



J. Dairy Sci. TBC

<https://doi.org/10.3168/jds.2022-23206>

© TBC, The Authors. Published by Elsevier Inc. and FASS Inc. on behalf of the American Dairy Science Association®.
This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

Single-step genome-wide association analyses for selected infrared-predicted cheese-making traits in Walloon Holstein cows

H. Atashi,^{1,2*} Y. Chen,¹ H. Wilmot,^{1,3} C. Bastin,³ S. Vanderick,¹ X. Hubin,⁴ and N. Gengler¹

¹TERRA Teaching and Research Center, Gembloux Agro-Bio Tech, University of Liège, 5030 Gembloux, Belgium

²Department of Animal Science, Shiraz University, 71441-13131 Shiraz, Iran

³National Fund for Scientific Research (F.R.S.-FNRS), B-1000 Brussels, Belgium

⁴Elevéo asbl Awé Group, 5590 Ciney, Belgium

ABSTRACT

This study aimed to perform genome-wide association study to identify genomic regions associated with milk production and cheese-making properties (CMP) in Walloon Holstein cows. The studied traits were milk yield, fat percentage, protein percentage, casein percentage (CNP), calcium content, somatic cell score (SCS), coagulation time, curd firmness after 30 min from rennet addition, and titratable acidity. The used data have been collected from 2014 to 2020 on 78,073 first-parity (485,218 test-day records), 48,766 second-parity (284,942 test-day records), and 21,948 third-parity (105,112 test-day records) Holstein cows distributed in 671 herds in the Walloon Region of Belgium. Data of 565,533 single nucleotide polymorphisms (SNP), located on 29 *Bos taurus* autosomes (BTA) of 6,617 animals (1,712 males), were used. Random regression test-day models were used to estimate genetic parameters through the Bayesian Gibbs sampling method. The SNP solutions were estimated using a single-step genomic BLUP approach. The proportion of the total additive genetic variance explained by windows of 50 consecutive SNPs (with an average size of ~216 KB) was calculated, and regions accounting for at least 1.0% of the total additive genetic variance were used to search for positional candidate genes. Heritability estimates for the studied traits ranged from 0.10 (SCS) to 0.53 (CNP), 0.10 (SCS) to 0.50 (CNP), and 0.12 (SCS) to 0.49 (CNP) in the first, second, and third parity, respectively. Genome-wide association analyses identified 6 genomic regions (BTA1, BTA14 [4 regions], and BTA20) associated with the considered traits. Genes including the *SLC37A1* (BTA1), *SHARPIN*, *MROH1*, *DGAT1*, *FAM83H*, *TIGD5*, *MROH6*, *NAPRT*, *ADGRB1*, *GML*, *LYPD2*, *JRK* (BTA14), and *TRIO* (BTA20) were identified as positional candidate genes

for the studied CMP. The findings of this study help to unravel the genomic background of a cow's ability for cheese production and can be used for the future implementation and use of genomic evaluation to improve the cheese-making traits in Walloon Holstein cows.

Key words: milk composition, cheese-making, genomic association, Holstein

INTRODUCTION

The cheese-making properties (CMP) of bovine milk are economically important for the dairy industry since a significant and growing fraction of the milk produced worldwide is used to make cheese (Wedholm et al., 2006; Cassandro et al., 2008; Food and Agriculture Organization of the United Nations, 2015; The World Dairy Situation, 2016). Milk coagulation is a key step that strongly influences the efficiency of cheese production. Furthermore, the composition, in particular casein and fat, and physicochemical properties of milk have considerable effects on its processing properties and the quality of resulting dairy products (Walstra et al., 2005; Visentin et al., 2017b; Nilsson et al., 2019). The coagulation ability of milk is also associated with the pH of milk, protein composition, SCC, milk fatty acids (FA), and milk calcium content (CC) (Bencini, 2002; Pastorino et al., 2003; Wedholm et al., 2006; Bobbo et al., 2016). The coagulation ability of milk can be evaluated using developed milk coagulation properties such as rennet coagulation time (RCT, min), which is the time from the addition of coagulant to milk to the beginning of coagulation, the time to a curd firmness of 20 mm, and curd firmness at 30 min after coagulant addition (a30, mm) (De Marchi et al., 2007; Bittante, 2011; Visentin et al., 2015). It has been documented that the CMP can be affected by environmental factors such as feeding, udder health, season, and physiological stage (e.g., parity, lactation stage), but they are also genetically influenced (De Marchi et al., 2007; Cassandro et al., 2008; Visentin et al., 2017a; Atashi et al., 2022a). Therefore, genetic selection can be used

Received December 28, 2022.

Accepted May 1, 2023.

*Corresponding author: atashi@shirazu.ac.ir

to improve cheese-making traits, and consequently, to improve the quality and the amount of cheese yield per volume of milk.

There are different laboratory methods developed for direct recording of individual milk coagulation properties (Preto et al., 2011; Bittante et al., 2012; Bobbo et al., 2016); however, they are time consuming and expensive to implement on a large scale; then, genetic selection based on direct measures of milk coagulation properties is limited (Sanchez et al., 2018, 2019). Fourier-transform mid-infrared (MIR) spectroscopy has been proposed as an alternative method for prediction of various milk characteristics, including fractions of protein, fat, casein, minerals, and milk FA contents (De Marchi et al., 2009, 2014; Soyeurt et al., 2009, 2011; Visentin et al., 2015). The MIR technology also is considered as a cheap method to measure individual CMP on a large scale, which is needed for breeding programs in dairy cattle (De Marchi et al., 2009, 2014; Sanchez et al., 2018).

However, enough knowledge on the genetic background of cheese-making traits is needed before they are included into the breeding program. Furthermore, identification of genomic regions and individual genes responsible for genetic variation in CMP will improve our understanding about the biological pathways involved and can be used for improving cheese-making traits. Atashi et al. (2022a) investigated the genetic parameters and genomic regions associated with selected infrared-predicted cheese-making traits in the Dual-Purpose Belgian Blue (DPBB) population, which is the second important cattle breed reared by dairy farmers in the Walloon Region of Belgium. However, to the best of our knowledge, no comprehensive study has been performed to investigate genetic background and the genetic architecture of cheese-making traits in Walloon Holstein cows. Therefore, the aim of this study was to estimate genetic parameters and identify genomic regions associated with milk production and selected infrared-predicted cheese-making traits in Walloon Holstein cows.

MATERIALS AND METHODS

Phenotypic Data

No human or animal subjects were used, so this analysis did not require approval by an Institutional Animal Care and Use Committee or Institutional Review Board. The data used consisted of test-day records of traits including milk yield (MY), SCC, MIR predicted fat percentage (FP), protein percentage (PP), casein percentage (CNP), CC, coagulation time (CT), a30,

and titratable acidity (TA). The FP and PP records were generated by the official milk recording in the Walloon Region of Belgium using MIR spectrometry and commercially available instruments and calibrations from FOSS (Foss Electric A/S). Test-day SCC records were log-transformed to SCS based on the following equation: $SCS = \log_2(SCC/100,000) + 3$.

The MIR prediction equations used to predict CNP, CC, CT, a30, and TA were obtained from various studies (Soyeurt et al., 2009; Colinet et al., 2010, 2013, 2015). Table 1 shows the calibration and the cross-validation statistics of the used MIR prediction equations. Milk MIR spectra were obtained by the analysis of individual milk samples on MilkoScan FT6000 spectrometer (Foss, Hillerød, Denmark). The MIR spectra were preprocessed to remove baseline variation and then were standardized. The MIR prediction equations were applied on standardized spectra from individual milk samples collected in the frame of the Walloon milk recording scheme. Full cross-validation was performed to assess the accuracy of the developed equations. The cross-validation coefficients of determination were 0.95, 0.82, 0.63, 0.42, and 0.68 for CNP, CC, CT, a30, and TA, respectively. The ratio performance/deviation of cross-validation, the ratio of standard deviation (SD) to standard error of cross-validation, was 4.47, 2.34, 1.64, 1.31, and 1.77 for CNP, CC, CT, a30, and TA, respectively.

Data were edited to include only cows with known birth date, calving date, and parity number. Only records from the first 3 parities that had data for all included traits on a given test-day were kept. Records from DIM lower than 5 d and greater than 365 d were eliminated. Age at the first calving was calculated as the difference between birth date and first calving date and restricted to the range of 540 to 1,200 d. Daily MY, FP, and PP were restricted to the range from 3 to 99 kg, 1 to 9%, and 1 to 7%, respectively (ICAR, 2022). Test-day records of the other considered traits were edited to remove records outside the range of mean \pm 5 SD. Within cow, if parity 3 was present, parities 1 and 2 were also present, and if parity 2 was present, parity 1 was also present. The number of test-day records in the first-, second-, and third-parity cows were 485,218 (on 78,073 cows), 284,942 (on 48,766 cows), and 105,112 (on 21,941 cows), respectively. The data were collected from 2014 to 2020 on 78,073 animals distributed in 671 herds in the Walloon Region of Belgium. On average across the data set, 5.88 test-day records were available per cow per lactation. Pedigree depth of the animals was traced back to 5 generations. The used pedigree consisted of 186,548 females and 10,076 males.

Genotypic Data

Genotypic data were available for 6,617 (1,712 males and 4,905 females) phenotyped animals or those animals included in the pedigree. Individuals were genotyped using the BovineSNP50 Beadchip v1 to v3 and EuroG MD (SI) v9 (Illumina, San Diego, CA). Single nucleotide polymorphisms in common among the 4 chips were kept. Nonmapped SNP, SNP located on sexual chromosomes, and triallelic SNPs were excluded. A minimum GenCall Score of 0.15 and a minimum GenTrain Score of 0.55 were used to keep SNP. Then, the genotypes were imputed to HD with a reference population of 4,352 HD individuals (1,046 males and 3,288 females) using FImpute V2.2 software (Sargolzaei et al., 2014). Single nucleotide polymorphisms with Mendelian conflicts and those with minor allele frequency (**MAF**) less than 5% were excluded. The difference between the observed and expected heterozygosity was estimated, and if the difference was greater than 0.15, the SNP was excluded (Wiggans et al., 2009). Finally, 565,533 SNPs located on 29 BTA were used in the genomic analyses.

Variance Component Estimation

The (co)variance components and breeding values for the considered cheese-making properties were estimated based on the integration of the random regression test-day model (**RR-TDM**) into the single-step GBLUP procedure (**SS RR-TDM**) using the following single-trait, multiple-lactation (first 3 lactations) model (Paiva et al., 2022):

$$y_{ijklmn} = \mu + \text{HTD}p_i + \sum_{b=0}^3 \text{AS}_j \Phi_b(t) + \sum_{b=0}^3 \text{HY}_k \Phi_b(t) + \sum_{b=0}^3 p_{e_m} \Phi_b(t) + \sum_{b=0}^3 a_m \Phi_b(t) + e_{ijklmn},$$

where y_{ijklmn} is the test-day record (milk yield, milk composition, and cheese-making traits) belonging to the DIM n of cow m in parity l , belonging to the i th class of HTDp, j th class of AS, and k th class of HY; HTDp is the fixed effect of herd-test-day-parity; AS is the fixed effect of age-season of calving defined as follows: age at calving class (6, 4, and 2 classes of age at calving were created for the first, second, and third parity, respectively) \times season of calving (4 seasons: winter from January to March, spring from April to June, summer from July to September, and autumn from October to December); $\sum_{b=0}^3 \text{AS}_j \Phi_b(t)$ is the fixed regression coefficients of the age-season of calving modeled using Legendre polynomials of order 3; $\sum_{b=0}^3 \text{HY}_k \Phi_b(t)$, $\sum_{b=0}^3 p_{e_m} \Phi_b(t)$, and $\sum_{b=0}^3 a_m \Phi_b(t)$ are, respectively, the random regression coefficients of herd-year of calving, permanent environmental, and additive effects modeled using Legendre polynomials of order 3; and e_{ijklmn} is the residual effect. The herd-year of calving, permanent environment, additive genetic, and residual variances were assumed as follows:

$$\text{Var} \begin{bmatrix} \text{hy} \\ \text{pe} \\ \text{a} \\ \text{e} \end{bmatrix} = \begin{bmatrix} \mathbf{HY} \otimes \mathbf{I} & 0 & 0 & 0 \\ 0 & \mathbf{P} \otimes \mathbf{I} & 0 & 0 \\ 0 & 0 & \mathbf{Ga} \otimes \mathbf{H} & 0 \\ 0 & 0 & 0 & \mathbf{R} \end{bmatrix},$$

Table 1. The calibration and the cross-validation statistics of the mid-infrared prediction equations used

Trait	Calibration statistics ¹							Cross-validation statistics ²		
	N	Mean	SD	Min	Max	SEC	R ² c	SECV	R ² cv	RPDcv
Calcium content (mg/kg)	1,094	1,149.9	124.7	593	1,743	51.8	0.83	53.38	0.82	2.34
Casein (%)	976	2.67	0.34	1.61	4.04	0.074	0.95	0.076	0.95	4.47
Titratable acidity ³ (°D)	930	15.97	1.84	9.50	23.40	1.00	0.71	1.040	0.68	1.77
Coagulation time ⁴ (s)	854	1,088.6	262.4	574.5	1,942	154.6	0.65	159.9	0.63	1.64
Curd firmness ⁵ (mm)	547	33.06	4.22	20.60	43.50	3.08	0.47	3.22	0.42	1.31

¹Min = minimum; Max = maximum; SEC = standard error of calibration; R²c = coefficient of determination of calibration.

