Single-step genome-wide association analyses for milk urea concentration in Walloon Holstein

H. Atashi,[1,2]* Y. Chen,[1] S. Vanderick,[1] X. Hubin,[3] and N. Gengler[1]
[1] TERRA Teaching and Research Center, Gembloux Agro-Bio Tech, University of Liège, 5030 Gembloux, Belgium.
[2] Department of Animal Science, Shiraz University, 71441-13131 Shiraz, Iran.
[3] Elevéo asbl Awé Group, 5590 Ciney, Belgium

ABSTRACT

The aims of this study were to conduct a single-step genome-wide association to identify genomic regions associated with milk urea (**MU**) and to estimate genetic correlations between MU and milk yield (**MY**), milk composition (calcium content (**CC**), fat percentage (**FP**), protein percentage (**PP**), and casein percentage (**CNP**)), and cheese-making properties (**CMP**) (coagulation time (**CT**), curd firmness after 30 min from rennet addition (**a30**), and titratable acidity (**TA**)). The used data have been collected from 2015 to 2020 on 78,073 first-parity (485,218 test-day records), and 48,766 s-parity (284,942 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 the top-3 genomic regions explaining the largest rate of the genetic variance were considered promising regions and used to identify potential candidate genes. Mean (standard deviation) MU was 25.38 (8.02) mg/dl and 25.03 (8,06) mg/dl in the first and second lactation, respectively. Mean heritability estimates for daily MU were 0.21 and 0.23 for the first and second lactation, respectively. The genetic correlations estimated between MU and MY, milk composition, and CMP were quite low (ranged from −0.10 (CC) to 0.10 (TA) and −0.05 (CT) to 0.13 (TA) for the first and second lactations, respectively). The top-3 regions associated with MU were located from 80.61 to 80.74 Mb on BTA6, 103.26 to 103.41 Mb on BTA11, and 1.59 to 2.15 Mb on BTA14. Genes including *KCNT1*, *MROH1*, *SHARPIN*, *TSSK5*, *CPSF1*, *HSF1*, *TONSL*, *DGAT1*, *SCX*, and *MAF1* were identified as positional candidate genes for MU. The findings of this study can be used for a better understanding of the genomic architecture underlying MU in Holstein cattle.

Key words: milk urea, genomic association, Holstein

INTRODUCTION

Nitrogen (**N**) emissions from livestock can have harmful effects on biodiversity, soil, water and air quality, as they account for 23% of global nitrous oxide and 60% of global ammonia emissions (Uwizeye et al., 2020). Nitrogen is released to the environment through the feces and urine of livestock as undigested N fractions and as the end products of N metabolism, respectively (Müller et al., 2021). Nitrogen, as oxides or ammonia, is one of the greenhouses gases contributing to air pollution and through leaching to rivers and ground water resources. Therefore, reducing N excretion by dairy cattle is desirable due to global concerns about the agricultural contribution to environmental pollution by N. The first requirement for conducting a successful genetic selection is to establish a method to measure the traits or phenotypes of interest on large number of animals with a low cost.

The breakdown of protein, both in the rumen and in the small intestine, results in the production of ammonia, which is then detoxified by conversion to urea in the liver or kidneys. The urea can then be recycled into the rumen via saliva or by direct diffusion across the rumen wall, be excreted in urine (urinary urea; **UU**), or be secreted in milk (milk urea; **MU**) (Alio et al., 2000). Although most urea is eliminated from blood with urine excretion, UU is difficult to measure. MU can be measured at a low cost and is included as a standard part in most milk recording systems and is positively correlated with UU (Burgos et al., 2007, Burgos et al., 2010). Furthermore, the content of urea in milk represents approximately 48% of **NPN** in milk and is known as an important parameter of milk quality. Moreover,

MU is associated with milk yield composition (e.g., fat and protein contents) (Chen et al., 2022a). It has been reported that MU is associated with cheese-making properties (**CMP**) and that higher MU is associated with greater coagulation time required for milk (Guinot-Thomas, 1992). In addition, MU has been shown to be correlated with nitrogen utilization efficiency and is considered as an on-farm indicator to guide nutritional strategies in dairy cows (Schepers and Meijer, 1998, Nousiainen et al., 2004, Chen et al., 2022a). There is a strong positive relationship between blood urea (**BU**) nitrogen and MU (Butler et al., 1996, van den Berg et al., 2021). Furthermore, it has been shown that an increased BU is associated with a decreased uterine pH and a decreased in the probability of conception in dairy cows (Elrod and Butler, 1993, Elrod et al., 1993, Butler et al., 1996). It has been reported that an increased MU is associated with a decreased reproductive performance in dairy cattle (Rajala-Schultz et al., 2001, Guo et al., 2004, Hojman et al., 2004).

Reducing MU can have many benefits, such as reducing N emissions, improving feed efficiency and animal health, milk quality and reproductive performance; therefore, this trait could be a potential selection criterion for breeding dairy cattle (Ariyarathne et al., 2021). However, a comprehensive knowledge of the genetic background of MU is needed before it is included into the breeding program. Furthermore, identification of genomic regions and individual genes responsible for genetic variation in MU will improve our understanding about the biological pathways involved and can be used for decreasing MU.

Although MU has been included in the regular milk recording of dairy cows in the Walloon Region of Belgium for more than 2 decades, the potential contribution of this trait in the breeding program of Walloon dairy cows is still under investigation. The genetic parameters for MU and its genetic correlations with traditional traits have previously been investigated in Walloon Holstein (Chen et al., 2021, 2022b); however, no comprehensive study has been conducted to investigate the genomic architecture of MU for the whole period of lactation and the potential genetic relationships between MU and milk technological traits (e.g., CMP) has not been investigated in this population of dairy cattle. Therefore, the aims of this study were to identify genomic regions associated with MU concentration and to estimate genetic correlation between MU and selected cheese-making traits in Walloon Holstein cows.

