Bovine chromosomal regions affecting rheological traits in acid-induced skim milk gels

M. Glantz,* F. Gustavsson,* H. P. Bertelsen,† H. Stålhammar,‡ H. Lindmark-Månsson,*§ M. Paulsson,* C. Bendixen,† and V. R. Gregersen†¹

*Department of Food Technology, Engineering and Nutrition, Lund University, PO Box 124, SE-221 00 Lund, Sweden

†Department of Molecular Biology and Genetics, Faculty of Science and Technology, Aarhus University, Blichers Allé 20, PO Box 50, DK-8830 Tjele, Denmark

‡VikingGenetics, PO Box 64, SE-532 21 Skara, Sweden

§Lantbrukarnas Riksförbund Dairy Sweden, Ideon Science Park, SE-223 70 Lund, Sweden

ABSTRACT

The production of fermented milk products has increased worldwide during the last decade and is expected to continue to increase during the coming decade. The quality of these products may be optimized through breeding practices; however, the relations between cow genetics and technological properties of acid milk gels are not fully known. Therefore, the aim of this study was to identify chromosomal regions affecting acid-induced coagulation properties and possible candidate genes. Skim milk samples from 377 Swedish Red cows were rheologically analyzed for acid-induced coagulation properties using low-amplitude oscillation measurements. The resulting traits, including gel strength, coagulation time, and yield stress, were used to conduct a genome-wide association study. Single nucleotide polymorphisms (SNP) were identified using the BovineHD SNPChip (Illumina Inc., San Diego, CA), resulting in almost 621,000 segregating markers. The genome was scanned for putative quantitative trait loci (QTL) regions, haplotypes based on highly associated SNP were inferred, and the additive genetic effects of haplotypes within each QTL region were analyzed using mixed models. A total of 8 genomic regions were identified, with large effects of the significant haplotype explaining between 4.8 and 9.8% of the phenotypic variance of the studied traits. One major QTL was identified to overlap between gel strength and yield stress, the QTL identified with the most significant SNP closest to the gene coding for κ -casein (*CSN3*). In addition, a chromosome-wide significant region affecting yield stress on BTA 11 was identified to be colocated with *PAEP*, coding for β -lactoglobulin. Furthermore, the coagulation properties of the genetic

variants within the 2 genes were compared with the coagulation properties identified by the patterns of the haplotypes within the regions, and it was discovered that the haplotypes were more diverse and in one case slightly better at explaining the phenotypic variance. Besides these significant QTL comprising the 2 milk proteins, 3 additional genes are proposed as possible candidates, namely *RAB22A*, *CDH13*, and *STAT1*, and all have previously been found to be expressed in the mammary gland. To our knowledge, this is the first attempt to map QTL regions for acid-induced coagulation properties.

Key words: acid-induced coagulation, genome-wide association study, milk protein, dairy cattle

INTRODUCTION

The production of fermented milk products, such as yogurt, has increased worldwide during the last decade and will most likely increase within the coming decade as well (International Dairy Federation, 2012). Today, more than 25 million tons of fermented products are produced every year throughout the world (International Dairy Federation, 2012), thus making it an economically vital part of the dairy industry. Acidification of milk is the basis in fermented milk production, where a lowering in pH causes insoluble calcium phosphate to dissolve within the CN micelles, and the net negative charge of the CN micelles is neutralized (Lucey, 2009). This results in aggregation as the isoelectric point of the CN micelles ($\approx$ pH 4.6) is reached (Lucey, 2009). During fermented milk production, the milk is normally subjected to extensive heat treatment ($\approx$ 90–95°C, 5–10 min), where whey proteins, mainly β -LG, are denatured (Walstra et al., 2006). The denatured β -LG will, via hydrophobic interactions and disulfide bonds, form either complexes with κ -CN on the surface of the CN micelles or soluble aggregates with themselves, and thereby form a protein gel network (Donato and Guyomarc'h, 2009) with viscoelastic properties (Robinson and Itsa-

Received March 14, 2014.

Accepted November 15, 2014.

¹Corresponding author: Vivi.Gregersen@agrsci.dk

ranuwat, 2006). Studies have shown that heat treatment of milk results in higher gel strength and thereby improved viscoelastic properties of acid gels compared with unheated milk (Donato and Guyomarc'h, 2009; Lucey, 2009).

Several factors have been shown to affect acid milk gels, including milk composition, processing conditions (Donato and Guyomarc'h, 2009; Lucey, 2009), as well as genetic factors (Jakob and Puhan, 1992; Allmere et al., 1998a; Hallén et al., 2009). Despite this, the relations between cow genetics and technological properties of acid milk gels are not fully known. Studies have so far focused on genetic polymorphism of milk proteins; however, the results are somewhat contradictory, mainly regarding the effects of the β -LG variants A and B on acid gel properties (Allmere et al., 1998a; Manderson et al., 1998; Hallén et al., 2009). No significant effects of β -CN (A1, A2, A3, or B; Hallén et al., 2009) or κ -CN [A or B (Allmere et al., 1998a); A, B or E (Hallén et al., 2009)] have been found on acid-induced gel strength or coagulation time (CT). On the contrary, α_{s1} -CN BB, κ -CN AB, and β -LG BB have been reported to be associated with better consistency and viscosity in fermented milk products (Jakob and Puhan, 1992). These contradictory results suggest that other genes may be of importance for acid milk gels; however, to our knowledge no studies have yet been conducted to explore this.

During the past years, several whole-genome scans and genome-wide association studies (GWAS) have been conducted for milk composition traits (e.g., Schopen et al., 2009; Gamba et al., 2013), but such studies are scarce for technological properties of milk. Thus far, only chromosomal regions affecting rennet-induced coagulation important for cheese production have been studied (Gregersen et al., 2015). In the study on rennet-induced coagulation, a total of 21 genomic regions were found for 4 different rennet-induced coagulation traits, and major QTL were identified around the CN gene (*CSN*) cluster on BTA 6. In addition, a candidate gene was identified in relation to O-glycosylation of κ -CN as well as 2 genes in relation to proteolysis of milk proteins. This suggests that chromosomal regions in addition to the regions harboring the milk-protein genes may also be of importance for acid-induced coagulation. The aim of this study was therefore to identify QTL within the bovine genome affecting 3 acid-induced coagulation traits important for viscosity and texture of acidified products, namely gel strength, CT, and yield stress (YS). To our knowledge, this is the first study to identify genomic regions related to genes or genomic regions that affect the acidification of milk with the potential to be used in breeding.