²SECV = standard error of cross-validation; R²cv = coefficient of determination of cross-validation; RPDcv = ratio performance/deviation (RPD) of cross-validation defined as the ratio SD to root mean square error of cross-validation.

³Milk titratable acidity in Dornic degrees (°D).

⁴Coagulation time is defined here as the sum of the rennet coagulation time plus the time to a curd firmness of 20 mm measured by the Computerized Renneting Meter.

⁵Curd firmness (a30) is defined as the curd firmness measured 30 min after enzyme addition by the Computerized Renneting Meter.

where \mathbf{HY} is the 12×12 covariance matrix of the herd-year of calving regression coefficients; \mathbf{I} is an identity matrix, \otimes represents the Kronecker product function, \mathbf{P} is the 12×12 covariance matrix of the permanent environmental regression coefficients; \mathbf{Ga} is the 12×12 covariance matrix of the additive genetic regression coefficients, blocks within $\mathbf{R} = \sum_{\mathbf{p}} \mathbf{r}^{\mathbf{p}}$ contain residual variance (\mathbf{r}) that depends on parity (\mathbf{p}). Residual variance was the same within each parity. The \mathbf{H} is a matrix that combines pedigree and genomic relationships, and its inverse consists of the integration of additive and genomic relationship matrices, \mathbf{A} and \mathbf{G} , respectively (Aguilar et al., 2010):

$$\mathbf{H}^{-1} = \mathbf{A}^{-1} + \begin{bmatrix} 0 & 0 \\ 0 & \mathbf{G}^{-1} - \mathbf{A}_{22}^{-1} \end{bmatrix},$$

where \mathbf{A} is the numerator relationship matrix based on the pedigree for all animals, \mathbf{A}_{22} is the numerator relationship matrix for genotyped animals, and \mathbf{G} is the weighted genomic relationship matrix obtained using the following function:

$$\mathbf{G} = \mathbf{G}^* \times 0.95 + \mathbf{A}_{22} \times 0.05.$$

The \mathbf{G}^* is the genomic relationship matrix obtained using the following function described by VanRaden (2008):

$$\mathbf{G}^* = \frac{\mathbf{ZDZ}'}{\sum_{i=1}^M 2p_i(1-p_i)},$$

where \mathbf{Z} is a matrix of gene content adjusted for allele frequencies (0, 1, or 2 for aa , Aa , and AA , respectively); \mathbf{D} is a diagonal matrix of weights for SNP variances ($\mathbf{D} = \mathbf{I}$); M is the number of SNPs, and p_i is the MAF of the i th SNP. The \mathbf{H} matrix was built scaling \mathbf{G} based on \mathbf{A}_{22} considering that the average of the diagonal of \mathbf{G} is equal to the average of the diagonal of \mathbf{A}_{22} , and the average of the off-diagonal \mathbf{G} is equal to the average of the off-diagonal \mathbf{A}_{22} .

The (co)variance components were estimated by Bayesian inference using the GIBBS3F90 software (Aguilar et al., 2018). Gibbs sampling was used to obtain marginal posterior distributions for the various parameters using a single chain of 200,000 iterates with a sampling interval of 20 samples. The first 50,000 iterates of the chain were regarded as a burn-in period to allow sampling from the proper marginal distributions. Post-Gibbs analysis was performed using the software POSTGIBBSF90 (Aguilar et al., 2018), using the re-

tained 7,500 samples. Genetic (co)variances on each test-day were calculated using the equation described by Jamrozik and Schaeffer (1997). Daily heritability was defined as the ratio of genetic variance to the sum of the additive genetic, permanent environmental, herd-year calving, and residual variances at a given DIM.

The vector of genomic estimated breeding values (\mathbf{GEBV}) of the included traits for each animal i , which included daily GEBV from all DIM (5 to 365) in each parity, was estimated by multiplying the vector of additive genetic predicted regression coefficients by the matrix of Legendre orthogonal polynomial covariates; that is, $\mathbf{GEBV}_i = \mathbf{T}\mathbf{g}_i$, where \mathbf{g}_i is the vector of additive genetic predicted regression coefficients for animal i , and \mathbf{T} is a matrix of orthogonal covariates associated with the Legendre orthogonal polynomial functions.

Genome-Wide Association Study

The GWAS analyses were performed for all included traits in the first 3 parities considering following 3 lactation stages: (1) from 5 to 60 DIM, representing the ascending production stage and lactation peak; (2) from 61 to 200 DIM, representing the middle lactation stage; and (3) from 201 to 365 DIM, representing the production decline up to the end of the lactation (Oliveira et al., 2019). Therefore, the GEBV for each lactation stage of each animal i (for each trait in each parity) were obtained by averaging (summing for MY) the daily GEBV solutions of the specific DIM; that is,

$$\mathbf{GEBV}_{1i} = \mathbf{GEBV}_{i5} + \mathbf{GEBV}_{i6} + \dots + \mathbf{GEBV}_{i60},$$

$$\mathbf{GEBV}_{2i} = \mathbf{GEBV}_{i61} + \mathbf{GEBV}_{i62} + \dots + \mathbf{GEBV}_{i200}, \text{ and}$$

$$\mathbf{GEBV}_{3i} = \mathbf{GEBV}_{i201} + \mathbf{GEBV}_{i202} + \dots + \mathbf{GEBV}_{i365},$$

where \mathbf{GEBV}_{1i} , \mathbf{GEBV}_{2i} , and \mathbf{GEBV}_{3i} are the GEBV for the first, second, and third lactation stages of animal i obtained by averaging (summing for MY) the GEBV from 5 to 60, 61 to 200, and 201 to 365 DIM, respectively. Furthermore, the GEBV of animal i through the entire lactation were obtained by summing (averaging) the daily GEBV solutions of all DIM; that is,

$$\mathbf{GEBV}_{ei} = \mathbf{GEBV}_{i5} + \mathbf{GEBV}_{i6} + \dots + \mathbf{GEBV}_{i365},$$

where \mathbf{GEBV}_{ei} is the GEBV of animal i through the entire lactation, obtained by averaging (summing for MY) the GEBV from 5 to 365.

The SNP effects for each lactation stage were estimated individually for each trait in each parity using

the postGSf90 software (Aguilar et al., 2014). The animal effects were decomposed into those for genotyped (\mathbf{a}_g) and ungenotyped animals (\mathbf{a}_n). The animal effects of genotyped animals are a function of the SNP effects, $\mathbf{a}_g = \mathbf{Z}\mathbf{u}$, where \mathbf{Z} is a matrix relating genotypes of each locus and \mathbf{u} is a vector of the SNP marker effect. The variance of animal effects was assumed as

$$\text{Var}(\mathbf{a}_g) = \text{Var}(\mathbf{Z}\mathbf{u}) = \mathbf{Z}\mathbf{D}\mathbf{Z}'\sigma_u^2 = \mathbf{G}\sigma_a^2,$$

where \mathbf{D} is a diagonal matrix of weights for variances of markers ($\mathbf{D} = \mathbf{I}$), and σ_u^2 is the additive genetic variance captured by each SNP marker when the weighted relationship matrix (\mathbf{G}) was built with no weight. The SNP effects were obtained using the following equation:

$$\hat{\mathbf{u}} = \lambda \mathbf{D}\mathbf{Z}' \mathbf{G}^{-1} \hat{\mathbf{a}}_g = \mathbf{D}\mathbf{Z}' [\mathbf{Z}\mathbf{D}\mathbf{Z}']^{-1} \hat{\mathbf{a}}_g,$$

where λ was defined by VanRaden (2008) as a normalizing constant, as described below:

$$\lambda = \frac{\sigma_u^2}{\sigma_a^2} = \frac{1}{\sum_{i=1}^M 2p_i(1-p_i)}.$$

The percentage of the total additive genetic variance explained by the i th genomic region was estimated as follows:

$$\frac{\text{Var}(\mathbf{a}_i)}{\sigma_a^2} \times 100\% = \frac{\text{Var}\left(\sum_{j=1}^{50} \mathbf{Z}_j \hat{\mathbf{u}}_j\right)}{\sigma_a^2} \times 100,$$

where \mathbf{a}_i is the genetic value of the i th region that consists of 50 adjacent SNPs, σ_a^2 is the total additive genetic variance, \mathbf{Z}_j is the vector of the SNP content of the j th SNP for all individuals, and $\hat{\mathbf{u}}_j$ is the marker effect of the j th SNP within the i th region. The additive genetic variance explained by 50-SNP moving windows, with an average size of ~ 216 KB, was calculated across the whole genome, and those windows explaining at least 1.0% of the total additive genetic variance were considered promising regions and used to identify positional candidate genes. The concept of grouping SNP into windows was adopted as a way to better capture the genetic information such as the extent of linkage disequilibrium (**LD**) in neighboring SNPs (Habier et al., 2011).

Identification of Positional Candidate Genes for the Studied Traits

The animals included in this study were genotyped using the BovineSNP50 Beadchip v1 to v3 and EuroG MD (SI) v9 (Illumina, San Diego, CA); then, the genotypes were imputed to the BovineHD Beadchip, which is based on the bovine reference genomes assembly UMD3.1. However, new bovine reference genome assembly ARS-UCD1.2, assembled using long sequencing reads, filled gaps and resolved repetitive regions of the UMD3.1 assembly, and has more credible annotation information (Rosen et al., 2020). The Lift Genome Annotations tool, available through a simple web interface (<https://genome.ucsc.edu/cgi-bin/hgLiftOver>), was used to convert coordinate ranges of the identified genomic regions from the UMD3.1 to the ARS-UCD1.2 assembly. Then, to identify possible candidate genes associated with the considered traits, genes located within the identified genomic regions (i.e., between the start and end of genomic coordinates of the identified regions based on the ARS-UCD1.2 assembly) were further investigated. We identified genes using the National Center for Biotechnology Information (NCBI) Map Viewer tool for the ARS-UCD1.2 assembly as the reference map.