MATERIALS AND METHODS

Phenotypic Data

The data used consisted of test-day records of traits including milk yield (**MY**), milk urea (**MU**), mid-infrared (**MIR**) prediction of fat percentage (**FP**), protein percentage (**PP**), casein percentage (**CNP**), calcium content (**CC**), titratable acidity (**TA**), coagulation time (**CT**), and curd firmness at 30 min after coagulant addition (**a30**). MIR prediction equations used to predict CNP, CC, TA, CT, and a30 were obtained from various studies (Soyeurt et al., 2009, Colinet et al., 2010, Colinet et al., 2013, Colinet et al., 2015). The data were collected from 2015 to 2020 on 78,073 animals distributed in 671 herds in the Walloon Region of Belgium. Data were edited to include only cows with known birth date, calving date, and parity number. Only records from the first 2 parities that had data for all included traits on a given test-day were kept. Records from days in milk (**DIM**) lower than 5 d and greater than 365 d were eliminated. Age at the first calving (**AFC**) was calculated as the difference between birth date and first calving date and restricted to the range of 540 to 1200 d. Daily MY, FP, and PP were restricted to range from 3 to 99 kg, 1.5 to 9% and 1 to 7%, respectively (ICAR, 2022). Test-day records of the rest of the considered traits were edited to remove records outside the range of mean ± 3 standard deviations (**SD**). The number of test-day records in the first- and second-parity cows were 485,218 (on 78,073 cows) and 284,942 (on 48,766 cows), respectively. On average across the data set, 5.51 test-day records were available per cow per lactation. Pedigree depth of the animals was traced back to 5 generations. Full pedigree records included 186,548 females and 10,076 males.

Genotypic Data

Genotypic data were available for 6,617 (1,712 males) 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, USA). Single nucleotide polymorphisms (**SNP**) in common among the 4 chips were kept. Non-mapped SNP, SNP located on sexual chromosomes, and triallelic SNPs were excluded. A minimum GenCall Score of 0.15 and a minimum Gen-Train 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) using FImpute V2.2 software (Sargolzaei et al., 2014). SNP 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 Bos taurus autosomes (**BTA**) were used in the genomic analyses.

Variance Component Estimation

The (co)variance components and breeding values for MU 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 2 lactations) model (Paiva et al., 2022):

$$y_{ijklmn} = \mu + HTDp_i + \sum_{b=0}^{3} AS_j \Phi_b(t) + \sum_{b=0}^{3} HY_k \Phi_b(t) +$$
$$\sum_{b=0}^{3} pe_m \Phi_b(t) + \sum_{b=0}^{3} a_m \Phi_b(t) + e_{ijklmn},$$

where y_{ijklmn} is the test-day record of MU belonging to the DIM n of cow m in parity l, belonging to ith class of HTDp, jth class of AS, and kth class of HY; HTDp is the fixed effect of herd-testday-parity; AS is the fixed effect of age-season of calving defined as follows: age at calving class (6, and 4 classes of age at calving were created for the first and second parity, respectively) × 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); $\sum_{b=0}^{3} AS_j \Phi_b(t)$ is the fixed regression coefficients of the age-season of calving modeled using Legendre polynomials of order 3; $\sum_{b=0}^{3} HY_k \Phi_b(t)$, $\sum_{b=0}^{3} pe_m \Phi_b(t)$, 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:

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

where **HY** is the 8 × 8 covariance matrix of the herd-year of calving regression coefficients; **I** is an identity matrix; ⊗ represents the Kronecker product function; **P** is the 8 × 8 covariance matrix of the permanent environmental regression coefficients; **Ga** is the 8 × 8 covariance matrix of the additive genetic regression coefficients; blocks within $\mathbf{R} = \sum \mathbf{r^p}^{+}$ contain residual variance (r) that depends on parity (p). Residual variance was assumed to be the same within each parity. The **H** is a matrix that combines pedigree and genomic relationships, which its inverse consists on the integration of additive and genomic relationship matrices, **A** and **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 **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 **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 **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 **Z** is a matrix of gene content adjusted for allele frequencies (0, 1 or 2 for *aa*, *Aa* and *AA*, respectively); **D** is a diagonal matrix of weights for SNP variances (**D** = **I**); M is the number of SNPs, and p_i is the MAF of the ith SNP. The **H** matrix was built scaling **G** based on $\mathbf{A_{22}}$ considering that the average of the diagonal of **G** is equal to the average of the diagonal of $\mathbf{A_{22}}$ and, the average of the off-diagonal **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. 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 (**GEBV**) of MU 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, $\mathbf{GEBV_i} = \mathbf{T}\hat{\mathbf{g}}_i$, where $\hat{\mathbf{g}}_i$ 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.

In addition, the same RR-TDM through multiple-trait (2 traits), multiple-lactation (first 2 lactations) was used to estimate correlation between MU and the studied traits of milk yield, milk composition, and CMP. Moreover, phenotypic trend for MU was obtained by linear regression of phenotypic values over calving year.

Genome-Wide Association Study

GEBV of animal i through the entire lactation (DIM 5 to 365) were obtained by averaging the daily GEBV solutions of all DIM; that is,

$$\mathrm{GE\hat{B}V_i} = \left(\mathrm{GE\hat{B}V_{i5}} + \mathrm{GE\hat{B}V_{i6}} + \ldots + \mathrm{GE\hat{B}V_{i365}}\right)/361,$$

where $\mathrm{GE\hat{B}V_i}$ is the GEBV of animal i through the entire lactation, obtained by averaging the GEBV from 5 to 365.

The SNP effects were estimated 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, $\mathbf{a_g} = \mathbf{Zu}$, 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:

$$\mathrm{Var}\left(\mathbf{a_g}\right) = \mathrm{Var}\left(\mathbf{Zu}\right) = \mathbf{ZDZ'}\sigma_u^2 = \mathbf{G}\sigma_a^2,$$

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:

$$\hat{u} = \lambda\mathbf{DZ'}\ \mathbf{G}^{-1}\ \hat{a}_g = \mathbf{DZ'}\left[\mathbf{ZDZ'}\right]^{-1}\hat{a}_g,$$

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

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

The percentage of the total additive genetic variance explained by the ith genomic region was estimated as following:

$$\frac{\mathrm{Var}\left(a_i\right)}{\sigma_a^2}\times100\% = \frac{\mathrm{Var}\left(\sum_{j=1}^{50}Z_j\hat{u}_j\right)}{\sigma_a^2}\times100,$$

where a_i is the genetic value of the ith 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 jth SNP for all individuals; and $\hat{u}_j$ is the marker effect of the jth SNP within the ith region. 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 the top-3 genomic regions explaining the largest rate of the total additive genetic variance were considered promising regions and used to identify potential candidate genes (Soares et al., 2021). 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). In addition, we performed a GWAS with frequentist p-values to detect single SNP associated with MU. The thresholds of the Bonferroni corrected p values for 5% and 1% genome-wide significance associations were set as $8.84\mathrm{E}^{-8}$ (0.05 divided by the number of SNPs) and $1.74\mathrm{E}^{-8}$ (0.01 divided by the number of SNPs), respectively.

