Estimation of genetic parameters and detection of quantitative trait loci for metabolites in Danish Holstein milk

A. J. Buitenhuis,^{*1} U. K. Sundekilde,[†] N. A. Poulsen,[‡] H. C. Bertram,[†] L. B. Larsen,[‡] and P. Sørensen^{*}

^{*}Department of Molecular Biology and Genetics, Aarhus University, Blichers Allé 20, PO Box 50, DK-8830 Tjele, Denmark

[†]Department of Food Science, Aarhus University, Kirstinebjergvej 10, DK-5792 Årlev, Denmark

[‡]Department of Food Science, Aarhus University, Blichers Allé 20, PO Box 50, DK-8830 Tjele, Denmark

ABSTRACT

Small components and metabolites in milk are significant for the utilization of milk, not only in dairy food production but also as disease predictors in dairy cattle. This study focused on estimation of genetic parameters and detection of quantitative trait loci for metabolites in bovine milk. For this purpose, milk samples were collected in mid lactation from 371 Danish Holstein cows in first to third parity. A total of 31 metabolites were detected and identified in bovine milk by using ¹H nuclear magnetic resonance (NMR) spectroscopy. Cows were genotyped using a bovine high-density single nucleotide polymorphism (SNP) chip. Based on the SNP data, a genomic relationship matrix was calculated and used as a random factor in a model together with 2 fixed factors (herd and lactation stage) to estimate the heritability and breeding value for individual metabolites in the milk. Heritability was in the range of 0 for lactic acid to >0.8 for orotic acid and β -hydroxybutyrate. A single SNP association analysis revealed 7 genome-wide significant quantitative trait loci [malonate: *Bos taurus* autosome (BTA)2 and BTA7; galactose-1-phosphate: BTA2; *cis*-aconitate: BTA11; urea: BTA12; carnitine: BTA25; and glycerophosphocholine: BTA25]. These results demonstrate that selection for metabolites in bovine milk may be possible.

Key words: metabolite, quantitative trait loci, bovine milk, nuclear magnetic resonance spectroscopy

INTRODUCTION

Bovine milk is processed into a variety of dairy products and ingredients; therefore, it is important to maintain high quality in the technological and nutritional properties of milk. Milk contains water, fat, protein, vitamins, carbohydrates, and metabolites. The metabolite profile of milk depends on different factors,

including lactation stage, feed, and genetic background of cattle (McJarrow and van Amelsfort-Schoonbeek, 2004; Klein et al., 2010), which can affect the technological and nutritional properties of milk. Some of the components in milk can be changed through different feeding regimens (e.g., Alizadeh et al., 2012), whereas others are mainly genetically regulated (Pimentel et al., 2011). Many milk components are secreted by the mammary epithelial cells by exocytosis, including caseins, whey proteins, lactose, and citrate (Shennan and Peaker, 2000). Metabolites in milk often reflect the metabolic activity of the mammary gland, and examples of metabolites being synthesized primarily in the mammary gland are lactose, citric acid (Linzell et al., 1976), and orotic acid (Ahmed et al., 1978). Moreover, in late lactation or during inflammation of the mammary gland, metabolites may also enter the milk from blood or interstitial fluid through the paracellular pathway (Shennan and Peaker, 2000).

Milk metabolites have a wide range of applications in the food industry and as biomarkers for intramammary infections (Åkerstedt et al., 2011). Furthermore, the different metabolites found in milk have been found to be associated with various technological or nutritional quality parameters (Sundekilde et al., 2011). Metabolomics analyses have been shown to be valuable in food and nutrition sciences (Bertram et al., 2007; Wishart et al., 2009). Liquid chromatography-mass spectrometry metabolomic analyses of milk metabolites have identified more than 100 metabolites in milk (Boudonck et al., 2009), whereas nuclear magnetic resonance (NMR) analyses have identified fewer metabolites (Hu et al., 2004; Klein et al., 2010; Sundekilde et al., 2011). Historically, metabolites have been difficult to measure in milk by NMR spectroscopy because the fat globules and casein micelles hold the majority of the fat and protein in the milk in large aggregates with restricted mobility, which leads to low resolution and broad resonances in NMR spectroscopy (Hu et al., 2004). Recently, however, proton NMR spectroscopy has shown to be very useful for measuring a large range of different metabolites in milk (Klein et al., 2010; Sundekilde et al., 2011).

Received July 6, 2012.

Accepted January 15, 2013.

¹Corresponding author: bart.buitenhuis@agrsci.dk

To date, breeding strategies have focused on the total fat and protein contents of milk. However, the potential exists to change breeding objectives to make milk a more valuable product through selective breeding (Boichard and Brochard, 2012). In the last decade, genetic parameters have been estimated for metabolites in milk such as urea (Miglior et al., 2007; Stoop et al., 2007; Mucha and Strandberg, 2011), lactose (Miglior et al., 2007), and acetone (Wood et al., 2004), as well as for metabolites in blood such as glucose and BHBA (Oikonomou et al., 2008). In addition, QTL for urea have been detected in Dutch Holstein cattle (Bouwman et al., 2010). In the present study, tissue and milk samples collected from 371 Danish Holstein cows were examined using a bovine 777K SNP chip (Illumina Inc., San Diego, CA; http://www.illumina.com/Documents/products/datasheets/datasheet_bovineHD.pdf) combined with detailed milk metabolite profiles measured by proton NMR. The aim of the present study was to (1) estimate genetic parameters for the individual metabolites, and (2) detect QTL for the different metabolites using a genome-wide association study (GWAS) approach.

MATERIALS AND METHODS

Animals

Within the Danish-Swedish Milk Genomics Initiative, morning milk samples were collected from 456 Danish Holstein cows from October to December 2009 (Jensen et al., 2012). The overall experimental strategy underlying this study was to minimize potential sources of environmental variation and maximize the genetic variation in the sample population. The animals were distributed over 20 commercial herds in Denmark. The herds were selected in such a way that most sires (200 sires in total) from the Danish Holstein cattle population were represented. From these herds, between 19 and 24 cows were sampled. All cows were housed in loose housing systems, fed according to standard practice, and milked twice a day. The cows sampled were all in mid lactation (d 129 to 228 of lactation) and within parity 1 to 3. Immediately after milking, milk samples were placed on ice for transport to the laboratory. Once at the laboratory, the milk samples were aliquoted and skimmed, as described by Jensen et al. (2012). Skim milk samples were kept at -80°C until further analyses were conducted.

Milk Production Traits

Milk yield (MY; kg) was recorded on the day of sampling for each animal. Fat percentage (FP) and protein percentage (PP) were determined on fresh full milk

samples by MilkoScan FT2 (Foss Electric, Hillerød, Denmark). In addition, each milk sample was analyzed for SCC by using the Fossomatic 5000 (Foss Electric) at Eurofins Laboratory (Holstebro, Denmark). Milk samples with high SCC ($>500,000$ cells/mL) were excluded from the sampling population.