RESULTS AND DISCUSSION

The lactation curves of daily average of phenotypic records for the considered traits are presented in Figure 1. The peak of MY occurred at the DIM 41 (27.59 kg), 35 (35.03 kg), and 32 (39.40 kg) for the first, second, and third parity, respectively. The FP, PP, and CNP curves were observed to be lower for first-parity cows than for second- and third-parity cows. The TA curves were higher for first-parity cows than for second- and third-parity cows. The descriptive statistics for studied traits in the first 3 lactations are presented in Table 2. Daily MY averaged 24.2, 27.9, and 30.0 kg in the first 3 lactations, which is comparable to those previously reported for Walloon Holstein cows (Bastin et al., 2013). Somatic cell score was the trait with the greatest coefficient of variation (53.54 to 59.45%), whereas a30 had the lowest coefficient of variation (7.17 to 7.54%). Cassandro et al. (2008) reported that among milk coagulation and production traits (MY, FP, PP, CNP, SCS, TA, RCT, and TA) of Italian Holstein, SCS has the highest and casein percentage has the lowest coefficient of variation. The average a30 ranged from 32.11 to 32.33 mm, which is in agreement with previous studies (Cassandro et al., 2008; Atashi et al., 2022a). However, mean a30 reported for Finnish Ayrshire cows ranged from 25 to 27 mm (Ikonen et al., 2004). The av-

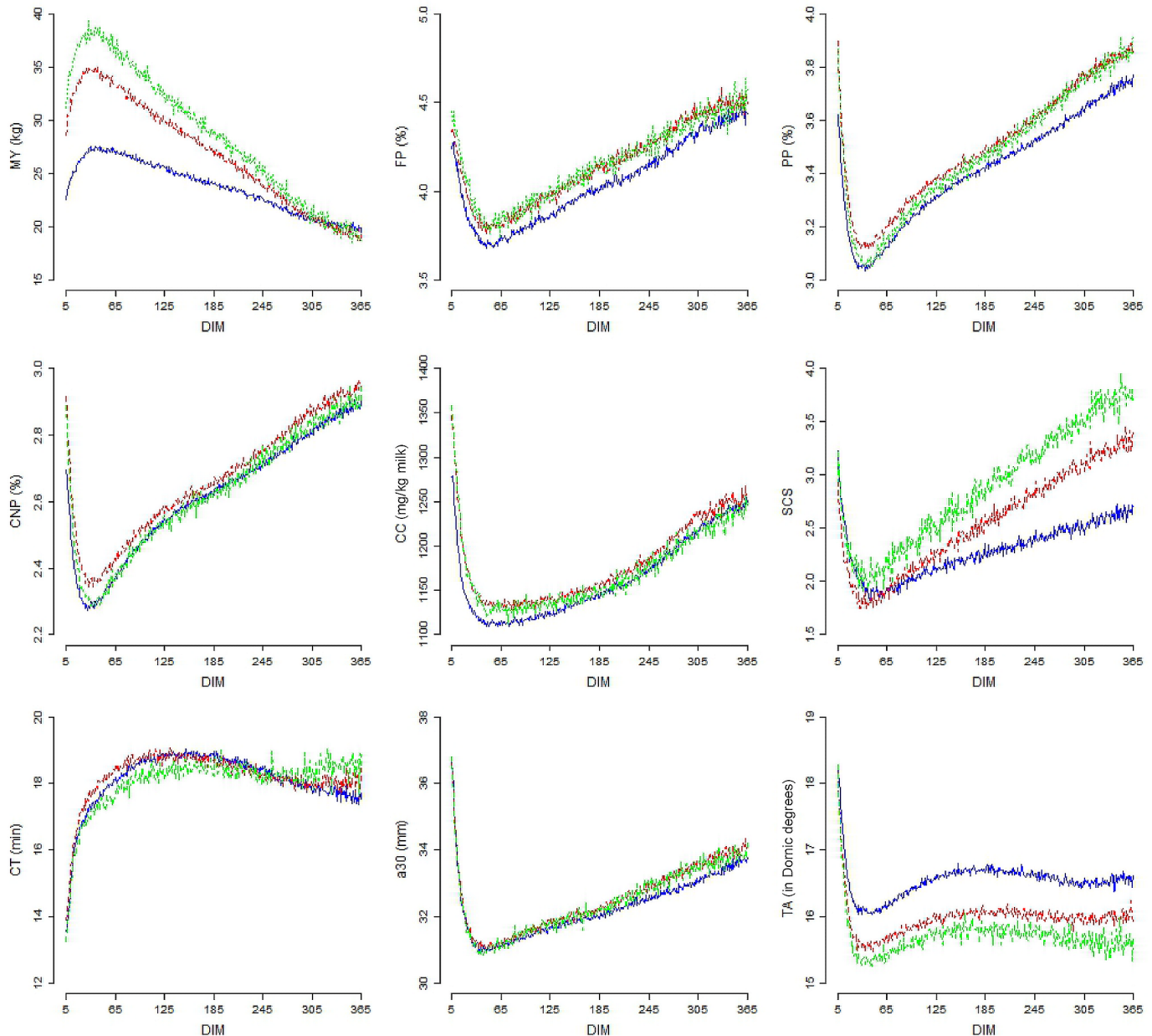


Figure 1. Lactation curves for milk yield (MY), fat percentage (FP), protein percentage (PP), casein percentage (CNP), milk calcium content (CC) expressed as milligrams per kilogram of milk, SCS, coagulation time (CT) expressed in minutes, curd firmness after 30 min from rennet addition (a30) expressed in millimeters, and titratable acidity (TA) in Dornic degrees for the first (blue), second (red), and third (green) parity in Walloon Holstein cows.

erage (SD) SCS were 2.25 (1.36), 2.46 (1.46), and 2.77 (1.48) in the first, second, and third parity, respectively. Averaged MY, and SCS were higher with increasing parity in line with previous studies (Bastin et al., 2013; Atashi and Hostens, 2021), whereas averaged TA was lower with increasing parity as has been reported for DPBB cows (Atashi et al., 2022a).

Average daily heritability (h^2) estimates for the studied traits are shown in Table 3. Heritability estimates ranged from 0.10 (SCS) to 0.53 (CNP), 0.10 (SCS) to 0.50 (CNP), and 0.12 (SCS) to 0.49 (CNP) in the first, second, and third parity, respectively. Heritability estimates were generally higher in first parity than in later parities. The mean daily h^2 estimates for CT, a30, and TA ranged from 0.47 to 0.52, 0.46 to 0.47, and 0.46 to

Table 2. Descriptive statistics for milk yield traits and cheese-making properties in Walloon Holstein cows¹

Trait	First lactation			Second lactation			Third lactation		
	Mean	SD	CV (%)	Mean	SD	CV (%)	Mean	SD	CV (%)
Milk yield (kg)	24.2	6.01	24.86	27.9	8.29	29.68	30.0	9.11	30.34
Fat (%)	4.01	0.68	15.94	4.10	0.67	16.37	4.11	0.67	16.25
Protein (%)	3.38	0.34	10.04	3.46	0.37	10.73	3.43	0.37	10.92
Casein (%)	2.59	0.29	11.23	2.63	0.31	11.69	2.59	0.31	11.85
Calcium content (mg/kg milk)	1,155	100.0	8.66	1,172	106.9	9.13	1,164	105.4	9.06
SCS ²	2.25	1.36	53.54	2.46	1.46	59.45	2.77	1.48	53.54
Coagulation time ³	18.16	2.99	16.46	18.25	3.05	16.72	18.00	3.64	17.47
Curd firmness ⁴ (mm)	32.11	2.30	7.17	32.31	2.44	7.54	32.33	2.39	7.41
Titratable acidity ⁵ (°D)	16.52	1.41	8.52	16.0	1.44	9.03	15.7	1.47	9.39

¹The number of data were 485,218 test-day records (on 78,073 animals), 284,942 test-day records (on 48,766 animals), and 105,112 test-day records (on 21,948 animals) in the first, second, and third lactations, respectively.

²SCS = $\log_2(\text{SCC}/100,000) + 3$.

³Coagulation time is defined here as the sum of the rennet coagulation time plus the time to a curd firmness of 20 mm measured by the Computerized Renneting Meter.

⁴Curd firmness (a30) is defined as the curd firmness measured 30 min after enzyme addition by the Computerized Renneting Meter.

⁵Milk titratable acidity in Dornic degrees (°D).

0.51 in agreement with previously reported estimates for dairy cattle (Vallas et al., 2010; Tiezzi et al., 2013). Atashi et al. (2022a) reported that h^2 of infrared-predicted CT, a30, and TA in DPBB ranged from 0.40 to 0.48, 0.36 to 0.39, and 0.41 to 0.50, respectively. The high h^2 estimated for milk coagulation properties supports possible genetic improvement of milk coagulation ability in the population of Walloon Holstein cows. Cecchinato et al. (2011) reported that h^2 of RCT (min), a30, and TA ranged, respectively, from 0.22 to 0.58, 0.05 to 0.32, and 0.11 to 0.42 in Italian Holstein cows. Colinet et al. (2012) reported that the daily h^2 of infrared-predicted TA in the first-parity Holstein cows in Wallonia ranged from 0.45 to 0.60 with an average

of 0.57. Mean daily h^2 of CC ranged from 0.47 to 0.50, which is in line with those reported for Montbéliarde dairy cows (Sanchez et al., 2021). The variation found for h^2 of CMP in the literature can be explained by the differences in the studied breeds, structure of the data, number of records, statistical models, and the length of the period of data collection.

Typically, GWAS methods are based on testing the significance of SNP effects on the traits of interest. However, SNPs within a genomic region can be highly correlated and jointly influence the phenotype. Furthermore, the genetic information in neighboring SNPs, such as the extent of LD, is not used in the GWAS depends on single SNP (Bao and Wang, 2017). Therefore,

Table 3. Mean (SD) daily heritability for milk yield traits and cheese-making properties estimated across the lactation in the first 3 parities in Walloon Holstein cows¹

Trait	First lactation		Second lactation		Third lactation	
	Mean (SD)	Range	Mean (SD)	Range	Mean (SD)	Range
Milk yield (kg)	0.25 (0.04)	0.16–0.29	0.25 (0.04)	0.11–0.24	0.25 (0.04)	0.10–0.22
Fat (%)	0.48 (0.10)	0.22–0.57	0.49 (0.10)	0.22–0.59	0.47 (0.10)	0.19–0.56
Protein (%)	0.51 (0.10)	0.22–0.60	0.48 (0.12)	0.17–0.60	0.47 (0.12)	0.17–0.59
Casein (%)	0.53 (0.10)	0.25–0.61	0.50 (0.12)	0.19–0.61	0.49 (0.12)	0.19–0.61
Calcium content (mg/kg milk)	0.50 (0.08)	0.31–0.57	0.47 (0.08)	0.28–0.56	0.47 (0.08)	0.26–0.57
SCS ²	0.10 (0.01)	0.07–0.11	0.10 (0.02)	0.05–0.14	0.12 (0.03)	0.07–0.17
Coagulation time ³	0.52 (0.08)	0.29–0.60	0.48 (0.07)	0.30–0.54	0.47 (0.07)	0.30–0.53
Curd firmness ⁴ (mm)	0.46 (0.09)	0.21–0.54	0.46 (0.09)	0.21–0.55	0.47 (0.09)	0.21–0.55
Titratable acidity ⁵ (°D)	0.51 (0.07)	0.32–0.57	0.47 (0.07)	0.28–0.53	0.46 (0.07)	0.27–0.52

¹The number of data were 485,218 test-day records (on 78,073 animals), 284,942 test-day records (on 48,766 animals), and 105,112 test-day records (on 21,948 animals) in the first, second, and third lactations, respectively.

²SCS = $\log_2(\text{SCC}/100,000) + 3$.

³Coagulation time is defined here as the sum of the rennet coagulation time plus the time to a curd firmness of 20 mm measured by the Computerized Renneting Meter.

⁴Curd firmness (a30) is defined as the curd firmness measured 30 min after enzyme addition by the Computerized Renneting Meter.

⁵Milk titratable acidity in Dornic degrees (°D).

window-based GWAS procedure have been proposed as an effective procedure to estimate the combined effect of several consecutive SNPs in a specific region and identify genomic regions explaining a given amount of genetic variance (Aguilar et al., 2019). Window-based GWAS may use different window types (distinct or sliding windows) and variable window sizes (defined as the number of SNP or the number of base pairs). However, the absence of a universal approach for hypothesis testing is an important challenge of window-based GWAS, even though it is quite a common procedure in genetic studies. The common form for declaring significance is to use a threshold on the additive genetic variance explained by individual window (Aguilar et al., 2019). However, it is unclear what window size is optimal, and no standard presently exists to define the threshold on explained genetic variance. Therefore, determining the proper window size is usually subjective, and researchers often have not justified their choices or sometimes have acknowledged that their choices are arbitrary (Beissinger et al., 2015). Medeiros de Oliveira Silva et al. (2017), using the BovineHD SNP panel, considered 50 adjacent SNP windows (with an average of 280 KB) that explained at least 0.50% of additive genetic variance as the threshold to declare significance. Atashi et al. (2020), using the BovineHD SNP panel, considered 50 adjacent SNP windows that explained more than 1% of the total additive genetic variance as the threshold to declare significance for milk production and lactation curve parameters. Han and Peñagaricano (2016) considered 1.5-MB windows that explained at least 0.50% of the total genetic variance as the threshold to declare significance. Suwannasing et al. (2018) considered windows that explained more than 1% of the total genetic variance as the threshold to declare significance. Tiezzi et al. (2015) calculated the variance absorbed by 10-SNP moving windows and reported the 10 windows explaining the largest amount of genomic variance as the most important windows. In this study, a window-based GWAS through the single-step genomic best linear unbiased predictor (**ssGBLUP**) was used. The results were presented by the proportion of total genetic variance explained by a window of 50 adjacent SNPs with an average size of ~216 KB and windows explaining for at least 1.0% of the total additive genetic variance were used to search for positional candidate genes. We used 1 SNP as the moving step of the window, which ensured that we do not miss genomic regions potentially associated with the traits due to the combination of SNPs.

