J. Dairy Sci. 105:6001–6020
https://doi.org/10.3168/jds.2021-21421
© 2022, 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/).

Role of *CSN2*, *CSN3*, and *BLG* genes and the polygenic background in the cattle milk protein profile

Nicolò Amalfitano,[1]* Lucio Flavio Macedo Mota,[1] Guilherme J. M. Rosa,[2] Alessio Cecchinato,[1] and Giovanni Bittante[1]
[1]Department of Agronomy, Food, Natural Resources, Animals and Environment (DAFNAE), University of Padova (Padua), 35020 Legnaro (PD), Italy
[2]Department of Animal and Dairy Sciences, University of Wisconsin-Madison, Madison 53706

ABSTRACT

To devise better selection strategies in dairy cattle breeding programs, a deeper knowledge of the role of the major genes encoding for milk protein fractions is required. The aim of the present study was to assess the effect of the *CSN2*, *CSN3*, and *BLG* genotypes on individual protein fractions (α_{S1}-CN, α_{S2}-CN, β-CN, κ-CN, β-LG, α-LA) expressed qualitatively as percentages of total nitrogen content (% N), quantitatively as contents in milk (g/L), and as daily production levels (g/d). Individual milk samples were collected from 1,264 Brown Swiss cows reared in 85 commercial herds in Trento Province (northeast Italy). A total of 989 cows were successfully genotyped using the Illumina Bovine SNP50 v.2 BeadChip (Illumina Inc.), and a genomic relationship matrix was constructed using the 37,519 SNP markers obtained. Milk protein fractions were quantified and the β-CN, κ-CN, and β-LG genetic variants were identified by reversed-phase HPLC (RP-HPLC). All protein fractions were analyzed through a Bayesian multitrait animal model implemented via Gibbs sampling. The effects of days in milk, parity order, and the *CSN2*, *CSN3*, and *BLG* genotypes were assigned flat priors in this model, whereas the effects of herd and animal additive genetic were assigned Gaussian prior distributions, and inverse Wishart distributions were assumed for the respective co-variance matrices. Marginal posterior distributions of the parameters of interest were compared before and after the inclusion of the effects of the 3 major genes in the model. The results showed that a high portion of the genetic variance was controlled by the major genes. This was particularly apparent in the qualitative protein profile, which was found to have a higher heritability than the protein fraction contents in milk and their daily yields. When the genes were included individually in the model, *CSN2* was the major gene controlling all the casein fractions except for κ-CN, which was controlled directly by the *CSN3* gene. The *BLG* gene had the most influence on the 2 whey proteins. The genetic correlations showed the major genes had only a small effect on the relationships between the protein fractions, but through comparison of the correlation coefficients of the proteins expressed in different ways they revealed potential mechanisms of regulation and competitive synthesis in the mammary gland. The estimates for the effects of the *CSN2* and *CSN3* genes on protein profiles showed overexpression of protein synthesis in the presence of the B allele in the genotype. Conversely, the β-LG B variant was associated with a lower concentration of β-LG compared with the β-LG A variant, independently of how the protein fractions were expressed, and it was followed by downregulation (or upregulation in the case of the β-LG B) of all other protein fractions. These results should be borne in mind when seeking to design more efficient selection programs aimed at improving milk quality for the efficiency of dairy industry and the effect of dairy products on human health.
Key words: milk protein profile, genetic parameter, major genes, genetic variants

INTRODUCTION

The protein content of milk is one of the major sources of variation in dairy industry profitability. In fact, together with fat, it is the most important determinant of milk price (Bailey et al., 2005; Atsbeha et al., 2016). However, it is not only the total content of this nutrient that affects industry profitability, but also its composition. The milk protein profile is composed principally of 4 casein fractions (α_{S1}-CN, α_{S2}-CN, β-CN, and κ-CN) and 2 major whey proteins (β-LG and α-LA). Previous studies have shown how qualitative and quantitative variations in the milk protein profile influence the coagulation pattern of milk and the final cheese yield (Wedholm et al., 2006; Cipolat-Gotet et al., 2018; Amalfitano et al., 2019). For example, increased

Received October 13, 2021.
Accepted February 28, 2022.
*Corresponding author: nicolo.amalfitano@unipd.it

contents of α_{S1}-CN and κ-CN were associated with the rapid formation of a firm coagulum, the basis of a high cheese yield; at the same time, higher contents and relative concentrations of α_{S2}-CN and β-LG negatively influenced the coagulation process, resulting in greater losses of nutrients in the whey.