NMR Spectroscopy

Nuclear magnetic resonance spectroscopy was performed as described previously (Sundekilde et al., 2013). In short, before NMR spectroscopy, skim milk samples were thawed and thoroughly shaken to homogenize the sample. The samples were filtered to remove residual lipids and protein using an Amicon Ultra 0.5 mL, 10-kDa cutoff spin filter (Millipore, Billerica, MA) at $10,000 \times g$ for 15 min. Four hundred microliters of filtered sample was mixed with 200 μL of deuterium oxide (D_2O) containing 0.025% sodium trimethylsilyl-[2,2,3,3- $^2\text{H}_4$]-1-propionate (Sigma-Aldrich, Brøndby, Denmark), which served as an internal chemical shift reference. The ^1H NMR spectroscopy was performed at 298 K on a Bruker Avance III 600 spectrometer operating at an ^1H frequency of 600.13 MHz and equipped with a 5-mm ^1H TXI probe (Bruker BioSpin, Ettlingen, Germany). Standard one-dimensional spectra were acquired using a single 90° pulse experiment with a relaxation delay of 5 s. Water suppression was achieved by irradiating the water peak during the relaxation delay, and a total of 64 scans were collected into 32,000 data points that spanned a spectral width of 12.15 ppm. Before Fourier transformation, the data were multiplied by a 0.3-Hz line-broadening function. The proton NMR spectra were phase- and baseline-corrected manually using Topspin 2.1 (Bruker Biospin), and the NMR signals were assigned in accordance with existing literature, the Human Metabolome Database (<http://www.hmdb.ca/>), and the Biological Magnetic Resonance Data Bank (<http://www.bmrwisc.edu/>; Hu et al., 2004; Ulrich et al., 2008; Wishart et al., 2009; Klein et al., 2010). Relative quantification of the following well-resolved resonances was performed in Matlab 7.9 (MathWorks Inc., Natick, MA): 2-oxoglutaric acid, δ 3.01 ppm (t, $J = 6.84$ Hz, 2H); alanine, δ 1.48 ppm (d, $J = 7.26$ Hz, 3H); acetate, δ 1.92 ppm (s, 3H); BHBA, δ 1.20 ppm (d, $J = 6.29$ Hz, 3H); carnitine, δ 3.23 ppm (s, 3^*3H); *cis*-aconitate, δ 3.15 ppm (s, 2H); citric acid, 2.53 ppm (d, $J = 16.51$ Hz, 2H); choline, δ 3.21 ppm (s, 3^*3H); creatinine, δ 3.04 ppm (s, 3H); glucose, δ 5.1 ppm (d, 3.77, 2H); fucose, δ 1.25 ppm (d, $J = 6.50$, 3H); fumarate, δ 6.52 ppm (s, 2H); galactose, δ 4.54 ppm (d, $J = 7.75$, 1H); galactose-1-phosphate, δ 5.38 ppm (dd, $J = 7.34$ Hz, 3.34 Hz, 1H); glucose-1-phosphate, δ 5.51 ppm (dd, $J = 7.36$ Hz, 3.30 Hz, 1H); glutamate, δ

2.36 ppm (m, 2H); glycerophosphocholine, δ 4.33 ppm (m, 2H); hippurate, δ 7.84 ppm (dd, $J = 8.40$ Hz, 1.46 Hz, 2H), δ 7.64 (tt, $J = 7.42$ Hz, 1.45 Hz, 1H), and δ 7.54 ppm (m, 2H); isobutyrate, δ 1.16 ppm (d, $J = 6.8$, 3H); isoleucine, δ 0.93 ppm (d, $J = 7$, 3H); lactate, δ 1.33 ppm (d, $J = 7.10$ Hz, 3H); lactose, δ 5.23 ppm (d, $J = 3.85$ Hz, 1H); methionine, δ 2.15 ppm (m, 5H); malonate, δ 3.11 ppm (s, 2H); butyric acid, δ 0.90 ppm (t, $J = 7.30$ Hz, 3H); *N*-acetylcarbohydrates, δ 2.05 ppm (d, $J = 1.4$, 2*3H); ornithine, δ 1.82 ppm (m, 2H); orotic acid, δ 6.19 ppm (s, 1H); valine, δ 1.05 ppm (d, $J = 7.04$, 3H); urea, δ 5.77 ppm (br. s, 4H); and unknown1.6, δ 1.6 ppm (m, ?H). Chemical shift (δ) is expressed as the shift in parts per million (ppm) from sodium trimethylsilyl-[2,2,3,3-2H₄]-1-propionate (TSP), in which protons, by definition, have a chemical shift of 0.00 ppm; s = singlet, d = doublet, t = triplet, m = multiple, J = coupling constant, H = H atom. In cases where a metabolite could not be assigned to a known entity, it was named "unknown" followed by the quantification (ppm).

SNP Chip and Genotyping

We used the BovineHD BeadChip (Illumina Inc., San Diego, CA; Van Tassell et al., 2008) for genotyping 371 animals. The bovineHD method assays 777,962 markers with a median interval of 2.68 kb between SNP (www.illumina.com/documents/products/datasheets/datasheet_bovineHD.pdf). Genomic DNA was extracted from ear tissue. The platform used was an Illumina Infinium II Multisample assay device; the SNP chips were scanned using iScan (Illumina Inc.) and analyzed using Beadstudio version 3.1 software (Illumina Inc.). The quality parameters used for the selection of SNP in the GWAS were minimum call rates of 80% for individuals and 95% for loci. Marker loci with minor allele frequencies (MAF) <1% were excluded. The quality of the markers was assessed using the GenCall data analysis software of Illumina. The GenCall software application automatically clusters, calls genotypes, and assigns confidence scores. Individuals with average GenCall scores <0.65 were excluded following Teo et al. (2007). The SNP positions within a chromosome were based on the *Bos taurus* genome assembly (Btau_4.0, Liu et al., 2009).

Statistical Analysis

The statistical analysis was based on the 371 animals that had both phenotypic and genotypic information.

Calculation of the Genomic Relationship Matrix. The genomic relationship matrix **G** was calculated for each chromosome separately, as described

by the first method presented in VanRaden (2008). In short, let **M** be a matrix with dimensions of the number of individuals (n) by the number of loci (m) that specifies which marker alleles each individual inherited. The elements of **M** were set to -1, 0, and 1 for the homozygote, heterozygote, and the other homozygote, respectively. The diagonals of **M'M** count the number of homozygous loci for each individual and off-diagonals measure the number of alleles shared by relatives. Let the frequency of the second allele at locus i be p_i , and let **P** contain the allele frequencies, such that column i of **P** equals $2(p_i - 0.5)$. Subtraction of **P** from **M** gives **Z**, which is needed to set the expected mean value to 0. The genomic relationship matrix **G** was then calculated as **ZZ'**/[$2\sum p_i(1 - p_i)$] (VanRaden, 2008).

Estimation of Heritability. To estimate the genetic parameters and variance components, the REML approach in DMU was used (Madsen and Jensen, 2007). The following model was used in the analysis:

$$Y_{ijk} = \mu + \text{herd}_i + \text{parity}_j + \text{animal}_k + e_{ijk}, \quad [1]$$

where Y_{ijk} is the phenotype of individual k in herd i and lactation j , μ is the fixed mean effect, herd is a fixed effect ($i = 1, 2, \dots, 20$), parity is a fixed effect ($j = 1, 2, 3$), and animal is the random additive genetic effect based on **G** of animal k .

Univariate analyses were performed to estimate the heritability, which was defined as

$$h^2 = \sigma_a^2 / (\sigma_a^2 + \sigma_e^2), \quad [2]$$

where σ_a^2 was the genetic variation and σ_e^2 was the residual variation.

Genetic and phenotypic correlations were studied by fitting a series of bivariate analyses with a REML approach in DMU (Madsen and Jensen, 2007).