General information (starts and end SNP numbers, window size, start and end genomic positions, and the variance explained by each windows) about the results of single-step GWAS for the included milk production

and cheese-making traits are presented in Supplemental Data S1–S108 (9 traits [MY, FP, PP, CNP, CC, SCS, CT, a30, and TA] \times 3 parities \times 4 stages per parity). The Manhattan plots of the proportion of total additive genetic variance explained by 50-SNP windows are shown in Supplemental Figures S1–S9. The windows associated with the studied traits along with corresponding genes are presented in Table 4. In total, 6 genomic regions distributed over 3 chromosomes (BTA1, BTA14 [4 regions], and BTA20) were identified that are associated with one or more of the included traits. However, there was no genomic region that explained more than 1.0% of the total additive genetic variance of SCS. The following are the results discussed by chromosome.

BTA1

The genomic region located from 144.38 to 144.47 MB (UMD3.1 assembly) on BTA1 was associated with TA and CT. This window was 84.82 KB in size and explained more than 2.0% of the total additive genetic variance of TA in all defined lactation stages and more than 1.0% of the total additive genetic variance of CT in the first 2 stages of first 3 lactations. Moderate negative genetic correlations were found between TA and CT (-0.46 in the first 3 lactations), which may explain why these traits are affected by the same genomic region (Supplemental Tables S1–S3). This region has been previously reported to be associated with PP, TA, CT, a30, and CNP in the DPBB cows (Atashi et al., 2022a,b). Iung et al. (2019) reported that this region is associated with SCS and MY in the Brazilian Holstein population. Sanchez et al. (2021) reported that this region is associated with milk mineral content of magnesium (Mg), potassium (K), sodium (Na), and phosphorus (P) in Montbéliarde dairy cows. This region harbors the solute carrier family 37 member 1 (*SLC37A1*) gene. *SLC37A1* is highly expressed in the mammary glands (Chamberlain et al., 2015; Raven et al., 2016) and encodes a glucose-6-phosphate transporter involved in the blood glucose homeostasis and sugar transport (Pan et al., 2011). It has been reported that the *SLC37A1* gene has a very strong effect on milk mineral contents (Sanchez et al., 2021; Zaalberg et al., 2021) and is associated with MY, FP, and PP (Kemper et al., 2015; Raven et al., 2016; Pausch et al., 2017; Sanchez et al., 2017), milk FA profile and SCS (Iung et al., 2019), and milk's CMP (Sanchez et al., 2019) in dairy cows. Sanchez et al. (2019) reported that the *SLC37A1* plays a role in inorganic anion transport and is a good candidate for CMP and milk composition. This gene has been shown to be associated with casein phosphorylation (Fang et al., 2019) and participates

Table 4. Genomic regions associated with milk yield traits and cheese-making properties in the first 3 parities in Wallow Holstein cows¹

Chromosome	Position ² (bp)	Position ³ (bp)	Gene ⁴	Trait ⁵ (parity, stage of lactation, ⁶ % variance explained)
BTA1	144,380,987–144,465,804	142,817,724–142,902,514	<i>SLC37A1</i>	CT (1, 1, 1.44), CT (1, 2, 1.02), CT (1, 3, 0.76), CT (1, e, 1.01), CT (2, 1, 1.35), CT (2, 2, 1.01), CT (2, 3, 0.62), CT (2, e, 0.85), CT (3, 1, 1.39), CT (3, 2, 1.01), CT (3, 3, 0.71), CT (3, e, 0.81), TA (1, 1, 4.52), TA (1, 2, 3.92), TA (1, 3, 3.05), TA (1, e, 3.65), TA (2, 1, 4.22), TA (2, 2, 3.48), TA (2, 3, 2.50), TA (2, e, 3.16), TA (3, 1, 4.13), TA (3, 2, 3.41), TA (3, 3, 2.39), TA (3, e, 3.08), MY (1, 1, 2.74), MY (1, 2, 3.58), MY (1, 3, 3.25), MY (1, e, 2.40), MY (2, 1, 3.09), MY (2, 2, 3.89), MY (2, 3, 2.37), MY (2, e, 3.31), MY (3, 1, 3.02), MY (3, 2, 3.86), MY (3, 3, 2.24), MY (3, e, 3.22), FP (1, 1, 15.51), FP (1, 2, 17.63), FP (1, 3, 17.56), FP (1, e, 17.65), FP (2, 1, 16.94), FP (2, 2, 17.75), FP (2, 3, 17.49), FP (2, e, 17.71), FP (3, 1, 16.61), FP (3, 2, 17.74), FP (3, 3, 17.48), FP (3, e, 17.68), PP (1, 1, 4.66), PP (1, 2, 6.34), PP (1, 3, 6.15), PP (1, e, 6.26), PP (2, 1, 5.15), PP (2, 2, 6.52), PP (2, 3, 5.98), PP (2, e, 6.32), PP (3, 1, 5.27), PP (3, 2, 6.57), PP (3, 3, 6.01), PP (3, e, 6.33), CNP (1, 1, 4.97), CNP (1, 2, 6.72), CNP (1, 3, 6.47), CNP (1, e, 6.59), CNP (2, 1, 5.51), CNP (2, 2, 6.89), CNP (2, 3, 6.30), CNP (2, e, 6.65), CNP (3, 1, 5.65), CNP (3, 2, 6.91), CNP (3, 3, 6.34), CNP (3, e, 6.65), CC (1, 1, 4.88), CC (1, 2, 6.83), CC (1, 3, 6.12), CC (1, e, 6.45), CC (2, 1, 5.73), CC (2, 2, 7.26), CC (2, 3, 6.44), CC (2, e, 6.83), CC (3, 1, 5.72), CC (3, 2, 7.28), CC (3, 3, 6.47), CC (3, e, 6.87), CT (1, 1, 1.97), CT (1, 2, 2.46), CT (1, 3, 2.39), CT (1, e, 2.43), CT (2, 1, 2.17), CT (2, 2, 2.59), CT (2, 3, 2.53), CT (2, e, 2.58), CT (3, 1, 2.13), CT (3, 2, 2.65), CT (3, 3, 2.61), CT (3, e, 2.66), a30 (1, 1, 2.74), a30 (1, 2, 3.94), a30 (1, 3, 4.12), a30 (1, e, 3.98), a30 (2, 1, 2.97), a30 (2, 2, 3.94), a30 (2, 3, 3.94), a30 (2, e, 3.91), a30 (3, 1, 3.01), a30 (3, 2, 3.98), a30 (3, 3, 3.90), a30 (3, e, 3.90), TA (1, 2, 1.20), TA (1, 3, 1.74), TA (1, e, 1.33), TA (2, 2, 1.43), TA (2, 3, 1.96), TA (2, e, 1.54), TA (3, 2, 1.46), TA (3, 3, 2.05), TA (3, e, 1.61), FP (1, 1, 3.33), FP (1, 2, 3.64), FP (1, 3, 3.64), FP (1, e, 3.66), FP (2, 1, 3.61), FP (2, 2, 3.68), FP (2, 3, 3.65), FP (2, e, 3.69), FP (3, 1, 3.61), FP (3, 2, 3.69), FP (3, 3, 3.65), FP (3, e, 3.71), IQANK1, RHPN1, ZNF623, GPIIBP1, GLI4, CCDC166, MAPK15, FAM83H, TIGD5, MROH6, NAPRT, TOP1MT
BTA14	1,517,553–2,150,825	481,648–870,500	<i>SHARPIN, MROH1, TSSK5, GPAA1, TMEM276-ZFTRAF1, PLEC, VPS28, CPSF1, HSF1, SLC39A4, OPLAH, GRINA, SMPD5, WDR97, FBXL6, TONSL, SLC52A2, ADCK5, SCRT1, TMEM249, CYC1, HGHI, SPATC1, EXOSC4, BOP1, PARP10, DGATI, SCX, MAF1, MIR1839</i>	
BTA14	2,194,228–2,568,452	999,900–1,366,351	<i>EEF1D, PYCR3, MIR193A-2, GFUS, MAFA, SCRIB, LY6H, GSDMD, ZC3H3, ZNF623, IQANK1, RHPN1, ZNF696, GPIIBP1, GLI4, CCDC166, MAPK15, FAM83H, TIGD5, MROH6, NAPRT, TOP1MT</i>	

Continued

Table 4 (Continued). Genomic regions associated with milk yield traits and cheese-making properties in the first 3 parities in Walloon Holstein cows¹

Chromosome	Position ² (bp)	Position ³ (bp)	Gene ⁴	Trait ⁵ (parity, stage of lactation, ⁶ % variance explained)
BTA14	2,673,388–2,978,629	1,435,245–1,818,117	<i>THEM6</i> , <i>LOC112441461</i> , <i>ARC</i> , <i>LY6K</i> , <i>SLURP1</i> , <i>LYNX1</i> , <i>LOC101905222</i> , <i>ADGRRB1</i> , <i>GML</i> , <i>PSCA</i> , <i>LYPD2</i> , <i>LOC112449566</i> , <i>CYP11B1</i> , <i>LY6D</i> , <i>JRK</i>	MY (1, 1, 0.80), MY (1, 2, 1.10), MY (1, 3, 0.94), MY (1, e, 1.02), MY (2, 1, 1.04), MY (2, 2, 1.19), MY (2, 3, 0.63), MY (2, e, 0.97), MY (3, 1, 1.04), MY (3, 2, 1.19), MY (3, 3, 0.58), MY (3, e, 0.94), FP (1, 1, 3.87), FP (1, 2, 4.20), FP (1, 3, 4.18), FP (1, e, 4.22), FP (2, 1, 4.18), FP (2, 2, 4.24), FP (2, 3, 4.19), FP (2, e, 4.25), FP (3, 1, 4.15), FP (3, 2, 4.25), FP (3, 3, 4.21), FP (3, e, 4.26), PP (1, 1, 1.01), PP (1, 2, 1.27), PP (1, 3, 1.20), PP (1, e, 1.24), PP (2, 1, 1.08), PP (2, 2, 1.32), PP (2, 3, 1.16), PP (2, e, 1.26), PP (3, 1, 1.12), PP (3, 2, 1.33), PP (3, 3, 1.17), PP (3, e, 1.27), CNP (1, 1, 1.03), CNP (1, 2, 1.34), CNP (1, 3, 1.27), CNP (1, e, 1.31), CNP (2, 1, 1.16), CNP (2, 2, 1.41), CNP (2, 3, 1.25), CNP (2, e, 1.34), CNP (3, 1, 1.19), CNP (3, 2, 1.40), CNP (3, 3, 1.27), CNP (3, e, 1.34), CC (1, 1, 1.38), CC (1, 2, 1.79), CC (1, 3, 1.61), CC (1, e, 1.71), CC (2, 1, 1.62), CC (2, 2, 1.93), CC (2, 3, 1.68), CC (2, e, 1.82), CC (3, 1, 1.62), CC (3, 2, 1.92), CC (3, 3, 1.72), CC (3, e, 1.84), FP (1, 1, 1.12), FP (1, 2, 1.21), FP (1, 3, 1.23), FP (1, e, 1.23), FP (2, 1, 1.22), FP (2, 2, 1.23), FP (2, 3, 1.24), FP (2, e, 1.25), FP (3, 1, 1.22), FP (3, 2, 1.24), FP (3, 3, 1.24), FP (3, e, 1.25), TA (1, 1, 1.02), TA (1, 2, 1.01), TA (1, 3, 1.02), TA (1, e, 1.02), TA (2, 1, 1.06), TA (2, 2, 1.01), TA (2, 3, 1.01), TA (2, e, 1.03), TA (3, 1, 1.00), TA (3, 2, 1.01), TA (3, 3, 1.02), TA (3, e, 1.01)
BTA14	3,129,420–3,376,224	1,977,348–2,221,511	<i>TSNARE1</i>	
BTA20	58,872,222–58,967,597	58,812,468–58,857,983	<i>TRIO</i>	

¹The number of data were 485,218 test-day records (on 78,073 animals), 284,942 test-day records (on 48,766 animals), and 105,112 test-day records (on 21,948 animals) in the first, second, and third lactations, respectively.