Many factors lie behind these variations in the milk protein profile, but the main one is genetic polymorphisms. In fact, it is well known that the genes encoding for the 4 casein fractions are located in a short genetic linkage group on a 250-kb region of chromosome 6. These genes are indicated and distributed in this region as *CSN1S1*, *CSN2*, *CSN1S2*, and *CSN3*, and they encode for α_{S1}-CN, β-CN, α_{S2}-CN, and κ-CN, respectively. The whey protein genes are located on 2 different chromosomes: the *LGB* gene encoding for β-LG on chromosome 11, and the *LAA* gene encoding for α-LA on chromosome 5 (Martin et al., 2002; Farrell et al., 2004; Caroli et al., 2009). All these loci are characterized by high polymorphism, with each gene having 3 to 12 variants. Variants of the same protein fraction differ in their AA chains in the substitution of one or more AA. These substitutions can lead to physico-chemical changes in the protein properties and also to different levels of protein expression. The former is a qualitative change, such as β-CN B, which exhibits more net positive charges than the A^1 and A^2 variants. Beta-casein is one of the major components of casein micelles, and its less negatively charged B variant could be associated with a lower repulsive force between micelles during the clotting phase, facilitating their aggregation (Mckenzie et al., 1984). The latter is a quantitative change, and it occurs with the B variant of κ-CN. In fact, it is well known that κ-CN B is associated with the overexpression of this protein fraction in milk compared with the A variant (Ikonen et al., 1997; Bobe et al., 1999; Hallén et al., 2008). The higher level of κ-CN caused by this variant can lead to faster and stronger coagulation of the milk, as suggested in previous works. Protein genetic variants also seem to be connected to some effects on human health. In particular, β-CN A^1 and B have been identified as precursors of the bioactive peptide β-casomorphin 7 (BCM-7); this peptide has been investigated as risk factor for the increasing incidence of ischemic heart disease and type 1 diabetes in ecological studies (Elliott et al., 1999; Laugesen and Elliott, 2003; Birgisdottir et al., 2006) and for general gastrointestinal discomfort in in vitro and in vivo trials (Barnett et al., 2014; Fiedorowicz et al., 2016; Milan et al., 2020). The proteolytic release of BCM-7 has not been observed in variant A^2, which has been associated with other health benefits and also considered to improve milk digestibility (Kamiński et al., 2007; Sebastiani et al., 2020). However, there is still not enough clear scientific evidence to establish a cause-effect relationship between the consumption of BCM-7 and incidences of human diseases (De Noni et al., 2009; Küllenberg de Gaudry et al., 2019; Summer et al., 2020). On the contrary, some recent studies suggest possible beneficial effects of the oral administration of BCM-7 on human health through the reduction of oxidative stress (Yin et al., 2010; Zhang et al., 2013; Zhu et al., 2018).

These effects on the cheese-making properties and on human health have sparked interest in milk protein genetic variants and their inclusion in the selection criteria for the genetic improvement of dairy cattle populations. This has led to many studies investigating the effects of milk protein genotypes on protein fraction concentrations in milk (Hallén et al., 2008; Bonfatti et al., 2010; Gustavsson et al., 2014). However, before including these traits into selection indices, it is important to understand the magnitude of the additive genetic variance underlying each protein fraction and the genetic relationships among them. Other studies, therefore, have focused attention on the genetic variability of the protein fractions, and have estimated the heritabilities of these traits and the correlations between them (Schopen et al., 2009; Bonfatti et al., 2011; Gebreyesus et al., 2016). However, genetic variations in the milk protein profile are the result of the additive action of the protein encoding genes and the remaining polygenic background of the animal (Bobe et al., 1999). Not many studies have considered the genotypes of these genes as a correction factor to assess the importance of their role in the composition of the milk protein profile compared with the polygenic background. Moreover, in studying the protein fractions, the different ways of expressing them, such as their relative proportions of total milk nitrogen (% N), their total content in milk (g/L), and their daily production (g/d), have not been simultaneously taken into account.

