release 103) was inspected for plausible candidate genes related to energy and especially phospholipid metabolism. The APOBR gene coding for the apolipoprotein B receptor and located on BTA25 within the most significantly associated QTL region for GPC and GPC/PC was selected for further analysis. We determined gene structure by aligning the

Table 2. PCR and sequencing primers used in comparative resequencing of the bovine APOBR gene

<table>
<thead>
<tr>
<th>Primer</th>
<th>Location</th>
<th>Forward (5’→3’)</th>
<th>Reverse (5’→3’)</th>
<th>Product Size, bp</th>
<th>Annealing Temp., °C</th>
</tr>
</thead>
<tbody>
<tr>
<td>APOBR_Exc1</td>
<td>exon 1</td>
<td>GGGCGAGGTGTAATACACAGACC</td>
<td>AGGAAAGGGCTTAGATGAGGTG</td>
<td>584</td>
<td>62</td>
</tr>
<tr>
<td>APOBR_Exc2A</td>
<td>exon 2</td>
<td>CTGGGATGCGCTGGCTGCTTC</td>
<td>TTCCTCTCGCTGGCTTC</td>
<td>702</td>
<td>62</td>
</tr>
<tr>
<td>APOBR_Exc2B</td>
<td>exon 2</td>
<td>GAGGCTGTCGATAGAGGAGGAG</td>
<td>GGGAGGCTTAGCTCGCTTC</td>
<td>702</td>
<td>62</td>
</tr>
<tr>
<td>APOBR_Exc2C</td>
<td>exon 2</td>
<td>GGAGATGAGAGGAGGACAGC</td>
<td>CTGGCTCTGCTGGCTTC</td>
<td>897</td>
<td>62</td>
</tr>
<tr>
<td>APOBR_Exc2D</td>
<td>exon 2</td>
<td>AGACAGACAGAGAGACGCTGGA</td>
<td>GCCCTCCTCCTCCTGCTCTG</td>
<td>751</td>
<td>62</td>
</tr>
<tr>
<td>APOBR_Exc2E</td>
<td>exon 2</td>
<td>2/3</td>
<td>AAGGAGGAGAGGAGGACCTGGA</td>
<td>CTAATCCCTGCTGCTGAC</td>
<td>738</td>
</tr>
<tr>
<td>APOBR_Exc3</td>
<td>exon 3</td>
<td>TGCAGGAGAGGAGGAGAGGAC</td>
<td>CCAGGAAGAGAGGAGGAGGAGGAC</td>
<td>469</td>
<td>62</td>
</tr>
<tr>
<td>APOBR_Exc4</td>
<td>exon 3</td>
<td>3/4</td>
<td>TGGCAATGGCAGGAGAGGAC</td>
<td>CAGACAGACAGAGAGGAGGAGGAC</td>
<td>575</td>
</tr>
</tbody>
</table>

Locations are given with respect to transcript variant 1 (ENSBTAAT00000037341).
Table 3. Estimates and SD of variance components and genetic parameters for the milk levels of GPC, PC, and GPC/PC, respectively

<table>
<thead>
<tr>
<th>Trait</th>
<th>Model Without Any Marker Effect</th>
<th>Model Including Top SNP</th>
<th>Variance Explained by QTL</th>
<th>QTL</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(\sigma_a^2 \pm SD)</td>
<td>(\sigma_{pe}^2 \pm SD)</td>
<td>(\sigma_e^2 \pm SD)</td>
<td>(h^2 \pm SD)</td>
<td>(r \pm SD)</td>
<td>(\sigma_a^2 \pm SD)</td>
<td>(\sigma_{pe}^2 \pm SD)</td>
<td>(\sigma_e^2 \pm SD)</td>
<td>(\sigma_{QTL}^2/\sigma_e^2)</td>
</tr>
<tr>
<td>GPC</td>
<td>0.018 ± 0.004</td>
<td>0.004 ± 0.003</td>
<td>0.020 ± 0.001</td>
<td>0.43 ± 0.08</td>
<td>0.52 ± 0.03</td>
<td>0.012 ± 0.003</td>
<td>0.003 ± 0.002</td>
<td>0.020 ± 0.001</td>
<td>0.33</td>
</tr>
<tr>
<td>PC</td>
<td>0.001 ± 0.001</td>
<td>0.006 ± 0.001</td>
<td>0.015 ± 0.001</td>
<td>0.07 ± 0.04</td>
<td>0.33 ± 0.04</td>
<td>0.164 ± 0.048</td>
<td>0.095 ± 0.046</td>
<td>0.300 ± 0.017</td>
<td>0.24</td>
</tr>
<tr>
<td>GPC/PC</td>
<td>0.215 ± 0.071</td>
<td>0.121 ± 0.001</td>
<td>0.301 ± 0.017</td>
<td>0.34 ± 0.10</td>
<td>0.53 ± 0.04</td>
<td>0.046</td>
<td>0.300</td>
<td>0.24</td>
<td>0.12</td>
</tr>
</tbody>
</table>