Association Mapping. The analysis was done using the following linear model:

$$Y_{ijk} = \mu + \text{herd}_i + \text{parity}_j + b \times \text{SNP} + \text{animal}_k + e_{ijk}, \quad [3]$$

where Y_{ijk} is the phenotype of individual k in herd i and lactation j , μ is the fixed mean effect, herd is a fixed effect ($i = 1, 2, \dots, 20$), parity is a fixed effect ($j = 1, 2, 3$), b is the allele substitution effect, which is a count in individual k of 1 of the 2 alleles (with arbitrary labeling), and animal is the random additive genetic effect based on **G** of animal k (Yang et al., 2010). The effect of the SNP was tested by a Wald test with a null hypothesis, H_0 , of $b = 0$. Significance thresholds were

Table 1. Estimates of heritability, genetic variances, and phenotypic mean, standard deviation, and coefficient of variation for relative metabolite levels and the milk production traits milk yield (kg), fat percentage, and protein percentage in Danish Holstein bovine milk¹

Trait	h ²	SE_h ²	Gen_VAR	SE_Gen_VAR	Mean	SD	CV
Milk yield	0.31	0.15	3.20	1.61	14.77	3.77	0.25
Fat percentage	0.39	0.17	0.21	0.09	3.99	0.78	0.19
Protein percentage	0.51	0.16	0.03	0.01	3.43	0.26	0.08
Lactic acid	2.77E-05	0.13	2.26E-08	1.50E-04	3.97E-02	3.13E-02	0.79
Isobutyrate	0.02	0.10	6.00E-07	7.75E-04	3.87E-03	6.44E-03	1.66
Acetic acid	0.02	0.10	1.70E-06	1.30E-03	1.25E-02	1.00E-02	0.8
Fumaric acid	0.06	0.00	1.67E-08	1.29E-04	6.68E-03	2.71E-03	0.41
Galactose	0.11	0.14	1.69E-05	4.11E-03	4.88E-02	1.24E-02	0.25
Lactose	0.17	0.12	1.14E+00	1.07E+00	3.81E+01	2.66E+00	0.07
Valine	0.17	0.16	6.56E-07	8.10E-04	2.28E-03	1.97E-03	0.86
Butyrate	0.19	0.13	5.27E-07	7.26E-04	5.35E-03	1.72E-03	0.32
2-Oxoglutaric acid	0.22	0.15	2.89E-05	5.38E-03	3.40E-02	1.27E-02	0.37
Unknown1.6	0.24	0.19	1.59E-06	1.26E-03	5.46E-03	2.68E-03	0.49
Urea	0.27	0.16	3.59E-04	1.90E-02	2.01E-01	4.47E-02	0.22
Methionine	0.27	0.16	9.58E-06	3.09E-03	3.24E-02	7.05E-03	0.22
Ornithine	0.33	0.17	1.46E-05	3.82E-03	2.76E-02	6.98E-03	0.25
Alanine	0.35	0.17	1.11E-05	3.33E-03	2.12E-02	5.95E-03	0.28
Isoleucine	0.39	0.17	1.52E-06	1.23E-03	4.51E-03	2.06E-03	0.45
Fucose	0.39	0.16	5.93E-05	7.70E-03	1.24E-01	1.40E-02	0.11
Galactose-1-phosphate	0.43	0.17	1.44E-04	1.20E-02	2.31E-02	1.99E-02	0.86
Creatinine	0.45	0.19	1.85E-03	4.30E-02	3.52E-01	6.75E-02	0.19
Choline	0.48	0.18	1.40E-02	1.18E-01	4.71E-01	1.78E-01	0.37
Glycerophosphocholine	0.48	0.19	1.72E-03	4.14E-02	3.82E-01	6.40E-02	0.17
cis-Aconitate	0.48	0.18	4.16E-04	2.04E-02	1.29E-01	5.46E-02	0.42
Malonate	0.48	0.21	6.22E-05	7.89E-03	4.73E-02	1.20E-02	0.25
Glucose-1-phosphate	0.49	0.18	5.41E-04	2.33E-02	1.16E-01	3.67E-02	0.31
N-Acetyl-carbohydrates	0.50	0.18	5.69E-03	7.55E-02	8.66E-01	1.07E-01	0.12
Glutamate	0.52	0.19	3.79E-04	1.95E-02	6.98E-02	2.98E-02	0.43
Hippuric acid	0.53	0.20	3.48E-04	1.87E-02	8.24E-02	2.86E-02	0.34
Citric acid	0.54	0.19	3.44E-01	5.87E-01	4.81E+00	8.71E-01	0.18
Glucose	0.55	0.19	4.25E-05	6.52E-03	4.92E-02	8.97E-03	0.18
Carnitine	0.63	0.21	4.69E-02	2.17E-01	1.36E+00	2.82E-01	0.21
Orotic acid	0.86	0.21	6.78E-04	2.60E-02	8.32E-02	3.22E-02	0.37
BHBA	0.87	0.25	1.89E-03	4.35E-02	2.13E-02	4.49E-02	2.11

¹SE_h² = standard error of the heritability estimate; Gen_VAR = genetic variation; SE_Gen_VAR = standard error of the estimate of the genetic variation.

determined using a Bonferroni correction; genome-wide significance thresholds were obtained by dividing the nominal significance threshold of 0.05 by the numbers of SNP included in the analysis. The analyses were performed using REML in the R interface of DMU (Madsen and Jensen, 2007; <http://dmu.agrsci.dk>).

Identification of Genes in QTL Regions. To identify biological information of the QTL, the mapped genes for those regions that had genome-wide significance at the $P < 0.05$ level (after Bonferroni correction for multiple testing) were downloaded from the *Bos taurus* genome 6.1 (<http://www.ncbi.nlm.nih.gov/genome/82>). All genes available between the upper and lower boundary markers were downloaded. Based on the map location of the genes, we identified whether a significant SNP was located in a gene or not.

RESULTS

Milk Metabolites

In total, 31 metabolites were identified in milk (Table 1). Each metabolite was relatively quantified by inte-

gration of NMR resonances. The metabolites with the highest CV were BHBA (2.11) and isobutyrate (1.66), whereas lactose had the lowest CV (0.08).

SNP Markers

A total of 588,528 SNP markers on 29 bovine autosomes passed quality control and had a minor allele frequency ≥ 0.01 and a GenCall score > 0.65 . The number of SNP markers ranged from 11,093 on BTA25 to 37,244 on BTA1.

Heritability

The heritabilities for the different metabolites are presented in Table 1. Five metabolites had low heritability (lactic acid, isobutyrate, acetic acid, fumaric acid, and galactose) in the range of 2.7×10^{-5} to 0.11, whereas the standard error was in the range of 0 to 0.14. The remaining metabolites had medium to high heritability, ranging from 0.17 for lactose to > 0.8 (orotic acid and BHBA). The phenotypic and genetic variation for the

metabolites was small (Table 1). The heritability for the production traits MY, FP, and PP were 0.31, 0.39, and 0.51, respectively.

Genetic and Phenotypic Correlations

Genetic and phenotypic correlations are presented in Supplementary Table S1 (available online at <http://www.journalofdairyscience.org/>). Phenotypic correlation was generally low (Supplementary Table S1). The standard error for the genetic correlation among the metabolites ranged from 0.66 to 0.98. The bivariate model for some of the combinations did not converge (denoted as Table S1 as NC). Carnitine was highly correlated with choline (0.86), isoleucine (0.79), ornithine (0.68), and lactose (0.58), whereas it showed low correlation with methionine and hippuric acid. In addition to carnitine, lactose was highly positively correlated with isoleucine (0.69) and orotic acid (0.70), whereas lactose showed high negative correlations with BHBA (−0.57), galactose-1-phosphate (−0.71), and urea (−0.57). Urea was highly positively correlated with galactose-1-phosphate (0.99) and butyrate (0.95), whereas it had a negative correlation with BHBA, ornithine, *N*-acetylcarbohydrate, methionine, and citric acid. In general, the phenotypic correlation between the metabolites was smaller compared with the genetic correlation. β -Hydroxybutyrate was highly positively correlated with choline (0.97) and citric acid (0.54), whereas it was highly negatively correlated with alanine (−0.82), valine (−0.85), and malonate (−0.93).