MATERIALS AND METHODS

Milk and Blood Samples

As part of the Danish–Swedish Milk Genomics Initiative, 377 Swedish Red (SR) cows from 21 herds were sampled for morning milk samples and blood samples during the indoor period between April to May 2010 and September 2010 to April 2011. Each herd was sampled in the same day, and a maximum of 24 individual milk samples were collected from each herd. The herds were conventional farms located within the same geographical area in the southern part of Sweden. The cows descended from in total 159 sires and 76 paternal grandsires within the sample population with the number of daughters per sire ranging from 1 to 17. As described in Gregersen et al. (2015), 26 of the bulls had both daughters and granddaughters among the sampled cows. Lactation number ranged from 1 to 4 (1% in lactation number 4) and lactation week from 7 to 40 (2% before lactation wk 7 and 10% after lactation wk 40) for the cows included in the study. The selection of the cows was designed to have a low degree of genetic relatedness between the cows and phenotypic characteristics that represent the current average SR population. None of the cows included in the study showed indication of impaired udder health; all cows had SCC <300,000 cells/mL with mean SCC 74,000 cells/mL. They were fed according to standard practices and milked 2 times (in total 274 of the sampled cows from 16 of the herds) or 3 times (in total 103 of the sampled cows from 5 of the herds) per day. Representative volumes in proportion to milk yield of each cow were sampled and carefully mixed before an aliquot was collected, cooled, and transported to Lund University, Sweden, the day of sampling. The samples were defatted by centrifugation at $2,000 \times g$ for 30 min at 4°C, subsampled, and stored at –20°C until rheological analyses. Milk yields were recorded for all cows on each sampling occasion. Fresh milk samples were analyzed for pH as well as SCC by using flow cytometry (CombiFoss 5000, Foss, Hillerød, Denmark) at a certified dairy analysis laboratory (Eurofins Steins Laboratory, Jönköping, Sweden).

Rheological Analyses

Rheological properties of acid-induced coagulation in individual skim milk samples were determined using low-amplitude oscillation measurements (Kinexus rheometer, Malvern Instruments Ltd., Worcestershire, UK). The milk samples were thawed at 4°C overnight and prewarmed at 30°C for 30 min, timed from the beginning of warming of refrigerated samples to ob-

tain representative samples. A total of 40 mL of milk was heat treated in a beaker at 95°C for 5 min (10 min in total including heating up) in a water bath and cooled on ice to 30°C. Glucono- δ -lactone (**GDL**; Sigma-Aldrich, St. Louis, MO) was added to the samples to 3% (wt/vol) and thoroughly mixed. The GDL concentration was chosen to have a firm gel within the time of measurement. The solution was immediately poured into the concentric cylinder cell (PC25) with the vane and cup geometry (4V21:PC25) of the rheometer. Low-amplitude oscillation measurements were carried out at 30°C for 60 min with a frequency of 1 Hz and controlled strain at 0.003, which was determined to be in the linear viscoelastic region of the acid milk gels. The addition of GDL corresponded to time 0. Gel strength (G'_{60}) was defined as the elastic or storage modulus, G' , after 60 min of coagulation, where $G' = \sigma_0/\gamma_0 \cdot \cos\delta$, and coagulation time was defined as $G'' = G'$, where $G'' = \sigma_0/\gamma_0 \cdot \sin\delta$ (σ_0 = maximum stress; γ_0 = maximum shear strain; δ = phase angle, phase difference between the strain and stress). After 60 min of coagulation, stress sweeps were run to obtain YS, where the shear stress was increased in 126 intervals from 1 to 300 Pa. Yield stress was defined as the shear stress at the first local maximum of the viscosity. All of the samples coagulated within time of measurement, and single measurements were performed for each cow. Following rheological analyses, pH was measured in the acid milk gels 75 min after GDL addition (mean pH 4.40).

Genotyping and Data Validation

Genotyping of the 377 animals was performed in cooperation with GenoScan A/S, Denmark, as described in Gregersen et al. (2015). The data were further analyzed by the Haploview 4.2 software (Barrett et al., 2005) to remove SNP with low minor allele frequency (<0.02) and deviation from Hardy-Weinberg proportions ($P < 0.001$; Wigginton et al., 2005). The SNP were ordered according to the position given by the manufacturer of the SNP chip, which were assigned to the *Bos taurus* UMD3 genome sequence (Zimin et al., 2009), and tag SNP were identified using Haploview 4.2's implementation of the Tagger algorithm for pairwise tagging using an r limit of 0.80 and a logarithm of the odds (LOD) threshold of 3 on the individual chromosomes (de Bakker et al., 2006). The tag SNP were used to generate a less-conservative genome and chromosome-wide Bonferroni significance level for the association and haplotype analysis as described in Gregersen et al. (2015).

Data Analysis

As outlined in Gregersen et al. (2015), the SNP were analyzed by a 2-step procedure. In the first step, the procedure consisted of a simple scan where the association of each SNP to the traits was evaluated. In the second step, SNP sets were selected and phased, and the corresponding haplotypes were evaluated by a mixed model. For details, see below.

Genome-Wide Association Analysis

In the first step in the analysis, a GWAS was performed as a simple scan to identify SNP highly associated with the rheological traits, i.e., CT, YS, and G'_{60} . The traits CT and YS were log-transformed to fit the Gaussian distribution, whereas the G'_{60} trait was found to be normally distributed without adjustments. The analysis was performed using the “-at” option of analyzing the total evidence of association conducted as a simple F -statistic implemented within the QTDT program package (version 2.5.1; QTDT, 2005) as described in detail in Gregersen et al. (2015). Because this step of the procedure was only performed to identify potentially interesting genomic regions, the 2-generation pedigree was modeled accounting for the population structure identified within the sampled population (Gregersen et al., 2015); however, the nongenetic factors potentially affecting the traits were not included in this statistical model. Note that such nongenetic factors are accounted for in subsequent analysis of the QTL regions, which is presented below.

Phasing Algorithm

In the second step of the procedure haplotypes were inferred by phasing of SNP from the SNP sets as described above and in detail in Gregersen et al. (2015). Briefly, SNP sets within regions highly associated with the traits were selected based on the criterion that the top SNP should be significant at least at the 5% chromosome-wide Bonferroni corrected level, and subsequent SNP should not be below the lowest level of chromosome-wide significance, identified to be $P < 1.1 \times 10^{-5}$ in this study. Furthermore, the distance between selected markers should not exceed 10 Mb. The phasing was conducted using a recursive algorithm constructing a binary tree and going through it twice. In the first run, the tree was built from the genotype combinations of the animals. A left and right branch represented allele 1 and allele 2, respectively. The leaf nodes contain information on the phase and a counter that summed

the probability of the specific phase within the population. In a second run, the most likely phase was selected, and the additional phase was calculated by subtracting the first phase from the genotypes. Thereafter, a mixed model was used to test for significance as described below.