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, USA); then, the genotypes were imputed to the BovineHD Beadchip which are 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 positional candidate genes associated with MU, 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. The list of identified positional candidate genes was uploaded to Enrichr to perform gene ontology (GO) enrichment analyses (Kuleshov et al., 2016). Significantly enriched terms were identified based on the retrieved adjusted P value.

RESULTS

The descriptive statistics for studied traits are presented in Table 1. Mean (SD) MU was 25.38 (8.02) and 25.03 (8.06) mg/dl for the first and second lactation, respectively. Daily MY averaged 24.2 kg (4.01% fat, 3.38% protein, and 2.59% casein), and 27.9 kg (4.10% fat, 3.46% protein, and 2.63% casein) in the first and second lactations. The trend in concentrations of MU across DIM is given in Figure 1. The lowest value for MU was found at the beginning of lactation, increased rapidly by increasing DIM, reached its peak at the middle of lactation, then slightly decreased by increasing DIM to the end of the lactation. The coefficient of variation (CV%) of MU was 32% in both lactations. Environmental factors including herd, calving year, calving season, age at first calving, parity and lactation stage affected MU ($P < 0.05$). Phenotypic trend, obtained by linear regression of MU phenotypic values over calving year, showed a positive trend of 0.73 and 0.79 mg/dl milk for the first and second lactations, respectively (Figure S1) (https://github.com/hadiatashi/Holstein-MU).

Mean heritability estimates for daily MU was 0.21 (SD = 0.02; range = 0.15 to 0.23), and 0.23 (SD = 0.02; range = 0.19 to 0.24) for the first and second lactation, respectively. Genetic correlations between MU in the first and second lactation ranged from 0.84 to 0.97 with a mean of 0.95. Genetic correlations estimated between MU and the studied traits are presented in Table 2. Genetic correlations between MU and milk yield traits ranged from −0.10 (CC) to 0.05 (MY) and −0.02 (CC) to 0.12 (FP) for the first and second lactation, respectively. Low genetic correlations ranged from 0.00 (a30) to 0.10 (TA) and from −0.05 (CT) to 0.13 (TA) were estimated between MU and cheese-making traits in the first and second lactation, respectively.

General information about the results of single-step GWAS for MU concentration are described in Data S1 and Data S2 (https://github.com/hadiatashi/Holstein-MU). The Manhattan plots of the proportion of total additive genetic variance explained by 50-SNP windows are shown in Figures 2. The top 3 windows explaining the highest percentage of the genetic variance for MU along with corresponding genes are presented in Table 3. The identified regions were located from 80.61 to 80.74 Mb on BTA6, 103.26 to 103.41 Mb on BTA11, and 1.59 to 2.15 Mb on BTA14 (UMD3.1 assembly) and combined explained 0.99% and 1.04% of the total additive genetic variance for MU at the first and second lactations, respectively. The region identified on BTA6 was 0.13 Mb in size and explained 0.34% of the total additive genetic

Table 1. Descriptive statistics for milk yield traits and cheese making properties in Walloon Holstein cows[1]

Trait	First lactation			Second lactation		
	Mean	SD	CV (%)	Mean	SD	CV (%)
Milk yield (kg)	24.2	6.01	24.86	27.9	8.29	29.68
Fat percentage (%)	4.01	0.68	15.94	4.10	0.67	16.37
Protein percentage (%)	3.38	0.34	10.04	3.46	0.37	10.73
Casein percentage (%)	2.59	0.29	11.23	2.63	0.31	11.69
Calcium content (mg/kg milk)	1155	100.0	8.66	1172	106.9	9.13
Titratable acidity (°D)[2]	16.52	1.41	8.52	16.0	1.44	9.03
Coagulation time[3]	18.16	2.99	16.46	18.25	3.05	16.72
Curd firmness (a30)[4]	32.11	2.30	7.17	32.31	2.44	7.54
Milk urea (mg/dl milk)	25.38	8.02	31.61	25.03	8.06	32.22

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

[2]Milk titratable acidity in Dornic degree (°D).

[3]Coagulation time is defined here as the sum of the rennet coagulation time (RCT) plus the time to a curd firmness of 20 mm (k20) measured by the Computerized Renneting Meter.

[4]Curd firmness is defined as the curd firmness measured 30 min after enzyme addition by the Computerized Renneting Meter.

Table 2. Genetic correlation estimated between MU and milk yield, milk composition and cheese-making traits in Walloon Holstein cows[1]

Trait	First lactation			Second lactation		
	Mean	SD	Range	Mean	SD	Range
Milk yield (kg)	0.05	0.01	0.03 to 0.09	−0.01	0.02	−0.04 to 0.08
Fat percentage (%)	0.04	0.01	−0.03 to 0.07	0.12	0.03	0.09 to 0.21
Protein percentage (%)	0.01	0.02	−0.02 to 0.04	0.06	0.06	−0.07 to 0.17
Casein percentage (%)	0.03	0.03	0.00 to 0.06	0.10	0.05	−0.02 to 0.21
Calcium content (mg/kg milk)	−0.10	0.01	−0.15 to −0.08	−0.02	0.02	−0.08 to 0.04
Titratable acidity (°D)[2]	0.10	0.04	0.02 to 0.19	0.13	0.01	−0.04 to 0.20
Coagulation time[3]	0.03	0.02	−0.02 to 0.04	−0.05	0.01	−0.07 to −0.02
Curd firmness (a30)[4]	0.00	0.04	−0.04 to 0.07	0.09	0.01	−0.08 to 0.18

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

[2]Milk titratable acidity in Dornic degree (°D).

[3]Coagulation time is defined here as the sum of the rennet coagulation time (RCT) plus the time to a curd firmness of 20 mm (k20) measured by the Computerized Renneting Meter.