Given are the additive genetic (\(\sigma_a^2\)), permanent environmental (\(\sigma_{pe}^2\)) and residual (\(\sigma_e^2\)) variance components as well as heritabilities (\(h^2\)) and repeatabilities (\(r\)) for each trait. For GPC and GPC/PC, variance components are also given as obtained from a model including the most significant marker within \(APOBR\). The amount of additive genetic and phenotypic (\(\sigma_e^2\)) variance explained by the QTL was calculated from the reduction of \(\sigma_e^2\) caused by the consideration of the effect of the top marker in the model. Variance components were rounded to 3 and ratios to 2 decimal places, respectively. \(^1\)Heritability, \(\sigma_a^2/(\sigma_a^2 + \sigma_{pe}^2 + \sigma_e^2)\). \(^2\)Repeatability, \((\sigma_a^2 + \sigma_{pe}^2)/\sigma_e^2\). \(^3\)\(\sigma_{QTL}^2\) was calculated as the difference between the additive genetic variance components and phenotypic variance, respectively, from the 2 models.

Available transcripts of two different splice variants (Ensembl transcript IDs ENSBTAT00000037341 and ENSBTAT00000065319; GenBank Acc. no. XM_002698084, see Fig. 2) to the current version of the cattle genome sequence (UMD3.1, GenBank Acc. no. AC_000182.1). All exons and adjacent intronic regions were then resequenced in 237 animals from the dataset used for the GWAS as described above. For sequence analysis, PCR primers were designed with primer3 (21, 31) (Table 2). The amplification was performed in a 12 µl reaction volume containing 20 ng of genomic DNA, 0.2 µM of each primer, 200 µM dNTPs, and 5 U of Taq polymerase (Invitrogen, Germany) in the reaction buffer supplied by the manufacturer. Cycling conditions were as follows: initial denaturation at 94°C for 5 min, 35 cycles each of 94°C for 30 s, 62°C for 60 s, 72°C for 90 s, and a final extension step at 72°C for 10 min. After purification of the PCR products with thermosensitive alkaline phosphatase (FastAP; Fermentas, St. Leon-Rot, Germany) and exonuclease I (Fermentas), sequencing was carried out with one of the correspondent primers using an ABI 3130xl Genetic Analyzer and BigDye Terminator v3.1 Cycle Sequencing Ready Reaction Kit (Applied Biosystems, Foster City, CA). Sequence analysis was done using the software Sequencher 4.9 (Gene Codes, Ann Arbor, MI); genotypes were directly called from the chromatograms.

The polymorphisms within the \(APOBR\) gene were tested for trait association with the GWAS model described above. Furthermore, the most significantly associated SNP for GPC and GPC/PC, respectively, were included in the model used to estimate the variance components.

Fig. 1. Results of the genome-wide association study for glycerophosphocholine (GPC), phosphocholine (PC), and GPC/PC. Left: Manhattan plots of the negative logarithm of the raw \(P\) values. The black line indicates the genome-wide significance threshold at a Bonferroni-corrected 5% level; the gray lines indicate the chromosome-wide significance threshold. Right: Q-Q plots comparing the observed and expected distribution of the \(P\) values along with the genomic inflation factor \(\lambda\) indicating a sufficient correction for population structure. Genome- and chromosome-wide significantly associated single nucleotide polymorphisms (SNPs) are highlighted in the Q-Q plot in red and blue, respectively.
The amount of variance explained by the QTL was determined from the reduction of additive genetic variance and heritability, respectively (Table 3).

RESULTS AND DISCUSSION

Genetic parameters. The estimated variance components and ratios for all traits analyzed are summarized in Table 3. The highest heritability of 0.43 was estimated for GPC, but its heritability was lower (0.34). The heritability of the PC level was close to zero ($h^2 = 0.07$). In summary, both the milk level of GPC and the GPC/PC ratio are highly heritable traits, while there seems to be no significant additive genetic effect for the level of PC. These findings indicate that genetic variation in GPC/PC is mainly explained by variation in GPC levels. However, a genetic correlation of $-0.59 \pm 0.10$ (phenotypic correlation $= -0.12$) between GPC and PC was estimated, which might point to a joint genetic control of both traits. This should though be interpreted with care, because such analysis requires the simultaneous estimation of several variance components in a small dataset.