Mixed Model

The most common haplotype within a given SNP set region was identified, and the effect of this haplotype (versus all other haplotypes) was analyzed to estimate the additive effect and the phenotypic variance explained by the haplotype. Animals with missing genotypes were placed within the “all other haplotypes” category. The lme4 package of R (Bates et al., 2008) was used for the purpose.

The additive haplotype effects were estimated using the following mixed model:

$$Y_i = \mu + H_{\text{herd}(i)} + w_{\text{week}(i)} + l_{\text{lactation}(i)} + \beta h(i) + e_i, \quad [1]$$

where Y_i is the trait value for the individual i , μ is the intercept, $H_{\text{herd}(i)}$ is the random effect of herd/test day (21 farms), and e_i is the random effect of the residual variance, for individual i , respectively. All random effects were assumed to be normally distributed with independent covariance structure. In addition, $w_{\text{week}(i)}$ and $l_{\text{lactation}(i)}$, the every-other-month week of lactations (5 intervals) and lactation number [1 or 2 (or more)], respectively, were considered fixed effects. The haplotype effect was a regression on $h(i)$, where $h(i) = 1$ if there were 2 copies of the most common haplotype, 0 if there was 1 copy, and -1 if there were 0 copies, and the parameter β was the additive genetic effect of the most common haplotype. A significance test for $\beta = 0$, i.e., no haplotype effect, was performed. No sire or animal effect was considered, as discussed in Gregersen et al. (2015). If the most common haplotype did not show any significance, the second-most-common haplotype was evaluated as well. Furthermore, if the mean value of the trait for a given haplotype was found to be extreme (lower or higher) compared with the rest of the mean values for the other haplotypes, this haplotype was evaluated by the model as described for the most common haplotype. Because 2 of the milk-protein genes were identified as likely candidates, the haplotypes within these regions were all analyzed for effects on the trait. This was done to compare the haplotypes to the variants of the genes as later described.

The phenotypic variance explained by the significant haplotype was estimated using the output from the

ANOVA statistics. The total variance was computed as the sum of the herd, residual, and haplotype variances for the model, where the variance of the haplotype effect (Var_{hap1}) was estimated according to Equation 2:

$$\text{Var}_{\text{hap1}} = 2 \times \text{freq}_{\text{hap1}} \times (1 - \text{freq}_{\text{hap1}}) \times (\text{effect}_{\text{hap1}})^2, \quad [2]$$

where $\text{freq}_{\text{hap1}}$ and $\text{effect}_{\text{hap1}}$ were the frequency and the estimated additive genetic effect of the most common haplotype, respectively. The variance of the haplotype effect was then divided by the total variance and multiplied by 100 to get the value in percentage.

The P -values obtained from the mixed model analysis were adjusted for multiple testing by the Bonferroni correction as described above. We decided to use the same level of adjustment as for the simple GWAS because we used the SNP for identifying haplotypes. In addition, an in-house comparison between this approach and traditional GWAS analysis had revealed that we were able to avoid too many false positives by selecting this strict level of correction (data not shown).

Genetic Variants of CSN3 and PAEP

To identify whether protein genetic variants could be related to the identified haplotypes, genetic variants of *CSN3* and *PAEP* were determined. Genomic DNA was purified from blood using the BioSprint 96 workstation (QIAGEN). Samples were treated with protease followed by isopropanol precipitation, and DNA was purified by means of magnetic-particle technology. All samples were genotyped for major genetic variants of *CSN3* and *PAEP* using custom Taqman SNP genotyping assays according to protocol (Applied Biosystems, Foster City, CA). Four assays have previously been constructed to distinguish between the genetic *CSN3* variants A, B, and E (Caroli et al., 2009), as described in a previous study by Poulsen et al. (2013). Furthermore, 4 assays (Supplemental Table S1; <http://dx.doi.org/10.3168/jds.2014-8137>) were designed to distinguish between the A and B genetic variant of *PAEP* (Caroli et al., 2009) using the same procedure as described in Poulsen et al. (2013).

Statistical Analyses

Box-and-whisker plots were used to report distributions of coagulation properties in relation to protein genetic variants and haplotypes identified. The box-and-whisker plots included the median and mean values, upper (75%) and lower (25%) quartiles, as well as minimum and maximum values of the data. The

Table 1. Descriptive statistics for rheological traits of acid-induced skim milk gels of Swedish Red cows (n = 377)

Trait ¹	Mean	SD	Minimum	Maximum	CV (%)
G' ₆₀ (Pa)	221.6	62.4	38.2	489.2	28
CT (min)	18.1	3.6	10.3	41.4	20
YS (min)	28.2	10.5	5.7	66.9	37
log(YS)	1.4	0.2	0.8	1.9	—
log(CT)	1.3	0.1	1.0	1.6	—

¹G'₆₀ = gel strength after 60 min of coagulation; CT = coagulation time; YS = yield stress.

general linear model procedure from Minitab (version 14, Minitab Ltd., Coventry, UK) was used for pairwise testing of significant differences between protein genetic variants or haplotypes. The models included the fixed effects of *CSN3* and *PAEP* genetic variants as well as haplotypes identified, respectively. The level of significance was set at $P < 0.05$.

RESULTS AND DISCUSSION

This study reports chromosomal regions affecting acid-induced coagulation properties and possible candidate genes. This is the first time a mapping approach has been performed for acid-induced coagulation properties, thus adding new knowledge to the dairy and breeding research community. The present study is based in a limited number of cows (377 cows); however, the results presented here give an indication of results to expect in a larger sample population.

Acid Milk Coagulation

Mean values, standard deviations, and coefficients of variation for the analyzed acid-induced coagulation traits are presented in Table 1. The results indicated a large variation within the coagulation traits between individual milk samples, which was also reflected in the coefficients of variation observed for the rheological traits. For G'₆₀ and YS the coefficients of variation were high, 28 and 37%, respectively, whereas a moderate coefficient of variation was found for CT (20%). No studies have, to our knowledge, earlier reported values of coefficients of variation for acid-induced coagulation traits, but the values are in the range previously reported for rennet-induced coagulation traits (Cassandro et al., 2008; Glantz et al., 2012a). Moderate to high heritabilities have been shown for acid-induced G'₆₀ and YS ($h^2 = 0.62$ and 0.38 , respectively), whereas the heritability of CT in acid-induced coagulation was estimated to be 0.0 (Gustavsson et al., 2014). The study by Gustavsson et al. (2014) thus indicates a possibility of changing acid-induced G'₆₀ and YS by genetic selection. Average values of fat, total protein, CN, and lactose contents

were 4.1, 3.6, 2.7, and 4.7%, respectively, within the sample population (Gregersen et al., 2015), which is in the range of previously reported values for SR cows (Lundén et al., 1997; Glantz et al., 2012b).