The genetic correlations between the metabolites and the production traits revealed some patterns of interest (Supplementary Table S1; <http://www.journalofdairyscience.org/>). Citric acid showed low correlation with MY (−0.04), whereas it was negatively correlated with FP (−0.47) and PP (−0.17). Choline showed high positive correlation with MY (0.43) but low correlation with FP (0.07) and PP (−0.09). Creatinine was highly positively correlated with MY (0.76), showed low correlation with FP (0.08), and was negatively correlated with PP (−0.51). Glycerophosphocholine was highly negatively correlated with MY (−0.81), moderately positively correlated with FP (0.2), and highly correlated with PP (0.62). Galactose showed a negative correlation with MY (−0.57), a low correlation with FP (−0.07), and a positive correlation with PP (0.12). Lactose was highly correlated with MY (0.64) and FP (0.53) and negatively correlated with PP (−0.19). Glucose was moderately positively correlated with MY (0.32), FP (0.38), and PP (0.45). Urea was negatively correlated with MY (−0.77), had a low correlation with FP (0.03), and was positively correlated with PP (0.26).

QTL

A whole-genome scan was performed for the 31 individual milk metabolites. In total, 29 QTL were detected for a total of 19 metabolites with SNP associations. Eight of these had genome-wide significant QTL (see Supplementary Figures S1 to S8; <http://www.journalofdairyscience.org/>), whereas the remaining 21 only had chromosome-wide significance (Table 2) and were distributed across 17 chromosomes. Three QTL were detected for orotic acid (BTA1, BTA5, and BTA19). Two QTL were detected for malonate, glycerophosphocholine, choline, fucose, isobutyrate, and glucose. Only 1 QTL was detected for citric acid, urea, *N*-acetylcarbohydrates, and glutamate. Moreover, carnitine and glycerophosphocholine had the same genome-wide SNP peak (BOVINEHD2500007432) on BTA25, and galactose-1-phosphate and glucose-1-phosphate exhibited the same SNP peak (BOVINEHD0200038319) on BTA2. The SNP peak on BTA2 for galactose-1-phosphate was assigned at the genome-wide significance level (Table 2), whereas the SNP peak for glucose-1-phosphate was detected only at the chromosome-wide significance level.

Underlying Genes

For the 8 genome-wide significant QTL, gene information between the upper and lower boundary markers (Table 2) was downloaded from the National Center for Biotechnology Information website (<http://www.ncbi.nlm.nih.gov/genome/?term=Bos+Taurus>). The most significant SNP for the QTL for orotic acid on BTA1 was located within *ITGB5* (integrin β 5). The most significant SNP for the QTL for malonate on BTA2 was located between *FN1* (fibronectin 1) and *LOC100848551* (40S ribosomal protein S20-like), and the most significant SNP for the QTL for galactose-1-phosphate on BTA2 was located within *RAP1GAP* (RAP1 GTPase activating protein). Furthermore, within the QTL region of galactose-1-phosphate on BTA2, *GALE* (UDP-galactose-4-epimerase) was located. The most significant SNP for the QTL for malonate on BTA7 was located close to *LOC100296630* (pancreatic progenitor cell differentiation and proliferation factor-like). In addition, the most significant SNP for the QTL for glucose on BTA11 was located between *LOC100848307* (uncharacterized) and *ABO* (ABO blood group). The *ABO* gene is a glycosyl transferase associated with ABO blood group but its function is relevant to the observed signal. The most significant SNP for the QTL for urea on BTA12 was located between *SERTM1* (serine-rich and transmembrane

Table 2. Quantitative trait loci detected for metabolites in the milk of Danish Holstein cattle

Trait ¹	Chr ²	ub_marker ³	ub_pos ⁴	lb_marker ³	lb_pos ⁴	Most significant SNP ³	Most significant SNP_pos ⁴	b ⁵	SE ⁶	P-value
N-Acetyl-carbohydrates	1	BOVINEHD0100026452	93073554	BOVINEHD0100008061	27089906	BOVINEHD0100015368	54565647	-0.046	0.008	5.33E-08
Orotic acid	1	BOVINEHD0100024509	85725626	BOVINEHD0100017968	63391092	BOVINEHD0100019957	69802046	-0.019	0.003	5.231E-12
Malonate	2	BOVINEHD0200030267	105149497	BOVINEHD0200029587	1.03E+08	BOVINEHD0200029956	104224898	-0.017	0.003	1.58E-08
Glucose-1-phosphate	2	BOVINEHD0200038635	132829873	BOVINEHD0200005158	18051846	BOVINEHD0200038319	131759549	-0.021	0.004	6.91E-08
Galactose-1-phosphate	2	BOVINEHD0200039895	136105398	BOVINEHD0200005158	18051846	BOVINEHD0200038319	131759549	-0.013	0.002	1.03E-10
Choline	3	BOVINEHD0300028487	99019530	BOVINEHD0300026266	91312413	BOVINEHD0300028476	98992612	0.074	0.014	5.31E-07
Orotic acid	5	BOVINEHD0500034735	119082435	BOVINEHD0500012380	43257749	BOVINEHD0500025542	89907365	-0.014	0.002	4.61E-08
Glucose	6	BOVINEHD0600033248	116813163	BOVINEHD0600011192	41252410	BOVINEHD0600011298	41789941	-0.005	0.001	1.68E-06
Glutamate	6	BOVINEHD0600012986	47480428	BOVINEHD0600001894	8300558	BOVINEHD0600012934	47259879	-0.012	0.002	1.82E-07
Malonate	7	BOVINEHD0700017272	60353695	BOVINEHD0700016099	55794763	BOVINEHD0700016099	55794763	-0.017	0.003	3.24E-09
Isoleucine	8	BOVINEHD0800009205	30342164	BOVINEHD0800006421	20882946	BOVINEHD0800006428	20912826	0.001	0.000	1.44E-07
Unknown1.6	8	BOVINEHD0800018119	60699612	BOVINEHD0800018119	60699612	BOVINEHD0800018119	60699612	-0.004	0.001	3.53E-07
Glucose	11	BOVINEHD1100031139	107078911	BOVINEHD1100023502	82044909	BOVINEHD1100030289	104215432	-0.009	0.001	1.16E-43
Unknown1.6	12	BOVINEHD1200005193	17053068	BOVINEHD1200003335	11647451	BOVINEHD1200005165	16985323	-0.005	0.001	5.76E-08
Urea	12	BOVINEHD1200008386	28150597	BOVINEHD1200007549	24956951	BOVINEHD1200007579	25062022	-0.021	0.003	4.25E-09
Choline	14	BOVINEHD1400001298	5238746	BOVINEHD1400000188	1588879	BOVINEHD1400000200	1640406	-0.074	0.014	2.32E-07
Orotic acid	19	BOVINEHD1900000983	4147200	BTA-45069-NO-RS	1439012	BTA-45069-NO-RS	1439012	-0.021	0.004	1.22E-07
Galactose-1-phosphate	19	BOVINEHD1900008997	30636595	BOVINEHD1900008993	30619505	BOVINEHD1900008997	30636595	0.009	0.002	4.95E-07
Galactose-1-phosphate	20	BOVINEHD2000008181	27721766	ARS-BFGL-NGS-110286	25689084	BOVINEHD2000007723	25780608	0.009	0.002	5.46E-07
Citric acid	20	BOVINEHD2000016721	59544815	BOVINEHD2000003036	9556688	BOVINEHD2000016330	58635840	-0.333	0.061	2.26E-07
Fucose	20	BOVINEHD2000020537	70239309	BOVINEHD2000017933	63187854	BOVINEHD2000020519	70210754	-0.005	0.001	1.48E-06
Fucose	23	BOVINEHD2300015161	52207889	BOVINEHD2300003454	13921728	BOVINEHD2300015161	52207889	0.005	0.001	5.42E-07
Isobutyrate	24	BOVINEHD2400009385	33884239	BOVINEHD2400009370	33832320	BOVINEHD2400009370	33832320	-0.010	0.002	4.61E-08
Carnitine	25	BOVINEHD2500007695	27555964	BOVINEHD2500006639	23698091	BOVINEHD2500007432	26308666	0.164	0.025	8.24E-10
Glycerophosphocholine	25	BOVINEHD2500008585	31276724	BOVINEHD2500005163	18386179	BOVINEHD2500007432	26308666	0.043	0.006	3.25E-13
2-Oxoglutaric acid	26	BOVINEHD2600015144	17949481	BOVINEHD2600003855	15094706	BOVINEHD2600003855	15094706	-0.008	0.002	2.30E-06
Isobutyrate	26	BOVINEHD2600007470	28048729	BOVINEHD2600004486	17628241	BOVINEHD2600007000	26365363	-0.010	0.002	1.71E-09
Glycerophosphocholine	29	BOVINEHD2900003286	11367118	BOVINEHD2900003279	11336592	BOVINEHD2900003279	11336592	0.062	0.011	1.86E-07
Unknown1.6	29	BOVINEHD2900013162	43955161	BOVINEHD2900013162	43955161	BOVINEHD2900013162	43955161	-0.005	0.001	3.23E-07