[4]Curd firmness is defined as the curd firmness measured 30 min after enzyme addition by the Computerized Renneting Meter.

variance of MU in the first and second lactation. The region found on BTA11 was 0.14 Mb in size, harbors the potassium sodium-activated channel subfamily T member 1 (*KCNT1*) gene, and explained 0.39% and 0.40% of the total additive genetic variance of MU in the first and second lactation, respectively. The region found of BTA14 explained 0.26% and 0.30% of the total additive genetic variance of MU in the first and second lactation, respectively. This region was 0.56 Mb in size and harbors genes including diacylglycerol O-acyltrans-ferase 1 (*DGAT1*), and scleraxis bHLH transcription factor (*SCX*). Regulation of protein polyubiquitination, regulation of gastrulation, regulation of double-strand break repair and monoacylglycerol biosynthetic process are among the most important biological process terms enriched by the candidate genes identified for MU. The results of the p-value based GWAS analyses showed that total number of SNP associated with MU differed slightly between lactations (Data S3; Data S4). However, the SNP identified using the single-SNP p-value
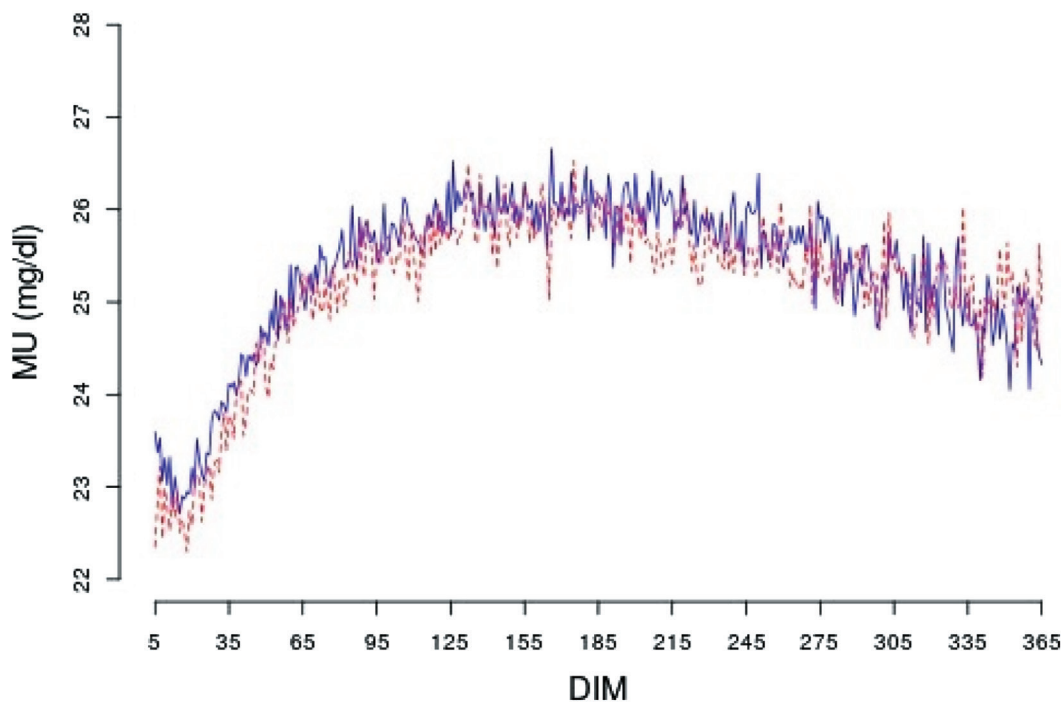

Figure 1. Lactation curves for milk urea (MU) concentration (mg/dl milk) for the first (blue), and second (red) parity in Walloon Holstein cows.

method were the same as those identified using the window-based GWAS procedure (Figure S2) (https://github.com/hadiatashi/Holstein-MU).

DISCUSSION

Mean MU in the first and second lactations was 25.38 and 25.03 mg/dl, respectively, which is in close agreement with those previously reported for Holstein (Bastin et al., 2009, Rzewuska and Strabel, 2015, Atashi and Hostens, 2021, Chen et al., 2021) and the Italian Brown Swiss cows (Samoré et al., 2007), but higher than that reported for the Dual-Purpose Belgian Blue (**DPBB**) cows (Atashi et al., 2022). The coefficient of variation was 32%. The CV for MU was 26% in Belgian Holstein (Atashi and Hostens, 2021), 33% in Dutch Holstein (Stoop et al., 2007), and 44% to 45% in the DPBB cows (Atashi et al., 2022). The trend in concentration of MU across DIM showed that lowest value for MU was found at the beginning of lactation, then increased rapidly by increasing DIM, reached the peak at the middle of lactation, then slightly decreased by increasing DIM to the end of the lactation. It is well documented that dairy cows are in a negative energy balance state in the early stage of the lactation when the feed intake cannot meet their requirements. During this time, it can be assumed that cows use N as efficiently as possible which can explain, at least in part, the obtained results (Chen et al., 2021).

Similar lactation curve pattern has been reported for MU by previous studies (Mucha and Strandberg, 2011, Rzewuska and Strabel, 2013, Atashi and Hostens, 2021, Atashi et al., 2022). Mean heritability estimates for daily MU were 0.21 and 0.23 for the first and second lactation, respectively. The heritability reported in the literature for MU range from 0.09 to 0.47, depending on the method used to measure the MU content, the population considered, the parity, and the model used to estimate variance components.

High genetic correlation was found between MU in first and second lactation which is in a close agreement previous studies (Miglior et al., 2007, Atashi and Hostens, 2021). The results showed that MU is weakly correlated with milk yield and composition. Genetic correlation between MU and milk yield traits has been investigated by several researchers, however the results are inconsistent. Atashi and Hostens (2021) reported that genetic correlations between MU with milk yield traits were close to zero in Holstein dairy cows. Godden et al. (2001) reported a positive genetic correlation between MY and MU. Diab and Hillers (1996) reported a negative relationship between MU and milk yield. Samoré et al. (2007) reported that MU had a positive genetic relationship with FY (0.12), null with PY (0.03)