In the production of fermented milk products, short CT as well as high gel strength and YS are desirable for an optimal gel. Industrially, milk is acidified with bacterial cultures, whereas the acidification in this study was simulated by GDL. Studies have shown that GDL-induced gels have shorter CT (van Marle and Zoon, 1995; Lucey et al., 1998) and higher gel strengths and YS compared with gels formed using bacterial cultures (Lucey et al., 1998). The chemical acidification with GDL was used to decrease the number of variables influencing acid-induced coagulation, i.e., to avoid variation in activity of cultures and differences between strains that could have a large effect on gel formation (Lucey et al., 1998; Robinson and Itsaranuwat, 2006; Lucey, 2009). In the present study, frozen milk samples were used for rheological analyses of acid-induced coagulation, which may influence gel formation. However, the samples were thawed at 4°C and prewarmed at 30°C before analyses to obtain an optimal equilibrium. The temperatures were carefully chosen to have the smallest effect on the structure of milk and hence the CN micelles.

Identified QTL Regions

The 377 animals used in this study were genotyped in a previous study (Gregersen et al., 2015), and milk samples from all animals were analyzed rheologically for acid-induced coagulation traits. The 3 resulting traits (CT, G'₆₀, and YS) were used in combination with the genotype data set to conduct a GWAS of the entire bovine genome.

In total, 9 of the 16 suggestive QTL representing 8 genomic regions on 6 different chromosomes were significant by the haplotype analysis (Table 2; Supplemental Table S2; <http://dx.doi.org/10.3168/jds.2014-8137>). The only overlapping region was identified in relation to G'₆₀ and YS on BTA 6, and in both cases with genome-wide significant P -values indicated a major lo-

cus. In general, the QTL intervals were broader for G'_{60} than for YS, but the significant signals accounted for approximately the same range of phenotypic variance explained. Furthermore, these findings are in agreement with the moderate to high heritabilities found for the traits (Gustavsson et al., 2014). One genomic region was identified associated with CT even though, as earlier mentioned, no heritability was observed for this trait (Gustavsson et al., 2014). However, as no confidence interval was reported for the heritability of the trait, the statistical uncertainty of the estimate is unknown. Moreover, a few additional samples were used in the present study indicating that the trait might be slightly heritable. For the other 2 traits, 5 of the suggestive QTL affecting YS located on BTA 1, 6a, 6b, 11, 17a, and 17c, and 3 QTL affecting G'_{60} on BTA 6, 13, and 18 were significant in the haplotype analysis. All regions were identified to be affected by the most common haplotype except for the region on BTA 17a that was identified to be affected by the fourth-most-frequent haplotype.

The fixed effect of the significant haplotypes was generally high, and the haplotypes were identified to

explain between 4.8 and 9.8% of the variance (Table 2). Nevertheless, because the mixed-model approach only represents the samples within the data set and calculates the variation among these, both the fixed effects and phenotypic variance explained are likely to be overestimated. In addition, several regions were close to significant after the rather conservative adjustment for multiple testing, indicating that these should be considered potential QTL regions. Some of the regions might also contain interesting candidate genes, such as the region affecting G'_{60} on BTA 28. Here *H2AFY2* is situated, which codes for H2A histone family, member Y2, which is part of the lactation gene set (Lemay et al., 2009). The gene is involved in chromatin modification affecting gene transcription. Previously, epigenetic changes to the chromatin structure surrounding the CSN cluster have been shown to unlock the genes for transcription during lactation (Rijnkels et al., 2013).

The overlapping QTL on BTA 6 affecting both G'_{60} and YS were found to be colocated with a QTL identified in a study by Schopen et al. (2011) affecting protein content, CN index, as well as the relative concentrations of κ -CN, β -LG, and α -LA in the Holstein-Friesian breed.

Table 2. Suggestive QTL regions identified in the Swedish Red breed for rheological parameters in acid-induced skim milk gels in addition to a haplotype analysis for an additive effect of the most frequent haplotype within the QTL¹

BTA	Top SNP (Mb)	QTL interval (Mb)	No. ²	HF ³ (%)	Haplotype analysis			Candidate gene	
					P-value	Effect	%Var ⁴	Symbol	Location (Mb)
Log(coagulation time) (min)									
2	79.067	79.067–79.152	3	61.3	2.38E-06*	−0.022	5.5	STAT1	79.894–79.926
13	81.071	81.069–81.078	5	47.6	3.22E-04	−0.018	3.6		
Gel strength (Pa)									
2	12.825	12.825–16.212	9	37.7	5.29E-06	17.54	4.1	CSN3	87.378–87.392
4	36.805	35.154–44.194	16	19.8	4.60E-03	16.78	2.5		
6	83.566	83.566–87.420	4	55.8	1.03E-06*	−21.44	6.3		
				11.2 ⁵	1.43E-07**	36.19	7.3		
13	62.619	53.137–62.974	12	20.9	5.65E-07*	22.53	4.8	RAB22A	58.537–58.585
18	9.847	8.998–11.675	10	41	1.53E-10**	26.92	9.8	CDH13	9.189–10.164
28	26.388	26.370–26.404	4	49.1	2.17E-05	−17.09	4.1		
Log(Yield stress) (Pa)									
1	9.940	9.262–9.997	6	36.2	4.63E-05	−0.05	4.5	CSN3	87.378–87.392
6a	70.699	70.263–71.442	8	43.3	4.35E-09**	−0.056	6.1		
6b	87.420	86.527–87.421	7	55.8	5.69E-11**	−0.072	9.8		
				12.4 ⁵	5.27E-08**	0.092	7.2		
11	103.302	103.286–103.317	4	51	5.25E-07*	0.055	5.9	PAEP	103.301–103.306
				32.5 ⁶	4.62E-07*	−0.061	6.4		
17a	3.286	2.847–3.286	4	2.8 ⁷	1.50E-06*	−0.169	6		
17b	54.337	53.886–54.347	4	70.8	1.59E-04	−0.046	3.4		
17c	71.43	71.101–71.799	7	52.2	4.56E-06*	0.054	5.6		
29	4.948	4.867–4.967	5	46.0	4.41E-04	0.041	3.2		

¹Candidate genes were selected based on Lemay et al. (2009).

²Number of identified haplotypes with a frequency above 1.5%.

³HF = haplotype frequency of the analyzed haplotype, the most common haplotype if not otherwise indicated.

⁴%Var = phenotypic variance explained by the haplotype.