The present work considered the yield, the total protein content, and the protein profile of milk from purebred Brown Swiss cows. The protein profile included all 6 major protein fractions and each was expressed as % N, g/L, and g/d. We were also able to identify the genotypes of 3 major milk protein genes (*CSN2*, *CSN3*, and *BLG*). The aims of this study were (1) to assess the relevance of the 3 major milk protein genes on the relevance of the additive genetic variance and heritabilities of each milk protein trait, and distinguish the influence of these genes from the remaining polygenic background; (2) to estimate the genetic correlations among the traits, and investigate the effects of the major genes on them; and (3) to quantify the additive

effects of the major gene alleles on the milk protein profile.

MATERIALS AND METHODS

Herds, Animals, and Sample Collection

This study is part of the multidisciplinary project Cowability-Cowplus. Milk samples were collected during the evening milking from 1,264 Italian Brown Swiss cows belonging to 85 commercial herds (maximum 15 cows/herd sampled). The herds were located in the province of Trento in northeast Italy. Details of the animals and the herds used in the study are reported in Cecchinato et al. (2015). Two aliquots were collected from the complete milking of each cow; the first 50-mL aliquot was taken to the Milk Quality Laboratory of the Trento Breeders Association (Trento, Italy) for milk composition analysis in accordance with the milk recording system, and the second 2.0-L aliquot was transported to the Milk Laboratory of the Department of Agronomy, Food, Natural Resources, Animals and Environment of the University of Padova (Legnaro, Padova, Italy) for processing into individual model cheeses and to record milk coagulation properties measurements within 20 h of collection. These data have been described in previous studies (Bittante et al., 2013, 2015; Cecchinato et al., 2013). In addition, 1.5 mL of milk was taken from the second aliquot and immediately stored at −80°C until analysis by reversed-phase HPLC (**RP-HPLC**).

Phenotypic and Genotypic Information

Milk composition traits were assessed from the 50-mL aliquot of fresh milk using a MilkoScan FT6000 apparatus (Foss Electric A/S). The analysis yielded the contents of 3 nitrogen compounds (protein, casein, and urea), fat, and lactose. Milk total protein was determined as the content of CP in the milk (i.e., the content of N multiplied by 6.38).

The contents of 4 casein fractions (α_{S1}-CN, α_{S2}-CN, β–CN, and κ-CN) and 2 whey proteins (α-LA and β-LG) were determined by validated RP-HPLC, as described by Bonfatti et al. (2008); this method enabled also the determination of the β-CN (A^1, A^2, and B), κ-CN (A and B), and β-LG (A and B) genetic variants, but not of the α_{S1}-CN and α_{S2}-CN. Total casein was defined as the sum of the 4 casein fractions, and total whey protein as the sum of α-LA and β-LG. To combine the data obtained by Fourier-transform infrared spectroscopy (total protein, casein, and urea) with the data obtained by RP-HPLC (individual protein fractions), the HPLC protein fractions were divided by the sum of all HPLC caseins and multiplied by the Fourier-transform infrared spectroscopy milk casein contents. Each milk protein fraction was then expressed as (1) the percentage of the total milk nitrogen content (% N), (2) g/L of milk, and (3) g/d of lactation.

A total of 1,050 cows were genotyped using the Illumina Bovine SNP50 v.2 BeadChip (Illumina Inc.). Autosomal SNP markers exhibiting minor allele frequencies lower than 0.05, a significant departure from Hardy–Weinberg proportions ($P \leq 10^{-5}$), and a call rate for markers and samples lower than 0.90 were removed. A total of 989 animals with phenotypic information and 37,519 SNP markers remained in the genomic data set.

Statistical Analysis

Phenotypic data for milk protein fractions outside the interval between ± 3.5 standard deviations from the mean for each herd were removed from the data set. The normality of the data was assessed with the Shapiro-Wilk test (data not shown).

The possible population structure of the genotyped animals was assessed by principal component analysis using the gaston R package (Perdry and Dandine-Roulland, 2018). The animals clustered into 2 groups. The first 2 components were included in the model as co-variables to correct for potential population effect.