²The positions of the identified genomic regions based on the UMD3.1 assembly.

³The positions of the identified genomic regions based on the ARS-UCDI.2 assembly.

⁴Genes inside the genomic region. Official gene symbol (Assembly ARS-UCDI.2, annotation release 105).

⁵MY = milk yield; FP = fat percentage; PP = protein percentage; CNP = casein percentage; CC = calcium content (mg/kg milk); SCS = somatic cell score defined as $\text{SCS} = \log_2(\text{SCC}/100,000) + 3$; CT = coagulation time defined as the sum of the rennet coagulation time plus the time to a curd firmness of 20 mm measured by the Computerized Renneting Meter; a30 = curd firmness defined as the curd firmness measured 30 min after enzyme addition by the Computerized Renneting Meter; and TA = milk titratable acidity measured in Dornic degrees (°D).

⁶The GWAS analyses were performed for all considered traits in each parity, considering 4 stages of lactation: (1) from 5 to 60 DIM, representing the ascending production stage and lactation peak; (2) from 61 to 200 DIM, representing the lactation persistency stage; (3) from 201 to 365 DIM, representing the production decline up to the end of the lactation, and (e) from 5 to 365 DIM representing the entire lactation.

in fat metabolism or mammary gland development in Holstein cows (Wang et al., 2022).

BTA14

Four genomic regions located from 1.52 to 2.15 MB, 2.19 to 2.57 MB, 2.67 to 2.98 MB, and 3.13 to 3.38 MB (UMD3.1 assembly) on BTA14 were identified to be associated with one or more of the included traits. Hereafter, these regions are identified as BTA14-I, BTA14-II, BTA14-III, and BTA14-IV. These 4 regions combined explained 23.82 to 26.92% of the total genetic variance of FP. Genomic regions of BTA14-I, BTA14-II, and BTA14-III combined explained 7.5 to 10.87%, 7.09 to 9.73%, and 6.72 to 9.29% of the total genetic variances of CC, CNP, and PP, respectively. Genomic regions of BTA14-I and BTA14-III combined explained 3.27 to 5.08% of the total genetic variances of MY. The association between this region (1.52 to 3.38 MB on BTA14) and milk production traits has been reported by previous studies (Pausch et al., 2015; Jiang et al., 2019; Pedrosa et al., 2021). Jung et al. (2019) reported that the region located from 1.19 to 3.70 MB (UMD3.1 assembly) on BTA14 explained the highest proportions of genetic variance for FP and milk FA profile in the Brazilian Holstein population. The following are the results discussed by regions identified on BTA14.

BTA14-I. The genomic region located from 1.52 to 2.15 MB on BTA14 was associated with all studied traits except for SCS. This region explained 2.40 to 3.89%, 15.51 to 17.75%, 4.66 to 6.57%, 4.97 to 6.91%, 4.88 to 7.28%, 1.97 to 2.66%, 2.74 to 4.12%, and 1.20 to 2.05% of the total additive genetic variance of MY, FP, PP, CNP, CC, CT, a30, and TA, respectively.

The results showed moderate negative genetic correlated between CT and milk composition traits, whereas high positive genetic correlations were estimated between milk composition traits and a30 and TA. Therefore, these traits are correlated and can be affected by the same genes. Chen et al. (2023) reported that this region is associated with nitrogen efficiency index and milk true protein nitrogen in Walloon Holstein cows. Clancey et al. (2019) reported that SNP inside this region are associated with MY in Holstein cows. This region was 633.27 KB in size and harbors 30 genes including the scratch family zinc finger 1 (*SCR1*), diacylglycerol O-acyltransferase 1 (*DGAT1*), cleavage and polyadenylation specific factor 1 (*CPSF1*), tonsoku like, DNA repair protein (*TONSL*), and spermatogenesis and centriole associated 1 (*SPATC1*). Most of the positional candidate genes found in this region support results from previous studies for MY and milk composition in dairy cattle (Kolbehdari et al., 2009; Li et al., 2010; Capomaccio et al., 2015; Jiang et al., 2019).

DGAT1, involved in the last step of the synthesis of triacylglycerol, has a major effect on milk production traits (Jiang et al., 2010; Maxa et al., 2012; Nayeri et al., 2016; Clancey et al., 2019; Cruz et al., 2019). Nayeri et al. (2016) reported that SNP located within *CPSF1* and *TONSL* are associated with MY in Canadian dairy Holstein cattle. The SHANK associated RH domain interactor (*SHARPIN*) gene product is inside this region. This product of this gene involved in the regulation of immune and inflammatory responses (Wang et al., 2012) and has been reported to be associated with the colostrum and serum albumin concentrations in Holstein cows (Lin et al., 2020). Sanchez et al. (2017) reported *DGAT1*, maestro heat like repeat family member 1 (*MROH1*), and BOP1 ribosomal biogenesis factor (*BOP1*) as the most important genes explaining the majority of the variability of milk protein composition in Montbéliarde, Normande, and Holstein dairy cattle.

BTA14-II. Furthermore, the region located from 2.19 to 2.57 MB on BTA14 was associated with milk composition traits including FP, PP, CC, and CNP. This region explained 3.33 to 3.71%, 1.08 to 1.38%, 1.09 to 1.42%, and 1.24 to 1.67% of the total additive genetic variance of FP, PP, CNP, and CC, respectively. Milk composition traits including FP, PP, CNP, and CC are strongly correlated and are expected to be influenced by the same genomic regions. The genetic correlations estimated among FP, PP, CNP, and CC ranged from 0.56 (FP and CC) to 0.97 (PP and CNP), 0.61 (FP and CC) to 0.97 (PP and CNP), and 0.58 (FP and CC) to 0.97 (PP and CNP) for the first, second, and third lactations, respectively. Atashi et al. (2020) reported that the region located from 1.86 to 2.12 MB on BTA14 is associated with 305-d MY and the lactation curve parameters in Holstein dairy cows. This region was 374.22 KB in size and harbors 22 genes. Among genes inside this region, nicotinate phosphoribosyltransferase (*NAPRT*), family with sequence similarity 83 member H (*FAM83H*), eukaryotic translation elongation factor 1 delta (*EEF1D*), pyrroline-5-carboxylate reductase 3 (*PYCR3*), scribbled planar cell polarity protein (*SCRIB*), lymphocyte antigen 6 family member H (*LY6H*), and mitogen-activated protein kinase 15 (*MAPK15*) have been previously reported as candidate genes for milk production traits in Holstein cows (Li et al., 2010; Buitenhuis et al., 2014; Ning et al., 2017; Wang et al., 2019). *MAPK15* may affect milk composition through downregulating transactivation of the glucocorticoid receptor, as glucocorticoid is an important hormone in maintaining milking (Saelzler et al., 2006). The pyrroline-5-carboxylate reductase 3 (*PYCR3*) gene is located inside this region. This gene encodes a protein that belongs to the pyrroline-5-carboxylate reductase family of enzymes that responds to inflammatory,

nutrient, and oxidative stress (Kuo et al., 2016) and has been reported to be associated with colostrum and serum albumin concentrations in Holstein cows (Lin et al., 2020).

BTA14-III. Additional region explaining large proportions of the genetic variance of MY and milk composition (CC, CNP, FP, PP) were found between 2.67 to 2.98 MB on BTA14, where 15 genes are located. Milk composition traits are strongly correlated and are expected to be affected by same polytrophic genes. This region contains lymphocyte-antigen-6 complex (*LY6*) including *LY6K*, *SLURP1*, *LYNX1*, *PSCA*, *LYPD2*, and *LY6D* in regulating the major histocompatibility complex. Tiezzi et al. (2015) reported that this region is associated with clinical mastitis in US Holsteins. Atashi et al. (2020) reported that this region is highly associated with 305-d MY and peak yield in Holstein cows. Among genes inside this region, adhesion G protein-coupled receptor B1 (*ADGRB1*), glycosylphosphatidylinositol anchored molecule like (*GML*), and lymphocyte antigen 6 family member D (*LY6D*) have been reported as candidate genes for MY, FP, PP, fat yield (**FY**), protein yield (**PY**), and milk FA profile in Holstein cows (Cole et al., 2011; Buitenhuis et al., 2014; Ning et al., 2017; Jiang et al., 2019). Costa et al. (2019) reported that the variant within *ADGRB1* accounted for 2.44% of the total additive genetic variance for lactose yield in Fleckvieh cattle. The *LY6D* gene has been reported to be associated with MY and FY in Holstein cows (Jiang et al., 2014; Suchocki et al., 2016). The cytochrome P450 family 11 subfamily B member 1 (*CYP11B1*) gene is involved in glucose and lipid metabolism and has been reported as a functional candidate gene for milk production in dairy cows (Bülöw and Bernhardt, 2002; Kaupe et al., 2007).

BTA14-IV. The genomic region located from 3.13 to 3.38 MB on BTA14, where the t-SNARE domain containing 1 (*TSNARE1*) gene is located, was associated with FP. This region was 246.80 KB in size and explained 1.12 to 1.25% of the total additive genetic variance of FP. *TSNARE1* plays a role in intracellular protein transport and synaptic vesicle exocytosis (Smith et al., 2012; Luo et al., 2021) and has been reported as a candidate gene for traits including MY, FY, FP, PY, PP, and milk FA profile in Holstein cows (Buitenhuis et al., 2014; Jiang et al., 2019; Freitas et al., 2020; Bohlouli et al., 2022). Luo et al. (2021) evaluated physiological indicators of heat stress in Holstein dairy cows and reported a strong association of *TSNARE1* with rectal temperature.

BTA20

On BTA20, the window located from 58.87 to 58.97 MB (UMD3.1 assembly), where the trio Rho guanine nucleotide exchange factor (*TRIO*) gene is located, was associated with TA. *TRIO* encodes a large protein that functions as a guanosine diphosphate (GDP) to guanosine triphosphate (GTP) exchange factor (Bateman et al., 2000). It has also been shown that the *TRIO* gene is associated with FY and milk BHB concentration in Holstein dairy cattle (Nayeri et al., 2019). The SNP inside this region has been reported to be associated with MY, FP, PP, FY, PY, SCS, and udder morphology traits in Holstein cows (Bennewitz et al., 2004; Schrooten et al., 2004; Höglund et al., 2009; Cole et al., 2011; Wang et al., 2022), and weaning weight and MY in beef cattle (Michenet et al., 2016).

CONCLUSIONS

Milk's CMP are among important breeding traits in dairy cattle breeds, especially in modern animal husbandry environments. This study aimed to estimate genetic parameters and to identify genomic regions associated with milk's cheese-making traits in Walloon Holstein cows. The findings showed that the included CMP are moderately heritable and could be included into the breeding program currently used for Walloon Holstein cows. Using available genotyping data, additional analyses were done to identify genomic regions associated with milk's cheese-making traits. Different milk production and cheese-making traits were associated with 6 genomic regions distributed over 3 chromosomes (BTA1, BTA14, and BTA20), which could further be used for genomic prediction purposes. The results confirmed most previously identified genes for milk production traits and identified several novel candidate genes (including *SLC37A1*, *TRIO*, and genes inside located from 1.52 to 2.15 MB on BTA14) for the studied cheese-making traits including CT, a30, and TA. Future research based on gene enrichment analysis might complement the GWAS results and help to deepen the understanding of the biological pathways related to the studied milk's cheese-making traits.