⁵Analysis of haplotype 3, the third-most-frequent haplotype.

⁶Analysis of haplotype 2, the second-most-frequent haplotype.

⁷Analysis of haplotype 4, the fourth-most-frequent haplotype.

*, **Significant at the 5% chromosome-wide and 5% genome-wide level, respectively.

The same overlap is valid for several SNP significantly associated with total protein content and concentrations of α_s -CN, β -CN, κ -CN, and β -LG identified for purebred Holstein and (Holstein $\times$ Jersey) $\times$ Holstein crossbred animals (Huang et al., 2012). Both the CN content (Walstra et al., 2006) and concentrations of β -LG and α -LA (Hallén et al., 2009) have earlier been reported to be associated with acid gel properties, thus strengthening the effect of milk protein composition on acid-induced coagulation and confirming the results found in the present study. Furthermore, Kolbehdari et al. (2009) found a genome-wide significant SNP on BTA 6 for fat content in Canadian Holstein bulls, which is in the same location found for YS on BTA 6 in the present study. As for protein composition, also fat content affects acid gel properties (Walstra et al., 2006; Lucey, 2009). A chromosome-wide significant QTL on BTA 11 (103.286 to 103.317 Mb) affecting YS in the present study is located in the same region as several SNP affecting β -LG concentration (Schopen et al., 2011; Huang et al., 2012; Gamba et al., 2013). A higher β -LG concentration has earlier been reported to be associated with shorter acid-induced CT and higher gel strength (Hallén et al., 2009), thus once again strengthening the results found in the present study. In addition, Schopen et al. (2011) also found several SNP within this QTL region that affect the concentration of α_s -CN, β -CN, κ -CN, as well as the CN index.

Major QTL on Chromosome 6

The major QTL on BTA 6 identified here is located within the same region as the major QTL identified for rennet-induced gels in Gregersen et al. (2015), but in the present study, only G'_{60} and YS were identified to be affected. This indicates that the CSN are equally important in both rennet- and acid-induced coagulation, which was expected. In the present study, the same candidate gene as identified in Gregersen et al. (2015), namely *CSN3* coding for κ -CN, was identified to affect both YS and G'_{60} as the only CSN gene with strong associations to the traits. Therefore, this gene was selected for further analysis to test whether one of the genetic variants (A, B, and E) of *CSN3* could be related to the identified haplotypes. Figures 1 and 2 represent boxplots of both the genetic variants and haplotypes within the gene in relation to YS and G'_{60} , respectively. In addition, a schematic representation of how the haplotypes relate to the variants of the gene is presented in Tables 3 and 4. From the figures it is evident that the B variant of the gene is the most beneficial for both traits, which is shown by the higher mean values of G'_{60} and YS for the BB, BE, and AB genotypes. Furthermore, by comparing the haplotypes

in the figures (Figures 1 and 2) to the genetic variants represented in the tables (Tables 3 and 4), it can be seen that for YS, the B variant is only represented by haplotype 3 (Table 3), whereas for G'_{60} , the B variant is represented by haplotype 3 and 4 (Table 4). Please note that the haplotypes are not comparable across traits because of differences in SNP sets (Supplemental Table S2; <http://dx.doi.org/10.3168/jds.2014-8137>) as well as differences in the phases (Tables 3 and 4). This should especially be noted when looking at Table 2, where it is shown that the haplotype analysis identified haplotype 1, the most common haplotype, in addition to haplotype 3 to be significant for YS and G'_{60} (Table 2). The beneficial effect of *CSN3* genotype AB has previously been found to affect both consistency and viscosity in fermented products, as reviewed by Jakob and Puhan (1992). However, other studies have found no effect of *CSN3* genetic variants on acid-induced gels (Allmere et al., 1998a; Hallén et al., 2009).

The haplotype analysis also indicates that the genetic variant *CSN3* A can be divided into subgroups such that the group also containing the E variant is associated with low G'_{60} and YS (Tables 3 and 4; split in 2 haplotypes in the analysis of YS). In both analyses, this haplotype has a significant negative additive effect on the traits. That the *CSN3* A variant can be split into subgroups indicates that something else than the genetic variants of *CSN3* contribute to the variation, such as differences in gene expression levels or differences in level of breakdown of the untranslated mRNA. However, it could also be due to linkage to the other CSN genes, but this is unlikely because none of the SNP located in close proximity of these other genes were found to be significantly associated with the traits. Because the most frequent haplotype comprises the most poorly coagulating samples, efforts to try to identify the underlying mechanisms behind this should be of high priority in the near future.

β -LG on BTA 11

Another interesting QTL was identified on BTA 11 to affect YS. This QTL region harbors *PAEP*, a gene known to play an important role in the formation of acid-induced gels (Lucey, 2009). Previously, the A and B genetic variants of *PAEP* have been identified in the SR breed (Lundén et al., 1997; Hallén et al., 2008; Glantz et al., 2012b), and hence this subpopulation was screened for these 2 variants. In Figure 3, a boxplot of both the genetic variants and haplotypes related to the genomic region are represented. This figure indicates that haplotype 1 and the *PAEP* B genetic variant are somewhat similar. But by looking in Table 5, it can be found that the *PAEP* B variant is represented by 2 dif-

Figure 1. Box-and-whisker plots representing the distribution of yield stress in relation to (A) the genetic variants and (B)–(D) the haplotypes (hap) identified within the QTL regions of *CSN3*. Individual haplotypes are represented if the frequency was above 5% and if differences were observed between the tested haplotype combinations. n = number of samples per group; a to c = mean values within each subplot with different letters differ ($P < 0.05$).

ferent haplotypes, likewise the A genetic variant. The table also indicates that the more frequent haplotype 1 relates to the genetic *PAEP* B variant and it has a higher mean value than the other 3 haplotypes that show approximately the same effects on YS. In addition, the haplotype was found to explain about 5.6% of the phenotypic variance of the trait. This could indicate that haplotype 1 is better at explaining the difference in phenotype than the genetic B variant, and this was confirmed by performing the mixed-model analysis for the genetic *PAEP* B variant, which identified the variant to explain 5.2% of the phenotypic variation. Furthermore, an analysis of haplotype 2 using the mixed model also revealed that this haplotype had a significant effect on the trait explaining about 6.4% of

the phenotypic variance. Previous studies have shown somewhat contradictory results regarding the effect of *PAEP* genetic variants on acid-induced coagulation properties. Hallén et al. (2009) found a favorable effect of AA and AB genotypes of *PAEP* on gel strength compared with β -LG BB, as well as a favorable effect of β -LG AA on coagulation time compared with β -LG AB and BB, possibly due to an increased concentration of β -LG in milk that was shown to be superior for β -LG A compared with the B variant (Allmere et al., 1998b; Hallén et al., 2008). In contrast, when adjusting for protein concentration, both Hallén et al. (2009) and Allmere et al. (1998a) found an association between β -LG B and higher gel strength compared with the A variant. This was also confirmed by Bikker et al. (2000)