and a negative genetic correlation with MY (−0.17) in Italian Brown Swiss dairy cattle. Research on the relationship between MU and cheese-making traits is limited. This study showed low genetic correlations between MU and cheese-making traits. Mean genetic correlation estimates between MU and TA ranged from 0.10 to 0.13. Martin et al. (1997) reported that MU influences directly acidification kinetics, chemical composition and texture characteristics of cheeses. Guinot-Thomas (1992) reported that higher MU is associated with greater coagulation time required for milk. The variation found for genetic correlation of MU with productive traits in the literature can be explained by the differences in structure of the data, number of records, the statistical models used, 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 considered in single SNP-based association study (Bao and Wang, 2017). Therefore, window-based GWAS procedure have been proposed as an effective method to estimate the combined effect of several consecutive SNPs in a specific region and to identify genomic regions that explain a given amount of genetic variance (Aguilar et al., 2019). 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 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 window of 50 adjacent SNP with an average size of ~216 Kb and he top-3 genomic regions explaining the largest rate of the total additive genetic variance were considered promising regions and used to identify potential 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.
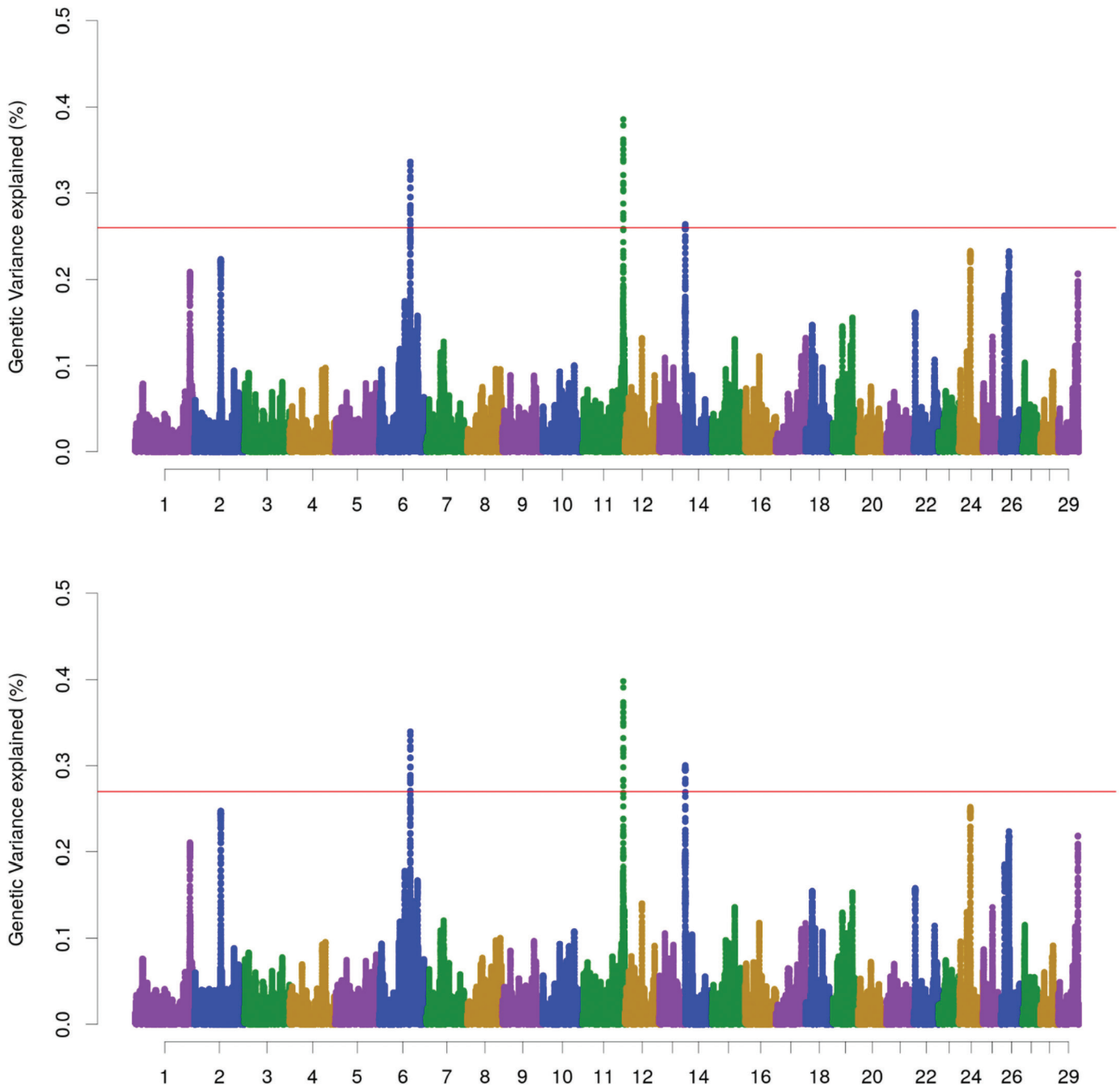
Figure 2. Additive genetic variance explained by windows of 50 adjacent SNP across chromosomes for milk urea concentration for the first (first row) and second (second row) lactation in Walloon Holstein cows. The red lines showed that the most important 3 windows explained more than 0.25% of the total additive genetic variance of MU in both lactations.

The top-3 identified genomic regions were located from 80.61 to 80.74 Mb on BTA6, 103.26 to 103.41 Mb on BTA11, and 1.59 to 2.15 Mb on BTA14 and combined explained 0.99% and 1.04% of the total additive genetic variance for MU at the first and second lactations, respectively. The region identified on BTA6 was 0.13 Mb in size and explained 0.34% of the total additive genetic variance of MU in the first and second lactation. It has been reported that SNPs inside this region are associated with traits including MY, fat yield (**FY**), protein yield (**PY**), milk fatty acid (**FA**) profile and CNP (Cole et al., 2011, Pausch et al., 2017, Sanchez et al., 2017b, Benedet et al., 2019). van den Berg et al. (2022) reported that this region is associated with MU nitrogen in Australian and New Zealand dairy cattle. Ma et al. (2023) reported that SNPs inside this region are associated with MU nitrogen in Holstein cows.

The genomic region located from 103.26 to 103.41 Mb on BTA11 was associated with MU in the first and second lactations. Previous studies showed that SNPs inside this region are associated with MY, FY, PY, FP, PP, CNP, somatic cell score (**SCS**), and CMP traits (Bennewitz et al., 2004, Chen et al., 2006, Kučerová et al., 2006, Schopen et al., 2009). This region harbors *KCNT1* gene which encodes a sodium-activated potassium channel subunit which is thought to function in ion conductance and developmental signaling pathways. *KCNT1* gene has been reported to be associated with MY, FY, and PY in US Holstein (Cole et al., 2011).