GWAS. The hypothesis that genetic variation of the GPC/PC ratio is due to GPC variation is supported by the results of the GWAS, which revealed three genome-wide significant ($P_{\text{adj}} \leq 0.05$) association signals on BTA7, 23, and 25 for GPC, while for PC only a single genome-wide significantly

Table 4. Results of the genome-wide association study

<table>
<thead>
<tr>
<th>BTAa</th>
<th>Positionb</th>
<th>SNP IDc</th>
<th>Allelesd</th>
<th>MAFe</th>
<th>$P$ Valuef</th>
<th>$\beta \pm SEg$</th>
<th>$P$ Valuef</th>
<th>$\beta \pm SEg$</th>
<th>$P$ Valuef</th>
<th>$\beta \pm SEg$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>41,732,937</td>
<td>ARS-BFGL-NGS-118356</td>
<td>A/G</td>
<td>0.42</td>
<td>1.85E-05</td>
<td>-0.04 \pm 0.01</td>
<td>1.89E-05</td>
<td>-0.09 \pm 0.02</td>
<td>6.60E-07</td>
<td>0.44 \pm 0.09</td>
</tr>
<tr>
<td>3</td>
<td>23,373,990</td>
<td>BTA-69698-no-rs</td>
<td>A/G</td>
<td>0.17</td>
<td>6.85E-07</td>
<td>-0.07 \pm 0.01</td>
<td>6.88E-06</td>
<td>-0.07 \pm 0.02</td>
<td>1.21E-05</td>
<td>-0.30 \pm 0.07</td>
</tr>
<tr>
<td>7</td>
<td>80,269,017</td>
<td>ARS-BFGL-NGS-116915</td>
<td>A/G</td>
<td>0.37</td>
<td>1.86E-05</td>
<td>-0.07 \pm 0.02</td>
<td>1.86E-05</td>
<td>-0.07 \pm 0.02</td>
<td>1.96E-07</td>
<td>-0.36 \pm 0.07</td>
</tr>
<tr>
<td>6</td>
<td>90,239,425</td>
<td>Hapmap58847-BTA-115377</td>
<td>A/G</td>
<td>0.47</td>
<td>4.90E-05</td>
<td>-0.27 \pm 0.07</td>
<td>4.90E-05</td>
<td>-0.48 \pm 0.11</td>
<td>2.91E-06</td>
<td>-0.30 \pm 0.06</td>
</tr>
<tr>
<td>3</td>
<td>23,737,990</td>
<td>BTA-66968-no-rs</td>
<td>A/G</td>
<td>0.17</td>
<td>1.72E-05</td>
<td>-0.19 \pm 0.06</td>
<td>1.72E-05</td>
<td>-0.20 \pm 0.06</td>
<td>2.89E-06</td>
<td>-0.20 \pm 0.06</td>
</tr>
</tbody>
</table>

Single nucleotide polymorphism (SNP) markers from the 50k SNP chip significantly associated with milk levels of GPC and PC as well as GPC/PC, respectively. aCattle chromosome. bGenomic position according to the bovine genome build UMD3.1. cIllumina SNP identifier. dThe minor allele is given first. eMinor allele frequency. fGiven is the raw $P$ value, boldfaced letters indicate genome-wide significance at the Bonferroni-corrected 5% level; all other listed results represent chromosome-wide significant associations. gGiven is the additive effect of the minor allele together with its SE.
Table 5. Polymorphisms identified within the APOBR gene and their association with milk levels of GPC and the ratio of GPC and PC, respectively