Table 3. Descriptive statistics of the haplotypes identified within the QTL (chromosome 6: 83.566–87.420 Mb) around *CSN3* affecting yield stress¹

Haplotype	Phase ²	Frequency (%)	Genetic variant	Mean yield stress $\pm$ SD (Pa)
Hap1	121212211111111111	55.79	A and E	25.8 $\pm$ 9.0
Hap2	211221211111111122	18.59	A	30.7 $\pm$ 10.7
Hap3	212121122222222222	12.40	B	34.6 $\pm$ 12.3
Hap4	211212211111111111	3.44	A and E	27.2 $\pm$ 10.9
Hap5	121221211111111122	1.79	A	27.3 $\pm$ 8.3
Hap6	211211211111111112	1.65	A	28.3 $\pm$ 6.5
Hap7	121221211111111112	1.51	A	26.3 $\pm$ 8.8

¹The table presents the haplotype phase, frequency, the distribution of the genetic variants of the gene, and mean $\pm$ SD of the trait. Haplotypes with a frequency above 1.5% are presented.

²A list of SNP can be found in Supplemental Table S2 (<http://dx.doi.org/10.3168/jds.2014-8137>).

Figure 2. Box-and-whisker plots representing the distribution of gel strength in relation to (A) the genetic variants and (B) and (C) the haplotypes (hap) identified within the QTL regions of *CSN3*. Individual haplotypes are represented if the frequency was above 5% and if differences were observed between the tested haplotype combinations. n = number of samples per group; a to d = mean values within each subplot with different letters differ ($P < 0.05$); G'_{60} = gel strength after 60 min of coagulation.

Table 4. Descriptive statistics of the haplotypes identified within the QTL (chromosome 6: 86.537–87.421 Mb) located around *CSN3* affecting gel strength after 60 min of coagulation (G'_{60})¹

Haplotype	Phase ²	Frequency (%)	Genetic variant	Mean $G'_{60} \pm$ SD (Pa)
Hap1	11221111111111	64.07	A and E	212.5 $\pm$ 59.8
Hap2	11221111111112	18.57	A	227.2 $\pm$ 61.1
Hap3	12112222222222	11.20	B	257.3 $\pm$ 68.0
Hap4	22112222222222	5.05	B	227.6 $\pm$ 62.1

¹The table presents the haplotype phase, frequency, the distribution of the genetic variants of the gene, and mean $\pm$ SD of the trait. Haplotypes with a frequency above 1.5% are presented.
²A list of SNP can be found in Supplemental Table S2 (<http://dx.doi.org/10.3168/jds.2014-8137>).

when studying the addition of pure variants of β -LG to skim milk. Furthermore, β -LG B was shown to be associated with higher aggregation rates toward κ -CN during heating compared with the A variant (Allmere

et al., 1998b), and a lower heat stability has been shown for β -LG B compared with β -LG A (Dannenberg and Kessler, 1988; Jakob and Puhane, 1992; Manderson et al., 1998), which thus favor the aggregation. In the pres-

Figure 3. Box-and-whisker plots representing the distribution of yield stress in relation to (A) the genetic variants and (B) and (C) the haplotypes (hap) identified within the QTL regions of *PAEP*. Individual haplotypes are represented if the frequency was above 5% and if differences were observed between the tested haplotype combinations. n = number of samples per group; a and b = mean values within each subplot with different letters differ ($P < 0.05$).

Table 5. Descriptive statistics of the haplotypes identified within the QTL (chromosome 11: 103.302–103.317 Mb) located around *PAEP* affecting yield stress¹

Haplotype	Phase ²	Frequency (%)	Genetic variant	Mean yield stress $\pm$ SD (Pa)
Hap1	12211	50.96	B	30.0 $\pm$ 10.9
Hap2	21122	32.51	A	25.5 $\pm$ 9.3
Hap3	11122	11.61	A	25.8 $\pm$ 7.3
Hap4	22211	3.55	B	25.7 $\pm$ 9.0

¹The table presents the haplotype phase, frequency, the distribution of the genetic variants of the gene, and mean $\pm$ SD of the trait. Haplotypes with a frequency above 1.5% are presented.

²A list of SNP can be found in Supplemental Table S2 (<http://dx.doi.org/10.3168/jds.2014-8137>).

ent study, β -LG B is associated with higher YS (Table 5), but no associations were found to gel strength or coagulation time.

In a study by Ganai et al. (2009), SNP within the *PAEP* gene were identified in relation to β -LG concentration. In that study, the A variant of *PAEP* was shown to have a significant negative effect on relative concentration of β -LG, whereas other studies have shown a negative effect of the B variant on β -LG concentration (Allmere et al., 1998b; Hallén et al., 2008). Ganai et al. (2009) could not identify a SNP for the B allele explaining a difference in concentration of β -LG. This might indicate that some other differences, e.g., the amino acid substitutions, influence acid gel formation. Furthermore, 2 SNP (BovineHD1100030062 and BovineHD1100030069) identified in this study, located within the promoter and intron 3 regions of *PAEP*, respectively, overlap with the study by Ganai et al. (2009). In the present study an additional SNP (BovineHD1100030066) affecting YS was identified, located in intron 1 in *PAEP*, which was not identified by Ganai et al. (2009). Whether the identified effect of the B allele of *PAEP* is caused by a difference in concentration or by a mutation within the coding part should be further investigated.

Additional Candidate Genes

In addition to the most obvious candidate genes already described, we can suggest additional candidate genes for some of the regions identified within this study. For instance, *RAB22A* is found in close proximity to the highly significant QTL affecting G'_{60} on BTA 13. Protein RAB22A is a membrane protein involved in trafficking and possibly the sorting of transferrin to recycling endosomes. Because lactoferrin is a member of the transferrin family, *RAB22A* could be a potential candidate gene. Furthermore, the protein has been identified in bovine milk fat globule membrane (Reinhardt and Lippolis, 2006). In addition, a potential candidate gene was selected because it was found to be part of the mammary lactation gene set identified by Lemay

et al. (2009). The interesting gene is *STAT1*, located within the genomic region affecting CT on BTA 2. A SNP located within this gene has been identified to affect the regulation of transcription of genes involved in milk protein synthesis and fat metabolism (Cobanoglu et al., 2006). Therefore, *STAT1* has the potential to influence the amount of milk proteins and, hence, is responsible for the QTL within this location. However, this needs to be further investigated. Furthermore, for the QTL affecting G'_{60} on BTA 18, one of the cadherins might be a candidate gene. This QTL is collocated within *CDH13*, called T- or H-cadherin, a gene that is expressed in multiple cell types in the mammary gland (Andrews et al., 2012). It is part of the virgin mammary gene set identified by Lemay et al. (2009). Whether these genes are true candidate genes needs to be further investigated.