¹Traits in bold have genome-wide significance (Bonferroni $P < 0.05$); the other traits have chromosome-wide significance (Bonferroni $P < 0.05$).

²Chr = chromosome (BTA).

³ub_marker = upper boundary marker name; lb_marker = lower boundary marker name; most significant SNP = marker name of the most significant marker.

⁴ub_pos = upper boundary marker position (bp); lb_pos = lower boundary marker position (bp); most significant SNP_pos = marker position (bp) of the most significant marker based on Btau_4 (Liu et al., 2009).

⁵b = allele-substitution effect.

⁶SE = standard error of the allele substitution effect.

domain containing 1) and *CCNA1* (cyclin A1). Finally, the most significant SNP for the QTL for carnitine and glycerophosphocholine on BTA25 was located within *CLN3* (ceroid-lipofuscinosis, neural 3).

DISCUSSION

Heritability and Correlation

New analytical techniques such as NMR spectroscopy have been successfully used to identify the metabolite milk profile (Hu et al., 2004). To date, knowledge regarding the genetic regulation of metabolites in milk has been limited. In this study, we have addressed the genetic influence on individual metabolites in milk based on a moderate sample size. This is reflected in the relatively large standard errors on the estimates. In general, we found large variation in the heritability of the individual components of the metabolic profile (Table 1). Metabolic profiles of bovine milk have focused on urea (Miglior et al., 2007; Stoop et al., 2007; Mucha and Strandberg, 2011) and lactose (Miglior et al., 2007). In our study, the heritability for urea was 0.27, which was higher than that found by Stoop et al. (2007; 0.14) and by Mucha and Strandberg, (2011; 0.16–0.18) but lower than that found by Miglior et al. (2007). For lactose, heritability was considerably lower in our study (0.17) compared with the study of Miglior et al. (2007). Two factors might explain the differences between the studies. First, Stoop et al. (2007), Miglior et al. (2007), and Mucha and Strandberg (2011) used infrared methods to measure urea and lactose, whereas we used NMR to measure the metabolites in the milk. Second, for Stoop et al. (2007), Miglior et al. (2007), and Mucha and Strandberg (2011), the spectrum of DIM spanned the entire lactation period, whereas in our study, data were collected only from mid-lactation cows.

Recently, Kettunen et al. (2012) estimated heritabilities of human serum metabolites using NMR. In general, they found that heritability estimates ranged from 0.23 to 0.55 for amino acids and other small molecule metabolites, whereas those for lipids ranged from 0.48 to 0.62 and for lipoproteins ranged from 0.50 to 0.76 (Kettunen et al., 2012). The heritability estimates for metabolites measured in our study ranged from 0 to 0.87. Although only 6 metabolites measured in our study were in common with those measured in Kettunen et al. (2012), the heritability estimates for alanine and creatinine were similar, whereas those for urea, isoleucine, and glutamate were lower and that for BHBA higher in our study compared with the previous study. Oikonomou et al. (2008) showed that the heritability of blood metabolites glucose and BHBA

was highly dependent on lactation stage. Heritability estimates for glucose in blood ranged from 0.12 at the beginning of the lactation to 0.39 at the end of the lactation (Oikonomou et al., 2008). These estimates are much lower compared with that estimated in our study for glucose in milk (0.55). The estimate of BHBA in the blood ranged from 0.08 after 7 wk of lactation to 0.40 in the first 7 wk of the lactation (Oikonomou et al., 2008). These estimates are much lower compared with the heritability estimates for BHBA in the milk estimated in our study (0.87). However, it has been shown in rats that the lactating gland is a potential major site of ketone-body utilization (Page and Williamson, 1972). Those authors observed that ketone bodies were used to supply carbon for fat synthesis when carbohydrates were limited.

The heritability estimates for the production traits (MY, FP, PP) were in line with those detected by Stoop et al. (2008). Citric acid was correlated with FP (−0.47) but showed lower correlations with MY and PP. This finding is in line with the fact that citric acid is involved in fatty acid metabolism. The conjugated base of citric acid is citrate, which is important in fat metabolism. Citrate is transported to the cytoplasm, where it is converted to acetyl CoA. Acetyl CoA is then converted into malonyl CoA by acetyl CoA carboxylase. Acetyl CoA carboxylase, in turn, is allosterically modulated by citrate.

Miglior et al. (2007) found a correlation between MUN and MY of −0.09, whereas the correlations of FP and PP with MUN were 0.43 and 0.20, respectively. Stoop et al. (2007) showed correlations between MUN and FP and PP of 0.19 and 0.27, respectively. Results from our study showed very different correlations for urea with MY and FP: −0.77 and 0.03, respectively, compared with those of Miglior et al. (2007) and Stoop et al. (2007). The correlation between urea and PP in our study was similar to that reported by Miglior et al. (2007) and Stoop et al. (2007).

QTL

Interpretation of results from a QTL experiment using metabolites as a phenotype is complicated, because metabolite concentrations are influenced indirectly by mRNA and protein expression. As a result, a one-to-one metabolite-to-gene correspondence is not known or expected a priori (Nicholson et al., 2011). In this study, we showed that carnitine and glycerophosphocholine had a high genetic correlation (−0.59). Furthermore, for both metabolites, a QTL was detected at BTA2, which showed the same SNP with the greatest significance. This SNP was located in the *CLN3* gene, which encodes a protein involved in lysosomal function. Mutations in

this gene cause neurodegenerative diseases, such as neuronal ceroid lipofuscinosis (http://www.ncbi.nlm.nih.gov/sites/entrez?db=gene&cmd=Retrieve&dopt=full_report&list_uids=1201). Humans with a juvenile form of neuronal ceroid lipofuscinosis have reduced plasma levels of carnitine (Katz, 1996), suggesting that *CLN3* regulates carnitine levels in plasma. Both carnitine and glycerophosphocholine play a role in fatty acid metabolism. Whether *CLN3* also regulates the carnitine or glycerophosphocholine in bovine milk remains unknown; however, these findings suggest that *CLN3* may be a good candidate gene underlying this QTL.