ACKNOWLEDGMENTS

H. Atashi acknowledges the support of the Walloon Government (Service Public de Wallonie, Direction Générale Opérationnelle Agriculture, Ressources Naturelles et Environnement, SPW-DGARNE, Namur, Belgium) for its financial support facilitating his stay in Belgium through the WALLSmart Project (D31-1392 and D65-1435). H. Wilmot, as a current research fellow,

and N. Gengler, as a former senior research associate, acknowledge the support of the Fonds de la Recherche Scientifique (FNRS; Brussels, Belgium). The authors thank the INTERREG VA France-Wallonie-Vlaanderen program and the Walloon Government (Service Public de Wallonie, Direction Générale Opérationnelle Agriculture, Ressources Naturelles et Environnement, SPW-DGARNE, Namur, Belgium) for their financial support through the BlueStar project and previous projects. The authors also acknowledge the technical support by the Walloon Breeders Association (elevéo). The University of Liège–Gembloux Agro-Bio Tech (Liège, Belgium) supported computations through the technical platform Calcul et Modélisation Informatique (CAMI) of the TERRA Teaching and Research Centre, partly supported by the Fonds de la Recherche Scientifique (FNRS) under grant no. T.0095.19 (PDR “DEEPSE-LECT”). Genotyping was facilitated through the support of the Fonds de la Recherche Scientifique (FNRS) under grant no. J.0174.18 (CDR “PREDICT-2”). Relevant information supporting the results not presented here is provided in supplemental data (<https://github.com/hadiatashi/Holstein-Cheese>). None of the data were deposited in an official repository because they are the property of the breeding organizations, and they are available upon reasonable request. Author contributions are as follows: Hadi Atashi: conceptualization, formal analysis, investigation, methodology, writing the original draft; Yansen Chen: reviewing and editing the draft; Hélène Wilmot: reviewing and editing the draft; Catherine Bastin: development of original trait definitions, reviewing and editing the draft; Sylvie Vanderick: Reviewing and editing the draft; Xavier Hubin: data curation, reviewing and editing the draft; and Nicolas Gengler: conceptualization, funding acquisition, project administration, resources, supervision, and reviewing and editing the draft. The authors have not stated any conflicts of interest.

REFERENCES

- Aguilar, I., A. Legarra, F. Cardoso, Y. Masuda, D. Lourenco, and I. Misztal. 2019. Frequentist p-values for large-scale-single step genome-wide association, with an application to birth weight in American Angus cattle. *Genet. Sel. Evol.* 51:28. <https://doi.org/10.1186/s12711-019-0469-3>.
- Aguilar, I., I. Misztal, D. Johnson, A. Legarra, S. Tsuruta, and T. Lawlor. 2010. Hot topic: a unified approach to utilize phenotypic, full pedigree, and genomic information for genetic evaluation of Holstein final score. *J. Dairy Sci.* 93:743–752. <https://doi.org/10.3168/jds.2009-2730>.
- Aguilar, I., I. Misztal, S. Tsuruta, A. Legarra, and H. Wang. 2014. PREGSF90–POSTGSF90: Computational tools for the implementation of single-step genomic selection and genome-wide association with ungenotyped individuals in BLUPF90 program. Paper presented at the Proceedings of the 10th World Congress on Genetics Applied to Livestock Production, Vancouver, BC, Canada.
- Aguilar, I., S. Tsuruta, Y. Masuda, D. Lourenco, A. Legarra, and I. Misztal. 2018. BLUPF90 suite of programs for animal breeding with focus on genomics. Paper presented at the Proceedings of the 11th World Congress on Genetics Applied to Livestock Production, Auckland, New Zealand.
- Atashi, H., C. Bastin, H. Wilmot, S. Vanderick, X. Hubin, and N. Gengler. 2022a. Genome-wide association study for selected cheese-making properties in Dual-Purpose Belgian Blue cows. *J. Dairy Sci.* 105:8972–8988. <https://doi.org/10.3168/jds.2022-21780>.
- Atashi, H., and M. Hostens. 2021. Genetic aspects of somatic cell count in Holstein dairy cows in Iran. *Animals (Basel)* 11:1637. <https://doi.org/10.3390/ani11061637>.
- Atashi, H., M. Salavati, J. De Koster, J. Ehrlich, M. Crowe, G. Opsomer, M. Hostens, N. McLoughlin, A. Fahey, E. Matthews, A. Santoro, C. Byrne, P. Rudd, R. O’Flaherty, S. Hallinan, C. Wathes, Z. Cheng, A. Fouladi, G. Pollott, D. Werling, B. Sanz Bernardo, A. Wylie, M. Bell, M. Vaneetvelde, K. Hermans, G. Opsomer, S. Moerman, J. Koster, H. Bogaert, J. Vandepitte, L. Vandeveld, B. Vanranst, J. Hoglund, S. Dahl, S. Ostergaard, J. Rothmann, M. Krogh, E. Meyer, C. Gaillard, J. Ettema, T. Rousing, F. Signorelli, F. Napolitano, B. Muioli, A. Crisà, L. Buttazzoni, J. McClure, D. Matthews, F. Kearney, A. Cromie, M. McClure, S. Zhang, X. Chen, H. Chen, J. Zhao, L. Yang, G. Hua, C. Tan, G. Wang, M. Bonneau, A. Pompozzi, A. Pearn, A. Evertson, L. Kosten, A. Fogh, T. Andersen, M. Lucey, C. Elsis, G. Conant, J. Taylor, N. Gengler, M. Georges, F. Colinet, M. Ramos Pamplona, H. Hammami, C. Bastin, H. Takeda, A. Laine, A.-S. Van Laere, M. Schulze, S. Palma Vera, C. Ferris, and C. Marchitelli. 2020. Genome-wide association for milk production and lactation curve parameters in Holstein dairy cows. *J. Anim. Breed. Genet.* 137:292–304. <https://doi.org/10.1111/jbg.12442>.
- Atashi, H., H. Wilmot, S. Vanderick, X. Hubin, and N. Gengler. 2022b. Genome-wide association study for milk production traits in Dual-Purpose Belgian Blue cows. *Livest. Sci.* 256:104831. <https://doi.org/10.1016/j.livsci.2022.104831>.
- Bastin, C., H. Soyeurt, and N. Gengler. 2013. Genetic parameters of milk production traits and fatty acid contents in milk for Holstein cows in parity 1–3. *J. Anim. Breed. Genet.* 130:118–127. <https://doi.org/10.1111/jbg.12010>.
- Bateman, J., H. Shu, and D. Van Vactor. 2000. The guanine nucleotide exchange factor trio mediates axonal development in the *Drosophila* embryo. *Neuron* 26:93–106. [https://doi.org/10.1016/S0896-6273\(00\)81141-1](https://doi.org/10.1016/S0896-6273(00)81141-1).
- Beissinger, T. M., G. J. Rosa, S. M. Kaeppeler, D. Gianola, and N. de Leon. 2015. Defining window-boundaries for genomic analyses using smoothing spline techniques. *Genet. Sel. Evol.* 47:30. <https://doi.org/10.1186/s12711-015-0105-9>.
- Bencini, R. 2002. Factors affecting the clotting properties of sheep milk. *J. Sci. Food Agric.* 82:705–719. <https://doi.org/10.1002/jsfa.1101>.
- Bennewitz, J., N. Reinsch, V. Guiard, S. Fritz, H. Thomsen, C. Looft, C. Kühn, M. Schwerin, C. Weimann, G. Erhardt, F. Reinhardt, R. Reents, D. Boichard, and E. Kalm. 2004. Multiple quantitative trait loci mapping with cofactors and application of alternative variants of the false discovery rate in an enlarged granddaughter design. *Genetics* 168:1019–1027. <https://doi.org/10.1534/genetics.104.030296>.
- Bittante, G. 2011. Modeling rennet coagulation time and curd firmness of milk. *J. Dairy Sci.* 94:5821–5832. <https://doi.org/10.3168/jds.2011-4514>.
- Bittante, G., M. Penasa, and A. Cecchinato. 2012. Invited review: Genetics and modeling of milk coagulation properties. *J. Dairy Sci.* 95:6843–6870. <https://doi.org/10.3168/jds.2012-5507>.
- Bobbo, T., C. Cipolat-Gotet, G. Bittante, and A. Cecchinato. 2016. The nonlinear effect of somatic cell count on milk composition, coagulation properties, curd firmness modeling, cheese yield, and curd nutrient recovery. *J. Dairy Sci.* 99:5104–5119. <https://doi.org/10.3168/jds.2015-10512>.
- Bohlouli, M., K. Halli, T. Yin, N. Gengler, and S. König. 2022. Genome-wide associations for heat stress response suggest potential candidate genes underlying milk fatty acid composition in dairy