CONCLUSIONS

A total of 9 individual genomic regions were identified on 6 different BTA to affect 1 or 2 of the acid-induced rheological traits, CT, G'_{60} , and YS. The significant haplotypes were found to explain between 4.8 and 9.8% of the phenotypic variance. The CSN cluster was harbored by the only overlapping QTL on BTA 6 where *CSN3*, the gene coding for κ -CN, was the only CSN gene significantly associated to the traits. A QTL for YS was also identified in relation to *PAEP*, coding for β -LG, which in addition to κ -CN, is known to affect acid-induced coagulation. A comparison between the genetic variants of *CSN3* and *PAEP* and the identified haplotypes revealed that the genetic variants alone cannot explain the QTL regions, which is in agreement with the contradictory results previously reported. Furthermore, 3 additional candidate genes were suggested, namely *RAB22A*, *CDH13*, and *STAT1*, all found to be expressed in mammary glands. However, these genes need to be verified in future studies. To our knowledge, this is the first attempt to map QTL regions for acid-induced coagulation properties. The results obtained in this study and the contradictory results found in

previous studies thus suggest that little is known about which genes affect acid-induced skim milk coagulation. The genomic regions identified in this study could serve as a basis for the identification of novel candidate genes affecting milk coagulation induced by acidification.

ACKNOWLEDGMENTS

This study is part of the Danish–Swedish Milk Genomics Initiative. The authors gratefully acknowledge TvärLivs [cooperation between the Swedish Research Council for Environment, Agricultural Sciences and Spatial Planning (FORMAS); Swedish Governmental Agency for Innovation Systems (VINNOVA); the Swedish Farmers' Foundation for Agricultural Research (SLF); Livsmedelsföretagen (Li) and Svensk Dagligvaruhandel], Stockholm, Sweden; Arla Foods a.m.b.a., Denmark; SLF, Stockholm, Sweden; the Danish Strategic Research Council, Copenhagen, Denmark; the Danish Cattle Federation, Aarhus, Denmark; and Aarhus University, Denmark, for financial support.

REFERENCES

- Allmere, T., A. Andren, M. Lindersson, and L. Björck. 1998a. Studies on rheological properties of stirred milk gels made from milk with defined genetic variants of κ -casein and β -lactoglobulin. *Int. Dairy J.* 8:899–905.
- Allmere, T., A. Andren, A. Lundén, and L. Björck. 1998b. Interactions in heated skim milk between genetic variants of β -lactoglobulin and κ -casein. *J. Agric. Food Chem.* 46:3004–3008.
- Andrews, J. L., A. C. Kim, and J. R. Hens. 2012. The role and function of cadherins in the mammary gland. *Breast Cancer Res.* 14:203.
- Barrett, J. C., B. Fry, J. Maller, and M. J. Daly. 2005. Haploview: Analysis and visualization of LD and haplotype maps. *Bioinformatics* 21:263–265.
- Bates, D. M., M. Maechler, and B. Dai. 2008. lme4: Linear mixed-effects models using Eigen and R package version 0.999375–28. Accessed Dec. 2, 2014. <http://lme4.r-forge.r-project.org/>.
- Bikker, J. F., S. G. Anema, and J. P. Hill. 2000. Rheological properties of acid gels prepared from heated milk fortified with whey protein mixtures containing the A, B and C variants of β -lactoglobulin. *Int. Dairy J.* 10:723–732.
- Caroli, A. M., S. Chessa, and G. J. Erhardt. 2009. Invited review: Milk protein polymorphisms in cattle: Effect on animal breeding and human nutrition. *J. Dairy Sci.* 92:5335–5352.
- Cassandro, M., A. Comin, M. Ojala, R. Dal Zotto, M. De Marchi, L. Gallo, P. Carnier, and G. Bittante. 2008. Genetic parameters of milk coagulation properties and their relationships with milk yield and quality traits in Italian Holstein cows. *J. Dairy Sci.* 91:371–376.
- Cobanoglu, O., I. Zaitoun, Y. M. Chang, G. E. Shook, and H. Khatib. 2006. Effects of the signal transducer and activator of transcription 1 (*STAT1*) gene on milk production traits in Holstein dairy cattle. *J. Dairy Sci.* 89:4433–4437.
- Dannenberg, F., and H. G. Kessler. 1988. Reaction kinetics of the denaturation of whey proteins in milk. *J. Food Sci.* 53:258–263.
- de Bakker, P. I., N. P. Burt, R. R. Graham, C. Guiducci, R. Yelensky, J. A. Drake, T. Bersaglieri, K. L. Penney, J. Butler, S. Young, R. C. Onofrio, H. N. Lyon, D. O. Stram, C. A. Haiman, M. L. Freedman, X. Zhu, R. Cooper, L. Groop, L. N. Kolonel, B. E. Henderson, M. J. Daly, J. N. Hirschhorn, and D. Altshuler. 2006. Transferability of tag SNPs in genetic association studies in multiple populations. *Nat. Genet.* 38:1298–1303.
- Donato, L., and F. Guyomarch. 2009. Formation and properties of the whey protein/ κ -casein complexes in heated skim milk—A review. *Dairy Sci. Technol.* 89:3–29.
- Gambra, R., F. Penagaricano, J. Kropp, K. Khateeb, K. A. Weigel, J. Lucey, and H. Khatib. 2013. Genomic architecture of bovine kappa-casein and beta-lactoglobulin. *J. Dairy Sci.* 96:5333–5343.
- Ganai, N. A., H. Bovenhuis, J. A. van Arendonk, and M. H. Visker. 2009. Novel polymorphisms in the bovine beta-lactoglobulin gene and their effects on beta-lactoglobulin protein concentration in milk. *Anim. Genet.* 40:127–133.
- Glantz, M., H. Lindmark Månsson, M. Paulsson, and H. Stålhammar. 2012a. Genomic selection in relation to bovine milk composition and processability. *J. Dairy Res.* 79:53–59.
- Glantz, M., H. Lindmark Månsson, H. Stålhammar, and M. Paulsson. 2012b. Effect of polymorphisms in the leptin, leptin receptor and acyl-CoA:diacylglycerol acyltransferase 1 (*DGAT1*) genes and genetic polymorphism of milk proteins on bovine milk composition. *J. Dairy Res.* 79:110–118.
- Gregersen, V. R., F. Gustavsson, M. Glantz, O. F. Christensen, H. Stålhammar, A. Andrén, H. Lindmark-Månsson, N. A. Poulsen, L. B. Larsen, M. Paulsson, and C. Bendixen. 2015. Bovine chromosomal regions affecting rheological traits in rennet-induced skim milk gels. *J. Dairy Sci.* 98:1261–1272. [10.3168/jds.2014-8136](https://doi.org/10.3168/jds.2014-8136).
- Gustavsson, F., M. Glantz, N. A. Poulsen, L. Wadsö, H. Stålhammar, A. Andrén, H. Lindmark Månsson, L. B. Larsen, M. Paulsson, and W. F. Fikse. 2014. Genetic parameters for rennet- and acid-induced coagulation properties in milk from Swedish Red dairy cows. *J. Dairy Sci.* 97:5219–5229.
- Hallén, E., T. Allmere, A. Lundén, and A. Andrén. 2009. Effect of genetic polymorphism of milk proteins on rheology of acid-induced milk gels. *Int. Dairy J.* 19:399–404.
- Hallén, E., A. Wedholm, A. Andrén, and A. Lundén. 2008. Effect of β -casein, κ -casein and β -lactoglobulin genotypes on concentration of milk protein variants. *J. Anim. Breed. Genet.* 125:119–129.
- Huang, W., F. Penagaricano, K. R. Ahmad, J. A. Lucey, K. A. Weigel, and H. Khatib. 2012. Association between milk protein gene variants and protein composition traits in dairy cattle. *J. Dairy Sci.* 95:440–449.
- International Dairy Federation. 2012. The World Dairy Situation 2012, Bulletin 458/2012. International Dairy Federation, Brussels, Belgium.
- Jakob, E., and Z. Puhani. 1992. Technological properties of milk as influenced by genetic polymorphism of milk proteins—A review. *Int. Dairy J.* 2:157–178.
- Kolbehdari, D., Z. Wang, J. R. Grant, B. Murdoch, A. Prasad, Z. Xiu, E. Marques, P. Stothard, and S. S. Moore. 2009. A whole genome scan to map QTL for milk production traits and somatic cell score in Canadian Holstein bulls. *J. Anim. Breed. Genet.* 126:216–227.
- Lemay, D. G., D. J. Lynn, W. F. Martin, M. C. Neville, T. M. Casey, G. Rincon, E. V. Kriventseva, W. C. Barris, A. S. Hinrichs, A. J. Molenaar, K. S. Pollard, N. J. Maqbool, K. Singh, R. Murney, E. M. Zdobnov, R. L. Tellam, J. F. Medrano, J. B. German, and M. Rijnkels. 2009. The bovine lactation genome: Insights into the evolution of mammalian milk. *Genome Biol.* 10:R43.
- Lucey, J. A. 2009. Milk protein gels. Pages 449–471 in *Milk Proteins—From Expression to Food*. Vol. 1. A. Thompson, M. Boland, and H. Singh, ed. Elsevier Inc., Burlington, MA.
- Lucey, J. A., M. Tamehana, H. Singh, and P. A. Munro. 1998. A comparison of the formation, rheological properties and microstructure of acid skim milk gels made with a bacterial culture or glucono- δ -lactone. *Food Res. Int.* 31:147–155.
- Lundén, A., M. Nilsson, and L. Janson. 1997. Marked effect of beta-lactoglobulin polymorphism on the ratio of casein to total protein in milk. *J. Dairy Sci.* 80:2996–3005.
- Manderson, G. A., M. J. Hardman, and L. K. Creamer. 1998. Effect of heat treatment on the confirmation and aggregation of β -lactoglobulin A, B and C. *J. Agric. Food Chem.* 46:5052–5061.
- Poulsen, N. A., H. P. Bertelsen, H. B. Jensen, F. Gustavsson, M. Glantz, H. L. Månsson, A. Andren, M. Paulsson, C. Bendixen,