The (co)variance components and genetic parameters were estimated by Bayesian inference in a multitrait animal model considering genomic information on (1) the milk protein fractions (α_{S1}-CN, α_{S2}-CN, β-CN, κ-CN, α-LA, and β-LG) expressed as % N; (2) total protein and milk protein fractions expressed as g/L, and (3) milk yield in kg/d and milk protein fractions expressed as g/d, separately, according to the following model:

$$\mathbf{y} = \mathbf{X\beta} + \mathbf{Wh} + \mathbf{Za} + \mathbf{e},$$

where $\mathbf{y}$ is the matrix for total protein, milk yield, and milk protein fractions expressed in g/L, g/d, and % N; $\boldsymbol{\beta}$ is the vector of fixed effects including days in milk (6 classes: class 1, less than 60 d; class 2, 60–120 d; class 3, 121–180 d; class 4, 181–240 d; class 5, 241–300 d; class 6, more than 300 d), parity of the cow (4 classes: 1, 2, 3, ≥4), and the first 2 principal components as linear co-variables to correct for population substructures; $\mathbf{h}$ is the random effect of herd, and $\mathbf{a}$ is the vector of additive genetic effects; $\mathbf{X}$, $\mathbf{W}$, and $\mathbf{Z}$ are incidence matrices relating $\mathbf{y}$ to effects in $\boldsymbol{\beta}$, $\mathbf{h}$, and $\mathbf{a}$, respectively. This was the base animal model; in a second model the β-CN (*CSN2*), κ-CN (*CSN3*), and β-LG (*BLG*) genotypes were also included separately as fixed effects to

estimate their individual effects; in a third model, these genotypes were included together as fixed effects.

The additive genetic and residual effects were assumed to be normally distributed, described as $\mathbf{a}=\{a_i\}\sim MN(\mathbf{0},\Sigma_{\mathbf{a}}\otimes\mathbf{G})$ and $\mathbf{e}=\{e_{ij}\}\sim MN(\mathbf{0},\Sigma_{\mathbf{e}}\otimes\mathbf{I})$, respectively, where

$$\Sigma_{\mathbf{a}}=\begin{bmatrix}\sigma^2_{a1} & \cdots & \sigma_{a1,n}\\ \vdots & \ddots & \vdots\\ \sigma_{a1,n} & \cdots & \sigma^2_{an}\end{bmatrix}$$

and

$$\Sigma_{\mathbf{e}}=\begin{bmatrix}\sigma^2_{e1} & \cdots & \sigma_{e1,n}\\ \vdots & \ddots & \vdots\\ \sigma_{e1,n} & \cdots & \sigma^2_{en}\end{bmatrix}$$

are the additive genetic and the residual (co)variance matrices, $\mathbf{I}$ is an identity matrix, and $\mathbf{G}$ is a genomic relationship matrix. The $\mathbf{G}$ matrix was constructed as follows (VanRaden, 2008): $\mathbf{G}=\mathbf{MM}'q$ where $\mathbf{M}$ is the SNP matrix coded as 0, 1, and 2 for genotypes AA, AB, and BB, respectively, and q is a weighting factor given as $q=1\Big/\sum_{j=1}^{n}2p_j(1-p_j)$, where p_j is the second allele frequency of the jth SNP marker.

Given the model specifications, the distribution of $\mathbf{y}$ given the random effects and residual (co)variance matrix is multivariate normal, described as $\mathbf{y}\mid\boldsymbol{\beta},\mathbf{h},\mathbf{a},\Sigma_{\mathbf{e}}\sim MVN(\mathbf{X}\boldsymbol{\beta}+\mathbf{Wh}+\mathbf{Za},\ \Sigma_{\mathbf{e}}\otimes\mathbf{I})$. This was the adopted sampling model on the first stage of the Bayesian hierarchical implementation. On a second stage, the previously described multivariate normal distributions were assumed for $\mathbf{h}$ and $\mathbf{a}$. To complete the Bayesian model specification, a flat prior was assumed for the fixed effects [i.e., $p(\beta)\propto$ constant], and inverse Wishart distributions were assumed for the genetic and residual (co)variance matrices, expressed as $\Sigma_{\mathbf{a}}\mid v_a,S_a\sim IW(v_a,v_aS_a)$ and $\Sigma_{\mathbf{e}}\mid v_e,S_e\sim IW(v_e,v_eS_e)$, respectively.