- cattle. *J. Dairy Sci.* 105:3323–3340. <https://doi.org/10.3168/jds.2021-21152>.
- Buitenhuys, B., L. L. Janss, N. A. Poulsen, L. B. Larsen, M. K. Larsen, and P. Sørensen. 2014. Genome-wide association and biological pathway analysis for milk-fat composition in Danish Holstein and Danish Jersey cattle. *BMC Genomics* 15:1112. <https://doi.org/10.1186/1471-2164-15-1112>.
- Bülöw, H. E., and R. Bernhardt. 2002. Analyses of the *CYP11B* gene family in the guinea pig suggest the existence of a primordial *CYP11B* gene with aldosterone synthase activity. *Eur. J. Biochem.* 269:3838–3846. <https://doi.org/10.1046/j.1432-1033.2002.03076.x>.
- Capomaccio, S., M. Milanese, L. Bomba, K. Cappelli, E. L. Nicolazzi, J. L. Williams, P. Ajmone-Marsan, and B. Stefanon. 2015. Searching new signals for production traits through gene-based association analysis in three Italian cattle breeds. *Anim. Genet.* 46:361–370. <https://doi.org/10.1111/age.12303>.
- 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. <https://doi.org/10.3168/jds.2007-0308>.
- Cecchinato, A., M. Penasa, M. De Marchi, L. Gallo, G. Bittante, and P. Carnier. 2011. Genetic parameters of coagulation properties, milk yield, quality, and acidity estimated using coagulating and noncoagulating milk information in Brown Swiss and Holstein-Friesian cows. *J. Dairy Sci.* 94:4205–4213. <https://doi.org/10.3168/jds.2010-3913>.
- Chamberlain, A. J., C. J. Vander Jagt, B. J. Hayes, M. Khansefid, L. C. Marett, C. A. Millen, T. T. Nguyen, and M. E. Goddard. 2015. Extensive variation between tissues in allele specific expression in an outbred mammal. *BMC Genomics* 16:993. <https://doi.org/10.1186/s12864-015-2174-0>.
- Chen, Y., H. Atashi, S. Vanderick, R. Mota, H. Soyeurt, H. Hammami, and N. Gengler. 2023. Genome-wide association study and functional annotation analyses for nitrogen efficiency index and its composition traits in dairy cattle. *J. Dairy Sci.* 106:3397–3410. <https://doi.org/10.3168/jds.2022-22351>.
- Clancey, E., J. N. Kiser, J. G. Moraes, J. Dalton, T. E. Spencer, and H. L. Neibergs. 2019. Genome-wide association analysis and gene set enrichment analysis with SNP data identify genes associated with 305-day milk yield in Holstein dairy cows. *Anim. Genet.* 50:254–258. <https://doi.org/10.1111/age.12792>.
- Cole, J. B., G. R. Wiggans, L. Ma, T. S. Sonstegard, T. J. Lawlor Jr., B. A. Crooker, C. P. Van Tassel, J. Yang, S. Wang, L. K. Matukumali, and Y. Da. 2011. Genome-wide association analysis of thirty one production, health, reproduction and body conformation traits in contemporary US Holstein cows. *BMC Genomics* 12:408. <https://doi.org/10.1186/1471-2164-12-408>.
- Colinet, F., H. Soyeurt, C. Anceau, A. Vanlierde, N. Keyen, P. Dardenne, N. Gengler, and M. Sindic. 2010. Potential estimation of titratable acidity in cow milk using mid-infrared spectrometry. *ICAR Tech. Ser.* 14:239–243.
- Colinet, F., T. Troch, V. Baeten, F. Dehareng, P. Dardenne, M. Sindic, and N. Gengler. 2015. Genetic variability of MIR predicted milk technological properties in Walloon dairy cattle. Book of Abstracts of the 66th Annual Meeting of the European Association for Animal Production. Wageningen Academic Publishers. <https://doi.org/10.3920/978-90-8686-816-2>.
- Colinet, F., T. Troch, B. Vanden, H. Soyeurt, O. Abbas, V. Baeten, F. Dehareng, E. Froidmont, G. Sinnaeve, and P. Dardenne. 2013. Mid-infrared prediction of cheese yield from milk and its genetic variability in first-parity cows. Book of Abstracts of the 64th Annual Meeting of the European Association for Animal Production. Wageningen Academic Publishers. <https://doi.org/10.3920/978-90-8686-782-0>.
- Colinet, F., A. Vanlierde, M. Sindic, F. Dehareng, G. Sinnaeve, J. Vandendplas, H. Soyeurt, C. Bastin, and N. Gengler. 2012. Mid-infrared prediction of milk titratable acidity and its genetic variability in first-parity cows. Book of Abstracts of the 63rd Annual Meeting of the European Association for Animal Production. Wageningen Academic Publishers.
- Costa, A., H. Schwarzenbacher, G. Mészáros, B. Fuerst-Waltl, C. Fuerst, J. Sölkner, and M. Penasa. 2019. On the genomic regions associated with milk lactose in Fleckvieh cattle. *J. Dairy Sci.* 102:10088–10099. <https://doi.org/10.3168/jds.2019-16663>.
- Cruz, V. A., H. R. Oliveira, L. F. Brito, A. Fleming, S. Larmer, F. Miglior, and F. S. Schenkel. 2019. Genome-wide association study for milk fatty acids in Holstein cattle accounting for the *DGAT1* gene effect. *Animals (Basel)* 9:997. <https://doi.org/10.3390/ani9110997>.
- De Marchi, M., R. Dal Zotto, M. Cassandro, and G. Bittante. 2007. Milk coagulation ability of five dairy cattle breeds. *J. Dairy Sci.* 90:3986–3992. <https://doi.org/10.3168/jds.2006-627>.
- De Marchi, M., C. C. Fagan, C. O'donnell, A. Cecchinato, R. Dal Zotto, M. Cassandro, M. Penasa, and G. Bittante. 2009. Prediction of coagulation properties, titratable acidity, and pH of bovine milk using mid-infrared spectroscopy. *J. Dairy Sci.* 92:423–432. <https://doi.org/10.3168/jds.2008-1163>.
- De Marchi, M., V. Toffanin, M. Cassandro, and M. Penasa. 2014. Invited review: Mid-infrared spectroscopy as phenotyping tool for milk traits. *J. Dairy Sci.* 97:1171–1186. <https://doi.org/10.3168/jds.2013-6799>.
- Fang, Z.-H., H. Bovenhuis, H. van Valenberg, P. Martin, S. Duchemin, T. Huppertz, and M. Visker. 2019. Genome-wide association study for α_{S1} - and α_{S2} -casein phosphorylation in Dutch Holstein Friesian. *J. Dairy Sci.* 102:1374–1385. <https://doi.org/10.3168/jds.2018-15593>.
- Food and Agriculture Organization of the United Nations. 2015. FAO-STAT. Accessed Mar. 27, 2015. <http://www.milb.com/index.jsp?sid=564>.
- Freitas, P. H., H. Oliveira, F. Silva, A. Fleming, F. Miglior, F. Schenkel, and L. Brito. 2020. Genomic analyses for predicted milk fatty acid composition throughout lactation in North American Holstein cattle. *J. Dairy Sci.* 103:6318–6331. <https://doi.org/10.3168/jds.2019-17628>.
- Habier, D., R. L. Fernando, K. Kizilkaya, and D. J. Garrick. 2011. Extension of the Bayesian alphabet for genomic selection. *BMC Bioinformatics* 12:186. <https://doi.org/10.1186/1471-2105-12-186>.
- Han, Y., and F. Peñagaricano. 2016. Unravelling the genomic architecture of bull fertility in Holstein cattle. *BMC Genet.* 17:143. <https://doi.org/10.1186/s12863-016-0454-6>.
- Höglund, J. K., A. Buitenhuys, B. Guldbbrandtsen, G. Su, B. Thomsen, and M. Lund. 2009. Overlapping chromosomal regions for fertility traits and production traits in the Danish Holstein population. *J. Dairy Sci.* 92:5712–5719. <https://doi.org/10.3168/jds.2008-1964>.
- ICAR. 2022. Procedure 2 of Section 2 of ICAR Guidelines - Computing of Accumulated Lactation Yield. Accessed Dec. 29, 2022. <https://www.icar.org/Guidelines/02-Procedure-2-Computing-Lactation-Yield.pdf>.
- Ikonen, T., S. Morri, A.-M. Tyrisevä, O. Ruottinen, and M. Ojala. 2004. Genetic and phenotypic correlations between milk coagulation properties, milk production traits, somatic cell count, casein content, and pH of milk. *J. Dairy Sci.* 87:458–467. [https://doi.org/10.3168/jds.S0022-0302\(04\)73185-9](https://doi.org/10.3168/jds.S0022-0302(04)73185-9).
- Iung, L. H., J. Petrini, J. Ramírez-Díaz, M. Salvian, G. Rovadoski, F. Pilonetto, B. Dauria, P. Machado, L. Coutinho, G. Wiggans, and G. B. Mourão. 2019. Genome-wide association study for milk production traits in a Brazilian Holstein population. *J. Dairy Sci.* 102:5305–5314. <https://doi.org/10.3168/jds.2018-14811>.
- Jamrozik, J., and L. Schaeffer. 1997. Estimates of genetic parameters for a test day model with random regressions for yield traits of first lactation Holsteins. *J. Dairy Sci.* 80:762–770. [https://doi.org/10.3168/jds.S0022-0302\(97\)75996-4](https://doi.org/10.3168/jds.S0022-0302(97)75996-4).
- Jiang, J., L. Ma, D. Prakapenka, P. M. VanRaden, J. B. Cole, and Y. Da. 2019. A large-scale genome-wide association study in US Holstein cattle. *Front. Genet.* 10:412. <https://doi.org/10.3389/fgene.2019.00412>.
- Jiang, L., J. Liu, D. Sun, P. Ma, X. Ding, Y. Yu, and Q. Zhang. 2010. Genome wide association studies for milk production traits in Chinese Holstein population. *PLoS One* 5:e13661. <https://doi.org/10.1371/journal.pone.0013661>.
- Jiang, L., X. Liu, J. Yang, H. Wang, J. Jiang, L. Liu, S. He, X. Ding, J. Liu, and Q. Zhang. 2014. Targeted resequencing of GWAS loci

- reveals novel genetic variants for milk production traits. *BMC Genomics* 15:1105. <https://doi.org/10.1186/1471-2164-15-1105>.
- Kaupe, B., H. Brandt, E. Prinzenberg, and G. Erhardt. 2007. Joint analysis of the influence of CYP11B1 and DGAT1 genetic variation on milk production, somatic cell score, conformation, reproduction, and productive lifespan in German Holstein cattle. *J. Anim. Sci.* 85:11–21. <https://doi.org/10.2527/jas.2005-753>.
- Kemper, K. E., C. M. Reich, P. J. Bowman, C. J. Vander Jagt, A. J. Chamberlain, B. A. Mason, B. J. Hayes, and M. E. Goddard. 2015. Improved precision of QTL mapping using a nonlinear Bayesian method in a multi-breed population leads to greater accuracy of across-breed genomic predictions. *Genet. Sel. Evol.* 47:29. <https://doi.org/10.1186/s12711-014-0074-4>.
- Kolbehdari, D., Z. Wang, J. Grant, B. Murdoch, A. Prasad, Z. Xiu, E. Marques, P. Stothard, and 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. <https://doi.org/10.1111/j.1439-0388.2008.00793.x>.
- Kuo, M.-L., M. B.-E. Lee, M. Tang, W. Den Besten, S. Hu, M. J. Sweredoski, S. Hess, C.-M. Chou, C. A. Changou, M. Su, W. Jia, L. Su, and Y. Yen. 2016. PYCR1 and PYCR2 interact and collaborate with RRM2B to protect cells from overt oxidative stress. *Sci. Rep.* 6:18846. <https://doi.org/10.1038/srep18846>.
- Li, H., Z. Wang, S. Moore, F. Schenkel, and P. Stothard. 2010. Genome-wide scan for positional and functional candidate genes affecting milk production traits in Canadian Holstein cattle. Paper presented at the Proceedings of the 9th World Congress on Genetics Applied to Livestock Production, Vancouver, BC Canada. Leipzig, Germany.
- Lin, S., Z. Wan, J. Zhang, L. Xu, B. Han, and D. Sun. 2020. Genome-wide association studies for the concentration of albumin in colostrum and serum in Chinese Holstein. *Animals (Basel)* 10:2211. <https://doi.org/10.3390/ani10122211>.
- Luo, H., X. Li, L. Hu, W. Xu, Q. Chu, A. Liu, G. Guo, L. Liu, L. F. Brito, and Y. Wang. 2021. Genomic analyses and biological validation of candidate genes for rectal temperature as an indicator of heat stress in Holstein cattle. *J. Dairy Sci.* 104:4441–4451. <https://doi.org/10.3168/jds.2020-18725>.
- Maxa, J., M. Neuditschko, I. Russ, M. Förster, and I. Medugorac. 2012. Genome-wide association mapping of milk production traits in Braunvieh cattle. *J. Dairy Sci.* 95:5357–5364. <https://doi.org/10.3168/jds.2011-4673>.
- Medeiros de Oliveira Silva, R., N. Bonvino Stafuzza, B. de Oliveira Fragomeni, G. Miguel Ferreira de Camargo, T. Matos Ceacero, J. Noely dos Santos Gonçalves Cyrillo, F. Baldi, A. Augusti Boligon, M. E. Zerlotti Mercadante, D. Lino Lourenco, I. Misztal, and L. Galvão de Albuquerque. 2017. Genome-wide association study for carcass traits in an experimental Nelore cattle population. *PLoS One* 12:e0169860. <https://doi.org/10.1371/journal.pone.0169860>.
- Michenet, A., M. Barbat, R. Saintilan, E. Venot, and F. Phocas. 2016. Detection of quantitative trait loci for maternal traits using high-density genotypes of Blonde d'Aquitaine beef cattle. *BMC Genet.* 17:88. <https://doi.org/10.1186/s12863-016-0397-y>.
- Nayeri, S., M. Sargolzaei, M. K. Abo-Ismael, N. May, S. P. Miller, F. Schenkel, S. S. Moore, and P. Stothard. 2016. Genome-wide association for milk production and female fertility traits in Canadian dairy Holstein cattle. *BMC Genet.* 17:75. <https://doi.org/10.1186/s12863-016-0386-1>.
- Nayeri, S., F. Schenkel, A. Fleming, V. Kroezen, M. Sargolzaei, C. Baes, A. Cánovas, J. Squires, and F. Miglior. 2019. Genome-wide association analysis for β -hydroxybutyrate concentration in milk in Holstein dairy cattle. *BMC Genet.* 20:58. <https://doi.org/10.1186/s12863-019-0761-9>.
- Nilsson, K., H. Stålhammar, M. S. Hansen, H. Lindmark-Månsson, S. Duchemin, F. Fikse, D. J. de Koning, M. Paulsson, and M. Glantz. 2019. Characterisation of non-coagulating milk and effects of milk composition and physical properties on rennet-induced coagulation in Swedish Red dairy cattle. *Int. Dairy J.* 95:50–57. <https://doi.org/10.1016/j.idairyj.2019.03.006>.
- Ning, C., H. Kang, L. Zhou, D. Wang, H. Wang, A. Wang, J. Fu, S. Zhang, and J. Liu. 2017. Performance gains in genome-wide association studies for longitudinal traits via modeling time-varied effects. *Sci. Rep.* 7:590. <https://doi.org/10.1038/s41598-017-00638-2>.
- Oliveira, H. R., J. Cant, L. Brito, F. Feitosa, T. Chud, P. Fonseca, J. Jamrozik, F. Silva, D. Lourenco, and F. Schenkel. 2019. Genome-wide association for milk production traits and somatic cell score in different lactation stages of Ayrshire, Holstein, and Jersey dairy cattle. *J. Dairy Sci.* 102:8159–8174. <https://doi.org/10.3168/jds.2019-16451>.
- Paiva, J. T., R. R. Mota, P. S. Lopes, H. Hammami, S. Vanderick, H. R. Oliveira, R. Veroneze, F. Fonseca e Silva, and N. Gengler. 2022. Random regression test-day models to describe milk production and fatty acid traits in first lactation Walloon Holstein cows. *J. Anim. Breed. Genet.* 139:398–413. <https://doi.org/10.1111/jbg.12673>.
- Pan, C.-J., S.-Y. Chen, H. S. Jun, S. R. Lin, B. C. Mansfield, and J. Y. Chou. 2011. SLC37A1 and SLC37A2 are phosphate-linked, glucose-6-phosphate antiporters. *PLoS One* 6:e23157. <https://doi.org/10.1371/journal.pone.0023157>.
- Pastorino, A. J., C. Hansen, and D. J. McMahon. 2003. Effect of pH on the chemical composition and structure-function relationships of Cheddar cheese. *J. Dairy Sci.* 86:2751–2760. [https://doi.org/10.3168/jds.S0022-0302\(03\)73871-5](https://doi.org/10.3168/jds.S0022-0302(03)73871-5).
- Pausch, H., R. Emmerling, B. Gredler-Grandl, R. Fries, H. D. Daetwyler, and M. E. Goddard. 2017. Meta-analysis of sequence-based association studies across three cattle breeds reveals 25 QTL for fat and protein percentages in milk at nucleotide resolution. *BMC Genomics* 18:853. <https://doi.org/10.1186/s12864-017-4263-8>.
- Pausch, H., C. Wurmser, F. Reinhardt, R. Emmerling, and R. Fries. 2015. Validation of 4 candidate causative trait variants in 2 cattle breeds using targeted sequence imputation. *J. Dairy Sci.* 98:4162–4167. <https://doi.org/10.3168/jds.2015-9402>.
- Pedrosa, V. B., F. S. Schenkel, S.-Y. Chen, H. R. Oliveira, T. M. Casey, M. G. Melka, and L. F. Brito. 2021. Genome-wide association analyses of lactation persistency and milk production traits in Holstein Cattle based on imputed whole-genome sequence data. *Genes (Basel)* 12:1830. <https://doi.org/10.3390/genes12111830>.
- Pretto, D., T. Kaart, M. Vallas, I. Jõudu, M. Henno, L. Ancilotto, M. Cassandro, and E. Pärna. 2011. Relationships between milk coagulation property traits analyzed with different methodologies. *J. Dairy Sci.* 94:4336–4346. <https://doi.org/10.3168/jds.2011-4267>.
- Raven, L.-A., B. G. Cocks, K. E. Kemper, A. J. Chamberlain, C. J. Vander Jagt, M. E. Goddard, and B. J. Hayes. 2016. Targeted imputation of sequence variants and gene expression profiling identifies twelve candidate genes associated with lactation volume, composition and calving interval in dairy cattle. *Mamm. Genome* 27:81–97. <https://doi.org/10.1007/s00335-015-9613-8>.
- Rosen, B. D., D. M. Bickhart, R. D. Schnabel, S. Koren, C. G. Elsik, E. Tseng, T. N. Rowan, W. Y. Low, A. Zimin, C. Coudrey, R. Hall, W. Li, A. Rhie, J. Ghurye, S. D. McKay, F. Thibaud-Nissen, J. Hoffman, B. M. Murdoch, W. M. Snelling, T. G. McDaniel, J. A. Hammond, J. C. Schwartz, W. Nandolo, D. E. Hagen, C. Dreischer, S. J. Schultheiss, S. G. Schroeder, A. M. Phillippy, J. B. Cole, C. P. Van Tassel, G. Liu, T. P. L. Smith, and J. F. Medrano. 2020. De novo assembly of the cattle reference genome with single-molecule sequencing. *Gigascience* 9:giaa021. <https://doi.org/10.1093/gigascience/gjaa021>.
- Saelzler, M. P., C. C. Spackman, Y. Liu, L. C. Martinez, J. P. Harris, and M. K. Abe. 2006. ERK8 down-regulates transactivation of the glucocorticoid receptor through Hic-5. *J. Biol. Chem.* 281:16821–16832. <https://doi.org/10.1074/jbc.M512418200>.
- Sanchez, M. P., M. El Jabri, S. Minéry, V. Wolf, E. Beuviel, C. Laithier, A. Delacroix-Buchet, M. Brochard, and D. Boichard. 2018. Genetic parameters for cheese-making properties and milk composition predicted from mid-infrared spectra in a large data set of Montbéliarde cows. *J. Dairy Sci.* 101:10048–10061. <https://doi.org/10.3168/jds.2018-14878>.
- Sanchez, M.-P., A. Govignon-Gion, P. Croiseau, S. Fritz, C. Hozé, G. Miranda, P. Martin, A. Barbat-Leterrier, R. Letaïef, D. Rocha, M. Brochard, M. Boussaha, and D. Boichard. 2017. Within-breed and multi-breed GWAS on imputed whole-genome sequence vari-