<table>
<thead>
<tr>
<th>Location*</th>
<th>CDS*</th>
<th>AA*</th>
<th>Minor Allele</th>
<th>MAF*</th>
<th>Genomic Position†</th>
<th>rs/ss Number</th>
<th>P Value‡</th>
<th>β ± SE*</th>
<th>P Value‡</th>
<th>β ± SE*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exon 1 (5'-UTR)</td>
<td>c.1-196G&gt;A</td>
<td>A</td>
<td>0.063</td>
<td>26,314,750</td>
<td>rs133727743</td>
<td>1.65E-05</td>
<td>0.13 ± 0.03</td>
<td>1.07E-05</td>
<td>0.54 ± 0.12</td>
<td></td>
</tr>
<tr>
<td>Intron 1</td>
<td>c.57 + 109&gt;C</td>
<td>T</td>
<td>0.057</td>
<td>26,314,795</td>
<td>rs208487368</td>
<td>9.36E-05</td>
<td>0.12 ± 0.03</td>
<td>4.31E-04</td>
<td>0.44 ± 0.13</td>
<td></td>
</tr>
<tr>
<td>Intron 2</td>
<td>c.57 + 157A&gt;G</td>
<td>G</td>
<td>0.233</td>
<td>26,315,159</td>
<td>rs2071228</td>
<td>3.28E-15</td>
<td>0.13 ± 0.02</td>
<td>3.95E-09</td>
<td>0.41 ± 0.07</td>
<td></td>
</tr>
<tr>
<td>Exon 2</td>
<td>c.58-20T&gt;C</td>
<td>C</td>
<td>0.233</td>
<td>26,315,323</td>
<td>rs2071227</td>
<td>1.62E-13</td>
<td>0.12 ± 0.02</td>
<td>4.39E-09</td>
<td>0.40 ± 0.07</td>
<td></td>
</tr>
</tbody>
</table>

The association signal on BTA25 was the most significant for GPC and equivalently the most prominent for the ratio trait (Fig. 1, Table 4). The association signal for PC on BTA3 was likewise present for the GPC/PC ratio, but only represented by a single SNP. Furthermore, there was a notable genome-wide significant association for GPC/PC on BTA9 that was neither present for GPC nor for PC (Fig. 1, Table 4).

A total of 10 SNPs reached genome-wide significance; four out of these SNPs were located on BTA25, two of which were concomitantly associated with GPC levels and the GPC/PC ratio (Table 4); these markers are in linkage disequilibrium with D' = 0.65 and r² = 0.38. The lowest nominal P values for GPC (P = 4.54E-09) and GPC/PC (P = 6.14E-09) were obtained for the SNP marker ARS-BFGL-NGS-100347 located at 26,982,725 bp on BTA25. The association signals for GPC and GPC/PC on BTA25 were essentially the same (Figs. 1 and 2), while there was no signal for PC. The direction of the allele substitution effects estimated for the most significantly associated markers on BTA25 was in line with the hypothesis, that

*Given is the location with respect to transcript ENSBTAT0000037341; exon 2 of this transcript is represented by 3 adjacent exons in the alternative transcript ENSBTAT00000065319; see Fig. 2 for details. †CDS = coding sequence, AA = amino acid sequence; nucleotide positions are with respect to transcript ENSBTAT00000065319, positions within the alternative transcript ENSBTAT00000065319 are given in parentheses; as both transcript have the same reading frame, amino acid changes are denoted equivalently. ‡Minor allele frequency. §Genomic position according to the bovine genome build UMD3.1. ★Given is the raw P value, boldfaced letters indicate genome-wide significance at the Bonferroni-corrected 5% level. ★Given is the additive effect of the minor allele together with its SE.

APOBR GENE IS HIGHLY ASSOCIATED WITH KETOSIS BIOMARKERS

Physiol Genomics • doi:10.1152/physiolgenomics.00126.2014 • www.physiolgenomics.org
variation in GPC levels determines variation in GPC/PC (Table 4). A QTL for milk levels of GPC has previously been reported on BTA25 (8), but that study did not analyze PC.

Identification of candidate genes. Klein et al. (2012) (20) hypothesized that the individual variability of the milk GPC/PC ratio might reflect different abilities to break down blood PtC as a fatty acid source for milk fat synthesis, thus reducing the amount of body fat mobilization. This hypothesis might be extended in so far as the increased PtC utilization reduces the need for de novo fatty acid synthesis in the mammary gland and, thus, the recruitment of NADPH from the pentose phosphate cycle. This in turn might lead to an increased C3 return from the mammary pentose phosphate cycle in favor of hepatic gluconeogenesis. It has been reported that this mechanism might contribute to the allocation of gluconic carbon in dairy cattle (2).