Samples of the joint posterior distribution of the model parameters were obtained by Gibbs sampling algorithm implemented in the GIBBS1F90 program (Misztal et al., 2018). The posterior analysis consisted of a single chain of 1,000,000 cycles after a burn-in of 100,000 iterations, with samples stored at every 10 cycles. Convergence was evaluated through visual inspection and using the Bayesian Output Analysis (Smith, 2007) and Geweke test (Geweke, 1992).

RESULTS

Descriptive Statistics of the Protein Profile

Table 1 presents descriptive statistics of the Brown Swiss milk yield, total protein, and protein profile. Briefly, the population presented an average milk yield of almost 25 kg/d, with a CV of 29%, and a total protein concentration of 36.7 g/L, with a CV of 11%. The most abundant protein fraction was β-CN (32.1% N), followed by α_{S1}-CN (25.8% N). The relative concentrations of α_{S2}-CN and κ-CN were smaller and similar (9.0 and 9.8% N, respectively). The CV of the casein fractions with higher relative concentrations were 6 to 7%, and around 9 to 11% for those with the lowest concentrations. The β-LG had a relative concentration of 8.6% N and a CV of almost 15%, and α-LA about 2.4% N with a CV of 17%. To calculate their contents in milk (g/L), the protein fractions expressed as % N were multiplied by the total protein content of each milk sample. Regarding the added variability of total protein, the CV of the protein fractions were higher when expressed as g/L than when expressed as % N. This was also the case when the proteins in g/L were multiplied by the milk yield to obtain their daily production levels (g/d). With the HPLC method validated by Bonfatti et al. (2008) it was possible to identify the following genotypes: for the β-CN we found A^1A^2 (96 cows), A^1B (22 cows), A^2A^2 (592 cows), A^2B (253 cows), and BB (18 cows) genotypes; for the κ-CN AA (48 cows), AB (354 cows), and BB (579 cows); for the β-LG AA (107 cows), AB (424 cows), and BB (450 cows).

Genetic Parameters and the Role of the Major Genes

Table 2 presents the posterior means and standard deviations (between parentheses) of additive genetic variance (σ^2_a), and intraherd heritability (h^2) for the relative proportions of the protein fractions (% N) obtained from an animal model before (base model) and after correction for the *CSN2*, *CSN3*, and *BLG* gene polymorphisms. The genes were included in the model one at a time (+ *CSN2,* + *CSN3,* or + BLG) or all together (+ *CSN2-CSN3-BLG*). The main effect of correction for the major gene genotypes was a reduction in the genetic variance ($\Delta\%\sigma^2_a$) and, as a consequence, in the heritability ($\Delta\%h^2$) for all the traits considered. This reduction represents the quota of the genetic variance explained by the major genes, whereas the remaining portion is that controlled by the residual polygenic background of the cow. It is worth mentioning that, compared with the original formula of the heritability

$\left(h^2 = \frac{\sigma_a^2}{\sigma_a^2 + \sigma_e^2}\right)$, the correction for the 3 major genes reduces the genetic variance present at the numerator and at the denominator of the equation, resulting in a new reduced heritability formula. This leads to a lower reduction of the h^2 compared with the reduction of the genetic variance.

This qualitative protein profile showed a moderate-high heritability, which was slightly higher for the casein fractions than for the whey proteins. In particular, the genetic variance in the casein fractions explained by the 3 major genes together was very high, ranging from −73% for α_{S2}-CN to −91% for β-CN, and caused a reduction in the heritabilities ranging from −45% for α_{S2}-CN and κ-CN to −68% for β-CN. This means that the remaining polygene effect is responsible for only 9% of the genetic variance in β-CN, and 27% of the genetic variance in α_{S2}-CN. The 3 major genes explained the same quota of genetic variance in both whey proteins (about −68%), but after correction β-LG showed a reduction in heritability of −42%, and α-LA a reduction of only −8%. This is explained by a considerable drop also in the residual phenotypic variance after inclusion of the major genes in the model. When included in the model individually, the sum of the genetic variance explained by the single genes was equal to the genetic variance explained when they were included together. Correction for the *CSN2* genotype resulted in a greater reduction in the genetic variance in the casein fractions than the other genes did in almost all the fractions, except for κ-CN, which was influenced more by its encoding *CSN3* gene. The whey protein α-LA was also influenced more by the *CSN2* gene, whereas β-LG was mainly affected by its encoding *BLG* gene.