- A. J. Buitenhuis, and L. B. Larsen. 2013. The occurrence of non-coagulating milk and the association of bovine milk coagulation properties with genetic variants of the caseins in 3 Scandinavian dairy breeds. *J. Dairy Sci.* 96:4830–4842.
- QTD. 2005. QTD—Download Binaries (updated Aug. 2008), version 2.5.1. Accessed Feb. 27, 2014. <http://www.sph.umich.edu/csg/abecasis/QTD/download/>.
- Reinhardt, T. A., and D. Lippolis. 2006. Bovine milk fat globule membrane proteome. *J. Dairy Res.* 73:406–416.
- Rijnkels, M., C. Freeman-Zadrowski, J. Hernandez, V. Potluri, W. L. Wang, and D. G. Lemay. 2013. Epigenetic modifications unlock the milk protein gene loci during mouse mammary gland development and differentiation. *PLoS ONE* 8:e53270.
- Robinson, R. K., and P. Itsaranuwat. 2006. Properties of yoghurt and their appraisal. Pages 76–94 in *Fermented Milks*. Vol. 1. A. Y. Tamime, ed. Blackwell Science Ltd., Oxford, UK.
- Schopen, G. C., P. D. Koks, J. A. van Arendonk, H. Bovenhuis, and M. H. Visker. 2009. Whole genome scan to detect quantitative trait loci for bovine milk protein composition. *Anim. Genet.* 40:524–537.
- Schopen, G. C. B., M. H. P. W. Visker, P. D. Koks, E. Mullaart, J. A. M. van Arendonk, and H. Bovenhuis. 2011. Whole-genome association study for milk protein composition in dairy cattle. *J. Dairy Sci.* 94:3148–3158.
- van Marle, M. E., and P. Zoon. 1995. Permeability and rheological properties of microbially and chemically acidified skim-milk gels. *Neth. Milk Dairy J.* 49:47–65.
- Walstra, P., J. T. M. Wouters, and T. J. Geurts. 2006. *Dairy Science and Technology*. CRC Press, Taylor & Francis Group, Boca Raton, FL.
- Wigginton, J. E., D. J. Cutler, and G. R. Abecasis. 2005. A note on exact tests of Hardy-Weinberg equilibrium. *Am. J. Hum. Genet.* 76:887–893.
- Zimin, A. V., A. L. Delcher, L. Florea, D. R. Kelley, M. C. Schatz, D. Puiu, F. Hanrahan, G. Pertea, C. P. Van Tassell, T. S. Sonstegard, G. Marcais, M. Roberts, P. Subramanian, J. A. Yorke, and S. L. Salzberg. 2009. A whole-genome assembly of the domestic cow, *Bos taurus*. *Genome Biol.* 10:R42.