Based on the above hypothesis, we screened the close vicinity of genome-wide significantly associated SNPs for possible candidate genes involved especially in phospholipid metabolism, but also in energy metabolism in general. There are no genes annotated within these intervals coding for phospholipases directly involved in PtC breakdown. An alternative mechanism involved in blood PtC breakdown as a fatty acid source in the mammary gland might be related to cellular low-density lipoprotein (LDL) uptake. Approximately 600 kb upstream of the most significant association signal on BTA25, the gene encoding for the apolipoprotein B-receptor (APOBR) is located. This receptor has been described as a macrophage receptor (6, 7), but it interacts with apolipoprotein B, a main constituent of LDL. Furthermore, the uptake and conversion of LDL-associated PtC have been shown at least for hepatocytes (24). The inspection of the genomic regions adjacent to associated SNPs on other chromosomes revealed no equally obvious candidates. However, on BTA23, ~800 kb downstream of the SNP Hapmap41168-BTA-117434 (Table 4), the bovine HMGCLL1 gene is located, which encodes the 3-hydroxy-3-methyl-3-methylglutaryl-CoA (HMG-CoA) lyase-like protein 1. HMG-CoA lyase catalyzes the cleavage of HMG-CoA into acetoacetate and acetyl-CoA (29), an essential step in ketogenesis. HMGCLL1 codes for an extramitochondrial form of this enzyme (25). This is an obvious link to ketosis, although a possible connection with PtC or GPC metabolism remains to be elucidated. Furthermore, the ITPR3 gene, coding for the inositol 1,4,5-triphosphate receptor type 3, is located ~100 kb upstream of the genome-wide significantly associated SNP ARS-BFGL-NGS-115177 on BTA23 (Table 4). This gene is also not directly linked to phospholipid metabolism but is strongly linked to ketosis, as it is known to regulate hepatic gluconeogenesis in fasting and diabetes (31).

Analysis of the APOBR gene. A total of 49 polymorphisms were identified within the APOBR gene (Table 5), which are almost exclusively SNPs. The only exception was a 3 bp insertion within the open reading frame leading to the insertion of an aspartic acid residue after position 163 of the protein (Table 5). With respect to transcript ENSBTAT0000037341

Fig. 2. Detailed view of the association results for GPC and GPC/PC within the QTL region on BTA25 and within the APOBR gene. Depicted are the negative decadic logarithms of the P values for both traits within the QTL region illustrating the matching association signals. Markers from the 50k SNP chip are depicted as dots, intragenic polymorphisms as triangles (A). Intragenic association results are shown with respect to 2 different splice variants (transcripts ENSBTAT0000037341 and ENSBTAT0000065319, see Table 4). Yellow asterisks indicate nonsynonymous mutations. The black lines in both A and B indicate the genome-wide significance threshold at a Bonferroni-corrected 5% level; the gray line indicates the chromosome-wide significance threshold.
(Table 5, Fig. 2), 36 of the 49 polymorphisms were located within the open reading frame, and 22 of these represent nonsynonymous mutations. A total of 29 polymorphisms were found to be genome-wide significantly associated at the 5%-Bonferroni level with both GPC and the GPC/PC ratio. Given the small dataset, the obtained \( P \) values were exceedingly small. For GPC, a minimum of 2.16E-18 was found for a synonymous G>C transversion in exon 3 at position 2955 of the coding sequence, while the best results for GPC/PC (\( P = 1.04E-12 \)) were obtained for two nonsynonymous A>G transitions causing the amino acid substitutions Glu518Lys and Val681Met in exon 2, respectively (Table 5, Fig. 2). The top SNPs explain 33 and 23% of the additive genetic variance for GPC and the GPC/PC ratio, corresponding to 17 and 12% of explained phenotypic variance, respectively.

Overall, the genome-wide significantly associated polymorphisms were scattered across the entire gene. They all showed a similar minor allele frequency of \( \approx 0.23 \) and almost identical absolute effect sizes of \( \approx 0.13 \) for GPC and \( \approx 0.4 \) to 0.5 for GPC/PC, respectively (Table 5) and probably represent the haplotype containing the causative mutation. Consequently, the pairwise linkage disequilibrium (LD) between the significantly associated markers averaged to an \( r^2 \) value of 0.96. Thus, it is not possible to determine an actual causative variant by a purely genetic approach. Also, tools to predict possibly deleterious variants are of limited informativeness for prioritization, as the variant underlying the analyzed QTL is segregating at high frequencies and might only be of functional impact in high-yielding animals, thus possibly being neutral under natural selection conditions. Furthermore, it cannot be excluded that the actual causative variant is a nearby regulatory mutation not captured by the current experiment.