The genomic region located from 1.59 to 2.15 Mb on BTA14 was associated with MU in the first and second lactation. Previous studies showed that SNPs inside this region are associated with MY, FY, PY, FP, PP, SCS, milk FA profile, and CNP in breeds of dairy cattle (Conte et al., 2010, Buitenhuis et al., 2016, Pedrosa et al., 2021). Chen et al. (2022b) reported that this region is associated with nitrogen efficiency index and milk true protein nitrogen in Holstein cows. Clancey

et al. (2019) reported that SNP inside this region are associated with MY in Holstein cows. This region was 0.56 Mb in size and harbors genes including the scratch family zinc finger 1 (*SCRT1*), 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*). The *DGAT1*, involves 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, 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 Holstein cattle. The SHANK associated RH domain interactor gene (*SHARPIN*) is mapped 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. (2017a) reported *DGAT1*, maestro heat like repeat family member 1 (*MROH1*), and 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.

CONCLUSION

This study showed that MU is moderately heritable and is weakly correlated with MY, milk composition, and cheese-making traits; therfore, selection for lower MU concentration would not influnce the production performance. In general, windows explained less than 0.50% of the total additive genetic variance of MU and these low contributing regions were spread across the genome. This indicates that MU is highly polygenic, in which many regions across the genome contribute to its genetic variation.

Table 3. Genomic regions associated to milk urea concentration in the first and second lactations for Walloon Holstein cows[1]

Chromosome	Position (bp)[2]	Position (bp)[3]	Gene[4]
BTA6	80,611,486 - 80,741,948	78,982,492 – 78,989,732	
BTA11	103,264,921– 103,409,247	103,309,149 – 103,336,835	*KCNT1*
BTA14	1,588,879 – 2,150,825	481,647 – 870,500	*MROH1, SHARPIN, TSSK5, GPAA1, PLEC, VPS28, CPSF1, SLC39A4, HSF1, OPLAH, GRINA, SMPD5, WDR97, FBXL6, TONSL, SLC52A2, ADCK5, SCRT1, TMEM249, CYC1, HGH1, SPATC1, EXOSC4, BOP1, PARP10, DGAT1, SCX, MAF1, MIR1839*

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

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

[3]The positions of the identified genomic regions based on the ARS-UCD1.2 assembly.

[4]Genes inside the genomic region. Official gene symbol (Assembly ARS-UCD1.2, annotation release 105).

ACKNOWLEDGMENTS

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. 51:1–8. 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 programs. in Proc. 10. World Congress on Genetics Applied to Livestock Production (WCGALP). American Society of Animal Science.

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. Pages 11–16 in Proc. Proceedings of the world congress on genetics applied to livestock production.

Alio, A., C. Theurer, O. Lozano, J. Huber, R. Swingle, A. Delgado-Elorduy, P. Cuneo, D. DeYoung, and K. Webb Jr.. 2000. Splanchnic nitrogen metabolism by growing beef steers fed diets containing sorghum grain flaked at different densities. J. Anim. Sci. 78:1355–1363. https://doi.org/10.2527/2000.7851355x.

Ariyarathne, H. B. P. C., M. Correa-Luna, H. Blair, D. Garrick, and N. Lopez-Villalobos. 2021. Can nitrogen excretion of dairy cows be reduced by genetic selection for low milk urea nitrogen concentration? Animals (Basel) 11:737. https://doi.org/10.3390/ani11030737.

Atashi, H., Y. Chen, H. Wilmot, S. Vanderick, X. Hubin, and N. Gengler. 2022. Genome-wide association for milk urea concentration in Dual-Purpose Belgian Blue cows. J. Anim. Breed. Genet. 139:710–722. https://doi.org/10.1111/jbg.12732.

Atashi, H., and M. Hostens. 2021. Genetic parameters for milk urea and its relationship with milk yield and compositions in Holstein dairy cows. PLoS One 16:e0253191. https://doi.org/10.1371/journal.pone.0253191.

Atashi, H., M. Salavati, J. De Koster, J. Ehrlich, M. Crowe, G. Opsomer, and M. Hostens. 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.

Bao, M., and K. Wang. 2017. Genome-wide association studies using a penalized moving-window regression. Bioinformatics 33:3887–3894. https://doi.org/10.1093/bioinformatics/btx522.

Bastin, C., L. Laloux, A. Gillon, F. Miglior, H. Soyeurt, H. Hammami, C. Bertozzi, and N. Gengler. 2009. Modeling milk urea of Walloon dairy cows in management perspectives. J. Dairy Sci. 92:3529–3540. https://doi.org/10.3168/jds.2008-1904.

Beissinger, T. M., G. J. Rosa, S. M. Kaeppler, D. Gianola, and N. de Leon. 2015. Defining window-boundaries for genomic analyses using smoothing spline techniques. Genet. Sel. Evol. 47:1–9. https://doi.org/10.1186/s12711-015-0105-9.

Benedet, A., P. Ho, R. Xiang, S. Bolormaa, M. De Marchi, M. Goddard, and J. Pryce. 2019. The use of mid-infrared spectra to map genes affecting milk composition. J. Dairy Sci. 102:7189–7203. https://doi.org/10.3168/jds.2018-15890.

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.

Buitenhuis, B., N. A. Poulsen, G. Gebreyesus, and L. B. Larsen. 2016. Estimation of genetic parameters and detection of chromosomal regions affecting the major milk proteins and their post translational modifications in Danish Holstein and Danish Jersey cattle. BMC Genet. 17:1–12. https://doi.org/10.1186/s12863-016-0421-2.

Burgos, S., N. Embertson, Y. Zhao, F. Mitloehner, E. DePeters, and J. Fadel. 2010. Prediction of ammonia emission from dairy cattle manure based on milk urea nitrogen: Relation of milk urea nitrogen to ammonia emissions. J. Dairy Sci. 93:2377–2386. https://doi.org/10.3168/jds.2009-2415.

Burgos, S., J. Fadel, and E. DePeters. 2007. Prediction of ammonia emission from dairy cattle manure based on milk urea nitrogen: Relation of milk urea nitrogen to urine urea nitrogen excretion. J. Dairy Sci. 90:5499–5508. https://doi.org/10.3168/jds.2007-0299.