Orotic acid was originally identified as vitamin B₁₃ (Novak and Hauge, 1948). However, evidence has emerged that orotic acid is not a vitamin and can accordingly be produced by the animal. Orotic acid is known to be synthesized and released by the mammary gland (Ahmed et al., 1978), and it is an intermediate in the biosynthesis of pyrimidine nucleotides (cytosine, thymine, and uracil; Jones, 1980). A large variability in orotic acid levels has been described in milk from individual cows (Jesse et al., 1980). This variability could to some extent be ascribed to breed, number of lactations, and stage of lactation (Jesse et al., 1980). Elevated levels of orotic acid have been shown to be associated with partial deficiency in uridine monophosphate synthase (Robinson et al., 1984). Thus, this deficiency decreases the synthesis of pyrimidine nucleotides (thymine, cytosine, and uracil), resulting in developmental defects. The QTL analysis identified 1 QTL with genome-wide significance on BTA1 and 2 QTL that had chromosome-wide significance associated with orotic acid (on BTA5 and BTA19). The most significant SNP of the QTL on BTA1 was located within *ITGB5*, whereas the QTL on BTA5 harbored the phosphodiesterase 3A gene (*PDE3A*) and the QTL at BTA19 harbored the carbonic anhydrase X (*CA10*) gene. Integrins are receptor proteins that mediate cell-extracellular matrix and cell-cell interactions; *ITGB5* is a receptor for fibronectin. The relation to orotic acid is unclear; however, *ITGB5* is a good candidate gene for the QTL on BTA1, because the most significant SNP for this QTL is located within *ITGB5*. Phosphodiesterases (PDE) are normally known to hydrolyze cyclic nucleotides such as cyclic AMP (cAMP) and cyclic guanosine monophosphate (cGMP), and recently evidence was provided for phosphodiesterases being able to hydrolyze other cyclic nucleotides as well, such as cyclic cytosine monophosphate MP or cyclic uridine monophosphate (Reinecke et al., 2011). Whether *PDE3A* directly affects the levels of orotic acid in milk remains unknown, but it is a candidate gene for further investigation.

Galactose-1-phosphate and glucose-1-phosphate were found to have a genetic correlation of -0.36 . Both metabolites, together with glucose and galactose, play an important role in galactose metabolism (KEGG map00052; http://www.genome.jp/dbget-bin/www_bget?map00052). The *GALE* gene is located within the QTL region for galactose-1-phosphate on BTA2; *GALE* is a uridine monophosphate-glucose 4-epimerase that catalyzes the reaction uridine diphosphate (UDP)-glucose into UDP-galactose. However, the most significant SNP of the QTL for galactose-1-phosphate on BTA2 was the same as that for glucose-1-phosphate. The function of the underlying gene *RAP1GAP* has been described as a GTPase activator for the nuclear Ras-related regulatory protein *RAP-1A* by converting it to an inactive *GDP*-bound state (<http://www.genecards.org/cgi-bin/carddisp.pl?gene=RAP1GAP&search=RAP1GAP>). However, the mechanism by which this gene is related to galactose-1-phosphate and glucose-1-phosphate levels in bovine milk remains unclear.

Previously, 6 QTL have been detected for MUN in Holstein cattle (Bouwman et al., 2010). However, none of these QTL overlapped the QTL for urea detected in this study on BTA12 (Table 2). It is difficult to pinpoint the reason for this. The number of animals in our study might have been too low to be able to detect QTL on BTA1, 6, 21, and 23, as described by Bouwman et al. (2010). Furthermore, the method to detect urea in the milk was different; Bouwman et al. (2010) used infrared spectroscopy, whereas we used NMR technology to detect urea. In our study, we chose mid-lactation cows distributed over different parities, whereas Bouwman et al. (2010) used first-lactation cows distributed over the whole lactation. The heritability of MUN is stable over lactation (Mucha and Strandberg, 2011), but this does not imply that the same genes are playing a role during the whole lactation. Thus, differences could exist between studies.

Metabolites and Disease

In dairy cattle, metabolites in the milk, together with other factors, are used as indicators of a disease state. It has been shown that ketone bodies, such as BHBA, acetone, and acetoacetate, are biomarkers for subclinical ketosis (Geishauser et al., 2000; Enjalbert et al., 2001). Recently, Klein et al. (2012) investigated the ratio of glycerophosphocholine to phosphocholine as a prognostic biomarker for risk of ketosis development in different cattle breeds. The advantage of using glycerophosphocholine and phosphocholine is that these factors are independent of acute ketosis. Glycerophosphocholine and phosphocholine are measures for the

availability of phospholipid breakdown as a fatty acid resource to meet the enhanced energy requirements of early lactation. Klein et al. (2012) suggested that the selection of animals with high glycerophosphocholine levels for a breeding strategy should generate animals that are better able to cope with a negative energy balance. Our study showed that glycerophosphocholine in the milk of Holstein cattle has a heritable component, suggesting the potential to select for increased glycerophosphocholine levels.

Nonesterified fatty acids and BHBA are the primary metabolites used to describe energy profiles in dairy cattle (Jackson et al., 2011). Urea is also produced in the mammary gland through arginase activity (Yip and Knox, 1972). Arginase is an enzyme that catalyzes the hydrolysis of arginine during the urea cycle, producing urea and ornithine. It is unclear, however, how much that contributes to milk urea. Urea is the metabolic end product of protein catabolism and is thought to reflect the dietary balance between effective rumen-degradable protein and fermentable ME (Butler, 1998). Jackson et al. (2011) suggested that high NEFA values and low urea-to-BHBA ratios can be used before calving as predictors of an increased risk of endometritis, whereas high NEFA and urea values and low urea-to-BHBA ratios are more useful after calving as predictors of an increased risk of cystic ovarian disease and delayed commencement of luteal activity. Our study showed that both urea and BHBA content in the milk are heritable, and that these 2 metabolites have a low genetic correlation (0.03). This is in line with the correlations with the milk production traits, which are very different for urea and BHBA (Supplementary Table S1; <http://www.journalofdairyscience.org/>). However, before selecting on indicator traits, more studies are needed to characterize the genetic correlation between the indicator metabolites and cystic ovarian disease and delayed commencement of luteal activity.

Metabolites Used in the Food Industry

In this study we found that many metabolites had an underlying heritable component. Some of these metabolites, such as choline and citric acid, play important roles in food quality and in the dairy industry. We showed that choline was heritable (0.48), and a chromosome-wide significant QTL was detected on BTA3 as well as on BTA14. Choline is an essential nutrient in the diet (Zeisel et al., 2003). It acts as a precursor of the messenger acetylcholine and is highly abundant in cell membranes. Choline is crucial during early development of offspring, and it prevents liver and muscle damage and other health risks in adults (Zeisel, 2006). Furthermore, we showed that citric acid

was heritable. Citric acid is a weak organic acid used to add a sour taste to food or soft drinks. It is known to reduce the storage modulus of rennet-induced gels, thereby affecting the technological properties of milk (Udabage et al., 2001), and it has been reported that that the citrate content is higher in milk with poor coagulation properties (Frederiksen et al., 2011). This tendency was confirmed in a subsequent study using NMR (Sundekilde et al., 2011), and it was postulated that citrate acts as a calcium chelating agent in milk, thereby solubilizing colloidal calcium phosphate and disrupting the casein micelle structure, which in turn affects the molecular interactions necessary for good coagulation properties (Udabage et al., 2001; Frederiksen et al., 2011).