- ants reveal candidate mutations affecting milk protein composition in dairy cattle. *Genet. Sel. Evol.* 49:68. <https://doi.org/10.1186/s12711-017-0344-z>.
- Sanchez, M.-P., Y. Ramayo-Caldas, V. Wolf, C. Laithier, M. El Jabri, A. Michenet, M. Boussaha, S. Taussat, S. Fritz, A. Delacroix-Buchet, M. Brochard, and D. Boichard. 2019. Sequence-based GWAS, network and pathway analyses reveal genes co-associated with milk cheese-making properties and milk composition in Montbéliarde cows. *Genet. Sel. Evol.* 51:34. <https://doi.org/10.1186/s12711-019-0473-7>.
- Sanchez, M.-P., D. Rocha, M. Charles, M. Boussaha, C. Hozé, M. Brochard, A. Delacroix-Buchet, P. Grosperin, and D. Boichard. 2021. Sequence-based GWAS and post-GWAS analyses reveal a key role of SLC37A1, ANKH, and regulatory regions on bovine milk mineral content. *Sci. Rep.* 11:7537. <https://doi.org/10.1038/s41598-021-87078-1>.
- Sargolzaei, M., J. P. Chesnais, and F. S. Schenkel. 2014. A new approach for efficient genotype imputation using information from relatives. *BMC Genomics* 15:478. <https://doi.org/10.1186/1471-2164-15-478>.
- Schrooten, C., M. Bink, and H. Bovenhuis. 2004. Whole genome scan to detect chromosomal regions affecting multiple traits in dairy cattle. *J. Dairy Sci.* 87:3550–3560. [https://doi.org/10.3168/jds.S0022-0302\(04\)73492-X](https://doi.org/10.3168/jds.S0022-0302(04)73492-X).
- Smith, J. J., K. Sumiyama, and C. T. Amemiya. 2012. A living fossil in the genome of a living fossil: Harbinger transposons in the coelacanth genome. *Mol. Biol. Evol.* 29:985–993. <https://doi.org/10.1093/molbev/msr267>.
- Soyeurt, H., D. Bruwier, J.-M. Romnee, N. Gengler, C. Bertozzi, D. Veselko, and P. Dardenne. 2009. Potential estimation of major mineral contents in cow milk using mid-infrared spectrometry. *J. Dairy Sci.* 92:2444–2454. <https://doi.org/10.3168/jds.2008-1734>.
- Soyeurt, H., F. Dehareng, N. Gengler, S. McParland, E. Wall, D. Berry, M. Coffey, and P. Dardenne. 2011. Mid-infrared prediction of bovine milk fatty acids across multiple breeds, production systems, and countries. *J. Dairy Sci.* 94:1657–1667. <https://doi.org/10.3168/jds.2010-3408>.
- Suchocki, T., K. Wojdak-Maksymiec, and J. Szyda. 2016. Using gene networks to identify genes and pathways involved in milk production traits in Polish Holstein dairy cattle. *Czech J. Anim. Sci.* 61:526–538. <https://doi.org/10.17221/43/2015-CJAS>.
- Suwannasing, R., M. Duangjinda, W. Boonkum, R. Taharnklaew, and K. Tuangsitthanon. 2018. The identification of novel regions for reproduction trait in Landrace and Large White pigs using a single step genome-wide association study. *Asian-Australas. J. Anim. Sci.* 31:1852–1862. <https://doi.org/10.5713/ajas.18.0072>.
- The World Dairy Situation. 2016. Bulletin of the International Dairy Federation 485/2016. Brussels, Belgium. Accessed <https://www.fil-idf.org/wp-content/uploads/2016/10/TEASER.pdf>.
- Tiezzi, F., K. L. Parker-Gaddis, J. B. Cole, J. S. Clay, and C. Maltecca. 2015. A genome-wide association study for clinical mastitis in first parity US Holstein cows using single-step approach and genomic matrix re-weighting procedure. *PLoS One* 10:e0114919. <https://doi.org/10.1371/journal.pone.0114919>.
- Tiezzi, F., D. Pretto, M. De Marchi, M. Penasa, and M. Cassandro. 2013. Heritability and repeatability of milk coagulation properties predicted by mid-infrared spectroscopy during routine data recording, and their relationships with milk yield and quality traits. *Animal* 7:1592–1599. <https://doi.org/10.1017/S1751731113001195>.
- Vallas, M., H. Bovenhuis, T. Kaart, K. Pärna, H. Kiiman, and E. Pärna. 2010. Genetic parameters for milk coagulation properties in Estonian Holstein cows. *J. Dairy Sci.* 93:3789–3796. <https://doi.org/10.3168/jds.2009-2435>.
- VanRaden, P. M. 2008. Efficient methods to compute genomic predictions. *J. Dairy Sci.* 91:4414–4423. <https://doi.org/10.3168/jds.2007-0980>.
- Visentin, G., M. De Marchi, D. Berry, A. McDermott, M. Fenelon, M. Penasa, and S. McParland. 2017a. Factors associated with milk processing characteristics predicted by mid-infrared spectroscopy in a large database of dairy cows. *J. Dairy Sci.* 100:3293–3304. <https://doi.org/10.3168/jds.2016-12028>.
- Visentin, G., A. McDermott, S. McParland, D. Berry, O. Kenny, A. Brodtkorb, M. Fenelon, and M. De Marchi. 2015. Prediction of bovine milk technological traits from mid-infrared spectroscopy analysis in dairy cows. *J. Dairy Sci.* 98:6620–6629. <https://doi.org/10.3168/jds.2015-9323>.
- Visentin, G., S. McParland, M. De Marchi, A. McDermott, M. Fenelon, M. Penasa, and D. Berry. 2017b. Processing characteristics of dairy cow milk are moderately heritable. *J. Dairy Sci.* 100:6343–6355. <https://doi.org/10.3168/jds.2017-12642>.
- Walstra, P., P. Walstra, J. T. Wouters, and T. J. Geurts. 2005. *Dairy Science and Technology*. CRC Press. <https://doi.org/10.1201/9781420028010>.
- Wang, D., C. Ning, J.-F. Liu, Q. Zhang, and L. Jiang. 2019. Replication of genome-wide association studies for milk production traits in Chinese Holstein by an efficient rotated linear mixed model. *J. Dairy Sci.* 102:2378–2383. <https://doi.org/10.3168/jds.2018-15298>.
- Wang, P., X. Li, Y. Zhu, J. Wei, C. Zhang, Q. Kong, X. Nie, Q. Zhang, and Z. Wang. 2022. Genome-wide association analysis of milk production, somatic cell score, and body conformation traits in Holstein cows. *Front. Vet. Sci.* 9:932034. <https://doi.org/10.3389/fvets.2022.932034>.
- Wang, Z., C. S. Potter, J. P. Sundberg, and H. Hogenesch. 2012. SHARPIN is a key regulator of immune and inflammatory responses. *J. Cell. Mol. Med.* 16:2271–2279. <https://doi.org/10.1111/j.1582-4934.2012.01574.x>.
- Wedholm, A., L. Larsen, H. Lindmark-Månsson, A. Karlsson, and A. André. 2006. Effect of protein composition on the cheese-making properties of milk from individual dairy cows. *J. Dairy Sci.* 89:3296–3305. [https://doi.org/10.3168/jds.S0022-0302\(06\)72366-9](https://doi.org/10.3168/jds.S0022-0302(06)72366-9).
- Wiggans, G. R., T. Sonstegard, P. VanRaden, L. Matukumalli, R. Schnabel, J. Taylor, F. Schenkel, and C. Van Tassell. 2009. Selection of single-nucleotide polymorphisms and quality of genotypes used in genomic evaluation of dairy cattle in the United States and Canada. *J. Dairy Sci.* 92:3431–3436. <https://doi.org/10.3168/jds.2008-1758>.
- Zaalberg, R. M., N. Poulsen, H. Bovenhuis, J. Sehested, L. Larsen, and A. Buitenhuis. 2021. Genetic analysis on infrared-predicted milk minerals for Danish dairy cattle. *J. Dairy Sci.* 104:8947–8958. <https://doi.org/10.3168/jds.2020-19638>.

ORCID

- H. Atashi  <https://orcid.org/0000-0002-6853-6608>
 Y. Chen  <https://orcid.org/0000-0002-8593-4384>
 H. Wilmot  <https://orcid.org/0000-0002-1075-088X>
 S. Vanderick  <https://orcid.org/0000-0003-4556-4718>
 X. Hubin  <https://orcid.org/0000-0002-2765-4771>
 N. Gengler  <https://orcid.org/0000-0002-5981-5509>