Butler, W., J. Calaman, and S. Beam. 1996. Plasma and milk urea nitrogen in relation to pregnancy rate in lactating dairy cattle. J. Anim. Sci. 74:858–865. https://doi.org/10.2527/1996.744858x.

Chen, H., Q. Zhang, C. Yin, C. Wang, W. Gong, and G. Mei. 2006. Detection of quantitative trait loci affecting milk production traits on bovine chromosome 6 in a Chinese Holstein population by the daughter design. J. Dairy Sci. 89:782–790. https://doi.org/10.3168/jds.S0022-0302(06)72140-3.

Chen, Y., H. Atashi, C. Grelet, S. Vanderick, H. Hu, and N. Gengler. 2022a. Defining a nitrogen efficiency index in Holstein cows and assessing its potential effect on the breeding program of bulls. J. Dairy Sci. 105:7575–7587. https://doi.org/10.3168/jds.2021-21681.

Chen, Y., H. Atashi, S. Vanderick, R. Mota, H. Soyeurt, H. Hammami, and N. Gengler. 2021. Genetic analysis of milk urea concentration and its genetic relationship with selected traits of interest in dairy cows. J. Dairy Sci. 104:12741–12755. https://doi.org/10.3168/jds.2021-20659.

Chen, Y., H. Atashi, S. Vanderick, R. Mota, H. Soyeurt, H. Hammami, and N. Gengler. 2022b. Genome-wide association study and functional annotation analyses for nitrogen efficiency index and its composition traits in dairy cattle. J. Dairy Sci. 105.

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 Tassell, J. Yang, S. Wang, L. K. Matukumalli, 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 12:1–17. 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. Page 400 in Proc. Book of Abstracts of the 66th Annual Meeting of the European Federation of Animal Science. Wageningen Academic Publishers.

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. .

Conte, G., M. Mele, S. Chessa, B. Castiglioni, A. Serra, G. Pagnacco, and P. Secchiari. 2010. Diacylglycerol acyltransferase 1, stearoyl-CoA desaturase 1, and sterol regulatory element binding protein 1 gene polymorphisms and milk fatty acid composition in Italian Brown cattle. J. Dairy Sci. 93:753–763. https://doi.org/10.3168/jds.2009-2581.

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.

Diab, I. A., and J. Hillers. 1996. Effect of selection for milk yield and dietary energy on yield traits, bovine somatotropin, and plasma urea nitrogen in dairy cows. J. Dairy Sci. 79:682–688. https://doi.org/10.3168/jds.S0022-0302(96)76414-7.

Elrod, C., and W. Butler. 1993. Reduction of fertility and alteration of uterine pH in heifers fed excess ruminally degradable protein. J. Anim. Sci. 71:694–701. https://doi.org/10.2527/1993.713694x.

Elrod, C., M. Van Amburgh, and W. Butler. 1993. Alterations of pH in response to increased dietary protein in cattle are unique to the uterus. J. Anim. Sci. 71:702–706. https://doi.org/10.2527/1993.713702x.

Godden, S., K. Lissemore, D. Kelton, K. Leslie, J. Walton, and J. Lumsden. 2001. Factors associated with milk urea concentrations in Ontario dairy cows. J. Dairy Sci. 84:107–114. https://doi.org/10.3168/jds.S0022-0302(01)74458-X.

Guinot-Thomas, P. 1992. Technological and microbiological consequences related to urea addition to milk. Aust. J. Dairy Technol. 47:58.

Guo, K., E. Russek-Cohen, M. Varner, and R. Kohn. 2004. Effects of milk urea nitrogen and other factors on probability of conception of dairy cows. J. Dairy Sci. 87:1878–1885. https://doi.org/10.3168/jds.S0022-0302(04)73346-9.

Habier, D., R. L. Fernando, K. Kizilkaya, and D. J. Garrick. 2011. Extension of the Bayesian alphabet for genomic selection. BMC Bioinformatics 12:1–12. 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. 17:1–11. https://doi.org/10.1186/s12863-016-0454-6.

Hojman, D., O. Kroll, G. Adin, M. Gips, B. Hanochi, and E. Ezra. 2004. Relationships between milk urea and production, nutrition, and fertility traits in Israeli dairy herds. J. Dairy Sci. 87:1001–1011. https://doi.org/10.3168/jds.S0022-0302(04)73245-2.

ICAR. 2022. Section 2 - Guidelines for Dairy Cattle Milk Recording. Section 2 – Cattle Milk Recording Version May, 2022. https://www.icar.org/Guidelines/02-Overview-Cattle-Milk-Recording.pdf.

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.

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.

Kučerová, J., M. Lund, P. Sørensen, G. Sahana, B. Guldbrandtsen, V. Nielsen, B. Thomsen, and C. Bendixen. 2006. Multitrait quantitative trait loci mapping for milk production traits in Danish Holstein cattle. J. Dairy Sci. 89:2245–2256. https://doi.org/10.3168/jds.S0022-0302(06)72296-2.

Kuleshov, M. V., M. R. Jones, A. D. Rouillard, N. F. Fernandez, Q. Duan, Z. Wang, S. Koplev, S. L. Jenkins, K. M. Jagodnik, A. Lachmann, M. G. McDermott, C. D. Monteiro, G. W. Gundersen, and A. Ma'ayan. 2016. Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. Nucleic Acids Res. 44(W1):W90–W97. https://doi.org/10.1093/nar/gkw377.

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.

Ma, L., H. Luo, L. F. Brito, Y. Chang, Z. Chen, W. Lou, F. Zhang, L. Wang, G. Guo, and Y. Wang. 2023. Estimation of genetic parameters and single-step genome-wide association studies for milk urea nitrogen in Holstein cattle. J. Dairy Sci. 106:352–363. https://doi.org/10.3168/jds.2022-21857.

Martin, B., J. Coulon, J. Chamba, and C. Bugaud. 1997. Effect of milk urea content on characteristics of matured Reblochon cheeses. Lait 77:505–514. https://doi.org/10.1051/lait:1997436.

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.

Miglior, F., A. Sewalem, J. Jamrozik, J. Bohmanova, D. Lefebvre, and R. Moore. 2007. Genetic analysis of milk urea nitrogen and lactose and their relationships with other production traits in Canadian Holstein cattle. J. Dairy Sci. 90:2468–2479. https://doi.org/10.3168/jds.2006-487.