CONCLUSIONS

This study provides evidence for genetic components underlying the metabolic profile in bovine milk. In addition, we found that specific chromosomal regions influence the metabolic profile of milk. Heritability estimates ranged from 0 for lactic acid to 0.87 for BHBA. In total, 7 genome-wide significant chromosomal regions were detected for 6 metabolites, and 1 candidate gene (*CLN3*) was identified for carnitine. Heritabilities, variances, correlations, and SNP associations from the present study are essential prerequisites for the evaluation and subsequent implementation of metabolic traits in breeding strategies.

ACKNOWLEDGMENTS

This study is part of the Danish-Swedish Milk Genomics Initiative (www.milkgenomics.dk) supported by the Danish Agency for Science, Technology and Innovation (Copenhagen); Danish Cattle Federation (Aarhus); and Arla Foods (Viby J, Denmark), as well as part of the project "Phenotypic and genetic markers for specific milk quality parameters" of Mælkeafgiftsfonden (Denmark).

REFERENCES

- Ahmed, A. A., G. A. Porter, and R. D. McCarthy. 1978. Isolation, quantification, and biosynthetic origin of orotic acid in milk. *J. Dairy Sci.* 61:39–43.
- Åkerstedt, M., L. Forsbäck, T. Larsen, and K. Svennersten-Sjaunja. 2011. Natural variation in biomarkers indicating mastitis in healthy cows. *J. Dairy Res.* 78:88–96.
- Alizadeh, A. R., M. Alikhani, G. R. Ghorbani, H. R. Rahmani, L. Rashidi, and J. J. Loo. 2012. Effect of feeding roasted safflower seeds (variety IL-111) and fish oil on dry matter intake, performance and milk fatty acid profiles in dairy cattle. *J. Anim. Physiol. Anim. Nutr. (Berl.)* 96:466–473.
- Bertram, H. C., C. Hoppe, B. O. Petersen, J. Ø. Duus, C. Mølgaard, and K. F. Michaelsen. 2007. An NMR-based metabonomic investi-

- gation on effects of milk and meat protein diets given to 8-year-old boys. *Br. J. Nutr.* 97:758–763.
- Boichard, D., and M. Brochard. 2012. New phenotypes for new breeding goals in dairy cattle. *Animal* 6:544–550.
- Boudonck, K. J., M. W. Mitchell, J. Wulff, and J. A. Ryals. 2009. Characterization of the biochemical variability of bovine milk using metabolomics. *Metabolomics* 5:375–386.
- Bouwman, A. C., G. C. B. Schopen, H. Bovenhuis, M. H. P. W. Visker, and J. A. M. van Arendonk. 2010. Genome-wide scan to detect quantitative trait loci for milk urea nitrogen in Dutch Holstein-Friesian cows. *J. Dairy Sci.* 93:3310–3319.
- Butler, W. R. 1998. Review: Effect of protein nutrition on ovarian and uterine physiology in dairy cattle. *J. Dairy Sci.* 81:2533–2539.
- Enjalbert, F., M. C. Nicot, C. Bayourthe, and R. Monocoulon. 2001. Ketone bodies in milk and blood of dairy cows: Relationship between concentrations and utilization for detection of subclinical ketosis. *J. Dairy Sci.* 84:583–589.
- Frederiksen, P. D., K. K. Andersen, M. Hammershøj, H. D. Poulsen, J. Sørensen, M. Bakman, K. B. Qvist, and L. B. Larsen. 2011. Composition and effect of blending of non-, poorly and well coagulating bovine milk from individual Danish Holstein cows. *J. Dairy Sci.* 94:4787–4799.
- Geishauser, T., K. Leslie, J. Tenhag, and A. Bashir. 2000. Evaluation of eight cow-side ketone tests in milk for detection of subclinical ketosis in dairy cows. *J. Dairy Sci.* 83:296–299.
- Hu, F., K. Furihata, M. Ito-Ishida, S. Kaminogawa, and M. Tanokura. 2004. Nondestructive observation of bovine milk by NMR spectroscopy: Analysis of existing states of compounds and detection of new compounds. *J. Agric. Food Chem.* 52:4969–4974.
- Jackson, R. A., J. R. Wills, N. R. Kendall, M. J. Green, R. D. Murray, and H. Dobson. 2011. Energy metabolites in pre- and postpartum dairy cattle as predictors of reproductive disorders. *Vet. Rec.* 168:562.
- Jensen, H. B., J. W. Holland, N. A. Poulsen, and L. B. Larsen. 2012. Milk protein variants and isoforms identified in bovine milk representing extremes in coagulation properties. *J. Dairy Sci.* 95:2891–2903.
- Jesse, B. W., C. R. Anderson, and J. L. Robinson. 1980. Bovine milk orotate: Differences between cows and changes during lactation. *J. Dairy Sci.* 63:235–242.
- Jones, M. E. 1980. Pyrimidine nucleotide biosynthesis in animals: Genes, enzymes, and regulation of UMP biosynthesis. *Annu. Rev. Biochem.* 49:253–279.
- Katz, M. L. 1996. Decreased plasma carnitine and trimethyl-L-lysine levels associated with lysosomal accumulation of trimethyl-L-lysine containing protein in Batten disease. *Biochim. Biophys. Acta* 1317:192–198.
- Kettunen, J., T. Tukiainen, A. P. Sarin, A. Ortega-Alonso, E. Tikkanen, L. P. Lyytikäinen, A. J. Kangas, P. Soiminen, P. Würtz, K. Silander, D. M. Dick, R. J. Rose, M. J. Savolainen, J. Viikari, M. Kähönen, T. Lehtimäki, K. H. Pietiläinen, M. Inouye, M. I. McCarthy, A. Jula, J. Eriksson, O. T. Raitakari, V. Salomaa, J. Kaprio, M. R. Järvelin, L. Peltonen, M. Perola, N. B. Freimer, M. Ala-Korpela, A. Palotie, and S. Ripatti. 2012. Genome-wide association study identifies multiple loci influencing human serum metabolite levels. *Nat. Genet.* 44:269–276.
- Klein, M. S., M. F. Almstetter, G. Schlamberger, N. Nürnberger, K. Dettmer, P. J. Oefner, H. H. D. Meyer, S. Wiedemann, and W. Gronwald. 2010. Nuclear magnetic resonance and mass spectrometry-based milk metabolomics in dairy cows during early and late lactation. *J. Dairy Sci.* 93:1539–1550.
- Klein, M. S., N. Buttcher, S. P. Miernicki, A. K. Immervoll, C. Louis, S. Wiedemann, W. Junge, G. Thaller, P. J. Oefner, and W. Gronwald. 2012. NMR metabolic analysis of dairy cows reveals milk glycerophosphocholine to phosphocholine ratio as prognostic biomarker for risk of ketosis. *J. Proteome Res.* 11:1373–1381.
- Linzell, J. L., T. B. Mepharm, and M. Peaker. 1976. The secretion of citrate into milk. *J. Physiol.* 260:739–750.
- Liu, Y., X. Qin, X. Z. Song, H. Jiang, Y. Shen, K. J. Durbin, S. Lien, M. P. Kent, M. Sodeland, Y. Ren, L. Zhang, E. Sodergren, P. Havelak, K. C. Worley, G. M. Weinstock, and R. A. Gibbs. 2009. *Bos taurus* genome assembly. *BMC Genomics* 10:180.
- Madsen, P., and J. Jensen. 2007. A User's Guide to DMU: A Package for Analysing Multivariate Mixed Models. Version 6, release 4.7. Aarhus University, Tjele, Denmark.
- McJarrow, P., and J. van Amelsfort-Schoonbeek. 2004. Bovine sialyl oligosaccharides: Seasonal variations in their concentrations in milk and a comparison to the colostrums of Jersey and Friesian cows. *Int. Dairy J.* 14:571–579.
- Miglior, F., A. Sewalem, J. Jamrozik, J. Bohmanova, D. M. Lefebvre, and R. K. Moore. 2007. Genetic analysis of milk urea nitrogen and lactose and their relationship with other production traits in Canadian Holstein cattle. *J. Dairy Sci.* 90:2468–2479.
- Mucha, S., and E. Strandberg. 2011. Genetic analysis of milk urea nitrogen and relationships with yield and fertility across lactation. *J. Dairy Sci.* 94:5665–5672.
- Nicholson, G., M. Rantalainen, J. V. Li, A. D. Maher, D. Malmmodin, K. R. Ahmadi, J. H. Faber, A. Barrett, J. L. Min, N. W. Rayner, H. Toft, M. Krestyaninova, J. Viksna, S. G. Neogi, M. E. Dumas, and U. Sarkans., the molPAGE ConsortiumDonnelly, P., T. Illig, J. Adamski, K. Suhre, M. Allen, K. T. Zondervan, T. D. Spector, J. K. Nicholson, J. C. Lindon, D. Baunsgaard, E. Holmes, M. I. McCarthy, and C. C. Holmes. 2011. A genome-wide metabolic QTL analysis in Europeans implicates two loci shaped by recent positive selection. *PLoS Genet.* 7:e1002270.
- Novak, A. F., and S. M. Hauge. 1948. Isolation of the unidentified growth factor (Vitamin B13) in distillers' dried solubles. *J. Biol. Chem.* 174:647–651.
- Oikonomou, G., G. E. Valergakis, G. Arsenos, N. Roubies, and G. Banos. 2008. Genetic profile of body energy and blood metabolic traits across lactation in primiparous Holstein cows. *J. Dairy Sci.* 91:2814–2822.
- Page, M. A., and D. H. Williamson. 1972. Lactating mammary gland of the rat: A potential major site of ketone-body utilization. *Biochem. J.* 128:459–460.
- Pimentel, E. C. G., M. Erbe, S. König, and H. Simianer. 2011. Genome partitioning of genetic variation for milk production and composition traits in Holstein cattle. *Front. Genet.* 2:19.
- Reinecke, D., H. Burhenne, P. Sandner, V. Kaever, and R. Seifert. 2011. Human cyclic nucleotide phosphodiesterases possess a much broader substrate-specificity than previously appreciated. *FEBS Lett.* 585:3259–3262.
- Robinson, J. L., D. B. Dombrowski, J. H. Clark, and R. D. Shanks. 1984. Orotate in milk and urine of dairy cows with a partial deficiency of uridine monophosphate synthase. *J. Dairy Sci.* 67:1024–1029.
- Shennan, D. B., and M. Peaker. 2000. Transport of milk constituents by the mammary gland. *Physiol. Rev.* 80:925–951.
- Stoop, W. M., H. Bovenhuis, and J. A. M. van Arendonk. 2007. Genetic parameters for milk urea nitrogen in relation to milk production traits. *J. Dairy Sci.* 90:1981–1986.
- Stoop, W. M., J. A. M. van Arendonk, J. M. L. Heck, H. J. F. van Valenberg, and H. Bovenhuis. 2008. Genetic parameters for major milk fatty acids and milk production traits of Dutch Holstein-Friesians. *J. Dairy Sci.* 91:385–394.
- Sundekilde, U. K., P. D. Frederiksen, M. R. Clausen, L. B. Larsen, and H. C. Bertram. 2011. Relationship between the metabolite profile and technological properties of bovine milk from two dairy breeds elucidated by NMR-based metabolomics. *J. Agric. Food Chem.* 59:7360–7367.
- Sundekilde, U. K., N. Poulsen, L. B. Larsen, and H. C. Bertram. 2013. Nuclear magnetic resonance metabolomics reveals strong association between milk metabolites and somatic cell count in bovine milk. *J. Dairy Sci.* 96:291–299.
- Teo, Y. Y., M. Inouye, K. S. Small, R. Gwilliam, P. Deloukas, D. P. Kwiatkowski, and T. G. Clark. 2007. A genotype calling algorithm for the Illumina BeadArray platform. *Bioinformatics* 23:2741–2746.