In their previous work, Klein et al. (20) demonstrated the predictive ability of the analyzed metabolites with respect to ketosis at a phenotypic level. They determined thresholds of 2.5 for the GPC/PC ratio in lactation weeks 1–4 and 1.2 mmol/l for GPC in lactation month six, respectively. Cows exceeding these limits were found to be less prone to ketosis. Compared with the mean phenotypic values of different QTL genotypes (Fig. 3), the thresholds are rather high. Although predominantly animals homozygous for the favorable QTL allele fall above the limits, the QTL genotype cannot be used to classify the animals with respect to ketosis liability. This result has to be expected, because the QTL only explains a part of the additive genetic variance, which in turn only explains half of the phenotypic variance of metabolite concentrations. The metabolites are, however, highly informative, reflecting both the genetic determination and the actual environment. This has also been found in a study on the prediction of individual ketosis risk by means of artificial neural networks using the very same dataset (11). The determination of the causative variant for the current and further identified QTL will enhance our knowledge about the mechanisms underlying metabolic stability in early lactation and help to determine combinations of metabolites, genetic markers, and possibly other phenotypes that are more predictive with respect to individual disease risk.

![Fig. 3. Milk GPC levels (A) and ratios of GPC to PC (B) by genotype across different lactation stages.](http://physiolgenomics.physiology.org/)

**Fig. 3.** Milk GPC levels (A) and ratios of GPC to PC (B) by genotype across different lactation stages. The phenotypic values are expressed with respect to the genotype of the polymorphism within the **APOBR** gene most significantly associated with GPC. For lactation month 6, only GPC values were available. The dashed lines indicate the thresholds determined by Klein et al. (19).
Conclusions

It has previously been reported that the ratio of milk GPC and PC is a prognostic biomarker for the risk of ketosis in dairy cattle. It was hypothesized that this might be due to a different ability of blood PtC breakdown as a source for mammary fatty acid synthesis. Here, using the same set of animals, we found that the milk level of GPC and the GPC/PC ratio are highly heritable traits. A major QTL on BTA25 significantly affects milk GPC and GPC/PC, but not the level of PC. Within the QTL region, the APOBR gene coding for the apolipoprotein B receptor is located. The analysis of this gene revealed highly significant associations with GPC and the GPC/PC ratio, suggesting APOBR harbors the causative variant underlying this QTL. These findings support the hypothesis that differences in the ability to take up blood PtC from LDL play a role in the metabolic stability of dairy cows during early lactation. How the APOBR gene coding for the apolipoprotein B receptor is located. The analysis of this gene revealed highly significant associations with GPC and the GPC/PC ratio, suggesting APOBR harbors the causative variant underlying this QTL. These findings support the hypothesis that differences in the ability to take up blood PtC from LDL play a role in the metabolic stability of dairy cows during early lactation. However, the identification of the causative mutation was not possible by purely genetic approaches; further studies will be necessary.

ACKNOWLEDGMENTS

The authors thank the staff at the Karkendamm research farm for providing milk and plasma specimens and NORD-OST GENETIC GmbH & Co. KG (Verden/Aller, Germany) for kind permission to include their animals in this study. Gabi Otzen is acknowledged for expert technical assistance. Current address for M. S. Klein: Faculty of Kinesiology, Univ. of Calgary, Calgary, Alberta, Canada.

GRANTS

This study was funded by the German Federal Ministry of Education and Research in context of the Fugato-plus MeGA-M framework and by the Bavarian Genome Network (BayGene), Munich, Germany. The genotyping of the cows was financially supported by the German Federal Ministry of Education and Research (project FUGATO-plus GENOTRACK, grant no. 0315134A) and the KMSH (Kompetenzzentrum Milch - Schleswig-Holstein, Kiel, Germany).

DISCLOSURES

No conflicts of interest, financial or otherwise, are declared by the author(s).

AUTHOR CONTRIBUTIONS

Author contributions: J.T., G.T., and N.K. conception and design of research; J.T., I.H., M.S.K., W.G., and P.J.O. performed experiments; J.T. and C.H. analyzed data; J.T. and N.K. interpreted results of experiments; J.T. prepared figures; J.T., M.S.K., W.G., W.J., P.J.O., G.T., and N.K. drafted manuscript; J.T., M.S.K., W.G., and P.J.O. performed experiments; J.T. and N.K. approved final version of manuscript.

REFERENCES

27. Reynolds CK, Huntington GB, Tyrrell HF, Reynolds PJ. Net portal-drained visceral and hepatic metabolism of glucose, L-lactate, and nitrog-