Mucha, S., and E. Strandberg. 2011. Genetic analysis of milk urea nitrogen and relationships with yield and fertility across lactation. J. Dairy Sci. 94:5665–5672. https://doi.org/10.3168/jds.2010-3916.

Müller, C. B. M., S. Görs, M. Derno, A. Tuchscherer, K. Wimmers, A. Zeyner, and B. Kuhla. 2021. Differences between Holstein dairy cows in renal clearance rate of urea affect milk urea concentration and the relationship between milk urea and urinary nitrogen excretion. Sci. Total Environ. 755:143198. https://doi.org/10.1016/j.scitotenv.2020.143198.

Nayeri, S., M. Sargolzaei, M. K. Abo-Ismail, 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. 17:1–11. https://doi.org/10.1186/s12863-016-0386-1.

Nousiainen, J., K. Shingfield, and P. Huhtanen. 2004. Evaluation of milk urea nitrogen as a diagnostic of protein feeding. J. Dairy Sci. 87:386–398. https://doi.org/10.3168/jds.S0022-0302(04)73178-1.

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.

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 18:1–11. https://doi.org/10.1186/s12864-017-4263-8.

Pedrosa, V. B., F. S. Schenkel, S.-Y. Chen, H. R. Oliveira, T. M. Casey, M. G. Melka, and L. F. Brito. 2021. Genomewide 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.

Rajala-Schultz, P., W. Saville, G. Frazer, and T. Wittum. 2001. Association between milk urea nitrogen and fertility in Ohio dairy cows. J. Dairy Sci. 84:482–489. https://doi.org/10.3168/jds.S0022-0302(01)74498-0.

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. Couldrey, R. Hall, W. Li, A. Rhie, J. Ghurye, S. D. McKay, F. Thibaud-Nissen, J. Hoffman, B. M. Murdoch, W. M. Snelling, T. G. McDaneld, 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 Tassell, 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/giaa021.

Rzewuska, K., and T. Strabel. 2013. Genetic parameters for milk urea concentration and milk traits in Polish Holstein-Friesian cows. J. Appl. Genet. 54:473–482. https://doi.org/10.1007/s13353-013-0159-8.

Rzewuska, K., and T. Strabel. 2015. The genetic relationship between reproduction traits and milk urea concentration. Anim. Sci. Pap. Rep. 33:243–256.

Samoré, A., C. Romani, A. Rossoni, E. Frigo, O. Pedron, and A. Bagnato. 2007. Genetic parameters for casein and urea content in the Italian Brown Swiss dairy cattle. Italian Journal of Animal Science 6(sup1):201–203.

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. 2017a. Within-breed and multi-breed GWAS on imputed whole-genome sequence variants reveal candidate mutations affecting milk protein composition in dairy cattle. Genet. Sel. Evol. 49:1–16. https://doi.org/10.1186/s12711-017-0344-z.

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. 2017b. Within-breed and multi-breed GWAS on imputed whole-genome sequence variants reveal candidate mutations affecting milk protein composition in dairy cattle. Genet. Sel. Evol. 49:1–16. https://doi.org/10.1186/s12711-017-0344-z.

Sargolzaei, M., J. P. Chesnais, and F. S. Schenkel. 2014. A new approach for efficient genotype imputation using information from relatives. BMC Genomics 15:1–12. https://doi.org/10.1186/1471-2164-15-478.

Schepers, A., and R. Meijer. 1998. Evaluation of the utilization of dietary nitrogen by dairy cows based on urea concentration in milk. J. Dairy Sci. 81:579–584. https://doi.org/10.3168/jds.S0022-0302(98)75611-5.

Schopen, G., P. Koks, J. Van Arendonk, H. Bovenhuis, and M. Visker. 2009. Whole genome scan to detect quantitative trait loci for bovine milk protein composition. Anim. Genet. 40:524–537. https://doi.org/10.1111/j.1365-2052.2009.01880.x.

Soares, R., G. Vargas, T. Duffield, F. Schenkel, and E. Squires. 2021. Genome-wide association study and functional analyses for clinical and subclinical ketosis in Holstein cattle. J. Dairy Sci. 104:10076–10089. https://doi.org/10.3168/jds.2020-20101.

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.

Stoop, W., H. Bovenhuis, and J. Van Arendonk. 2007. Genetic parameters for milk urea nitrogen in relation to milk production traits. J. Dairy Sci. 90:1981–1986. https://doi.org/10.3168/jds.2006-434.

Suwannasing, R., M. Duangjinda, W. Boonkum, R. Taharnklaew, and K. Tuangsithtanon. 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.

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.

Uwizeye, A., I. J. de Boer, C. I. Opio, R. P. Schulte, A. Falcucci, G. Tempio, F. Teillard, F. Casu, M. Rulli, J. N. Galloway, A. Leip, J. W. Erisman, T. P. Robinson, H. Steinfeld, and P. J. Gerber. 2020. Nitrogen emissions along global livestock supply chains. Nat. Food 1:437–446. https://doi.org/10.1038/s43016-020-0113-y.

van den Berg, I., P. N. Ho, M. Haile-Mariam, P. R. Beatson, E. O'Connor, and J. E. Pryce. 2021. Genetic parameters of blood urea nitrogen and milk urea nitrogen concentration in dairy cattle managed in pasture-based production systems of New Zealand and Australia. Anim. Prod. Sci. 61:1801–1810. https://doi.org/10.1071/AN21049.

van den Berg, I., P. N. Ho, T. V. Nguyen, M. Haile-Mariam, I. M. MacLeod, P. R. Beatson, E. O'Connor, and J. E. Pryce. 2022. GWAS and genomic prediction of milk urea nitrogen in Australian and New Zealand dairy cattle. Genet. Sel. Evol. 54:15. https://doi.org/10.1186/s12711-022-00707-9.

VanRaden, P. M. 2008. Efficient methods to compute genomic predictions. J. Dairy Sci. 91:4414–4423. https://doi.org/10.3168/jds.2007-0980.

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.

Wiggans, G., 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.

ORCIDS

H. Atashi ⦿ https://orcid.org/0000-0002-6853-6608
Y. Chen ⦿ https://orcid.org/0000-0002-8593-4384
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