- Udabage, P., I. R. McKinnon, and M. A. Augustin. 2001. Effects of mineral salts and calcium chelating agents on the gelation of renneted skim milk. *J. Dairy Sci.* 84:1569–1575.
- Ulrich, E. L., H. Akutsu, J. F. Doreleijers, Y. Harano, Y. E. Ioannidis, J. Lin, M. Livny, S. Mading, D. Maziuk, Z. Miller, E. Nakatani, C. F. Schulte, D. E. Tolmie, R. K. Wenger, H. Yao, and J. L. Markley. 2008. BioMagResBank. *Nucleic Acids Res.* 36(Suppl. 1):D402–D408.
- Van Tassell, C. P., T. P. L. Smith, L. K. Matukumalli, J. F. Taylor, R. D. Schnabel, C. T. Lawley, C. D. Haudenschild, S. S. Moore, W. C. Warren, and T. S. Sonstegard. 2008. SNP discovery and allele frequency estimation by deep sequencing of reduced representation libraries. *Nat. Methods* 5:247–252.
- VanRaden, P. M. 2008. Efficient methods to compute genomic predictions. *J. Dairy Sci.* 91:4414–4423.
- Wishart, D. S., C. Knox, A. C. Guo, R. Eisner, N. Young, B. Gautam, D. D. Hau, H. Psychogios, E. Dong, S. Bouatra, R. Mandal, I. Sinelnikov, J. Xia, L. Jia, J. A. Cruz, E. Lim, C. A. Sobsey, S. Shrivastava, P. Huang, P. Liu, L. Fang, J. Peng, R. Fradette, D. Cheng, D. Tzur, M. Clements, A. Lewis, A. De Souza, A. Zuni-
ga, M. Dawe, Y. Xiong, D. Clive, R. Greiner, A. Nazyrova, R. Shaykhutdinov, L. Li, H. J. Vogel, and I. Forsythe. 2009. HMDB: A knowledgebase for the human metabolome. *Nucleic Acids Res.* 37(Suppl. 1):D603–D610.
- Wood, G. M., P. J. Boettcher, D. F. Kelton, and G. B. Jansen. 2004. Phenotypic and genetic influences on test-day measures of acetone concentration in milk. *J. Dairy Sci.* 87:1108–1114.
- Yang, J., B. Benyamin, B. P. McEvoy, S. Gordon, A. K. Henders, D. R. Nyholt, P. A. Madden, A. C. Heath, N. G. Martin, G. W. Montgomery, M. E. Goddard, and P. M. Visscher. 2010. Common SNPs explain a large proportion of the heritability for human height. *Nat. Genet.* 42:565–569.
- Yip, M. C., and W. E. Knox. 1972. Function of arginase in lactating mammary gland. *Biochem. J.* 127:893–899.
- Zeisel, S. H. 2006. Choline: Critical role during fetal development and dietary requirements in adults. *Annu. Rev. Nutr.* 26:229–250.
- Zeisel, S. H., M. H. Mar, J. C. Howe, and J. M. Holden. 2003. Concentrations of choline-containing compounds and betaine in common foods. *J. Nutr.* 133:1302–1307.