As can be seen from Table 3, which shows the results for milk contents (g/L), total protein had a moderate-high heritability, but seemed to be much less affected by the major genes and much more controlled by the residual polygenic component compared with its composition. In fact, −22% of the genetic variance was explained by the genotypes, and only −14% of heritability loss. Comparing the quantitative and qualitative protein profiles, we found the initial heritability (no major genes included) of the protein fractions in g/L to be similar to their relative proportions (except β-CN, from 0.79–0.59), but the major genes had less of an influence on the genetic variance, due to an effect of dilution after the multiplication for total protein content. The only exception was the κ-CN content, for which the portion of the genetic variance explained by the major genes was similar to its relative proportion (−86 vs. −84%). The other caseins ranged from −48% for α_{S1}-CN to −69% for β-CN, whereas the whey proteins ranged from −44% for β-LG to −55% for α-LA. The heritability estimates were also less affected by the correction for the major genes. When the genes were tested individually, *CSN2* still had the greatest influence on all the casein fractions except κ-CN, for which the *CSN3* gene was more important. The *BLG* gene, on the other hand, was responsible for more genetic variance in both whey proteins than the other genes.

Table 1. Descriptive statistics for milk protein profile in Brown Swiss cattle

Trait[1]	n	Mean	SD	Minimum	Maximum	CV (%)
Protein fraction proportions (% N)						
α_{S1}-CN	933	25.75	1.64	21.48	30.04	6.4
α_{S2}-CN	907	9.00	1.07	6.54	11.25	9.7
β-CN	919	32.07	2.14	27.53	38.21	6.7
κ-CN	926	9.76	1.07	6.01	12.35	11.0
α-LA	905	2.37	0.41	1.38	3.63	17.3
β-LG	901	8.64	1.28	5.44	12.25	14.8
Milk contents (g/L)						
Total protein	951	36.70	4.10	25.80	58.10	11.2
α_{S1}-CN	928	9.40	1.11	6.53	12.41	11.8
α_{S2}-CN	926	3.32	0.54	2.08	4.81	16.3
β-CN	959	11.75	1.32	8.48	15.43	11.2
κ-CN	903	3.73	0.83	1.86	5.20	15.5
α-LA	943	0.85	0.15	0.43	1.28	17.6
β-LG	925	3.13	0.62	1.45	4.94	19.8
Daily yields (g/d)						
Milk yield (kg/d)	952	24.9	7.2	10.1	45.3	29.1
α_{S1}-CN	956	228.0	62.1	80.9	412.6	27.3
α_{S2}-CN	938	82.5	24.3	27.2	150.3	29.5
β-CN	920	288.7	74.6	103.8	497.6	25.8
κ-CN	949	84.0	24.7	25.5	150.7	29.4
α-LA	916	21.1	7.5	4.0	43.6	35.4
β-LG	929	76.9	23.8	15.1	148.0	31.0

[1]g/d: grams of protein secreted per day; % N: percentage of nitrogen; g/L: grams per L of milk.

The daily yields are shown in Table 4. The heritability of milk yield was lower (0.14), and the genetic variance explained by the 3 major genes was slightly higher than that of total protein $\left(\Delta\%\sigma_a^2 \text{ of } -27 \text{ vs. } -22\%\right)$. The daily yields of the protein fractions had lower heritabilities than their contents in milk, but the quotas of the genetic variance explained by the major genes together were similar in size. The effect of the genes on the heritability was again lower, indicating an increase in the residual phenotypic variance explained by them. With regard to the individual effects of the major genes, again the *CSN2* gene exerted the greatest influence on caseins, and the *BLG* gene the greatest influence on whey proteins. The *BLG* and *CSN3* genes explained a similar portion of genetic variance only in κ-CN.

Genetic Correlations

Figure 1 shows the additive genetic correlations between the individual protein fractions and milk yield (A) or total protein (B). The negative correlations are indicated in red, the positive correlations in green. The darker colors indicate the correlations before the inclusion of the 3 major genes in the model, the paler colors the correlations after inclusion. The correlations

Table 2. Posterior means (SD) of additive genetic variances $\left(\sigma_a^2\right)$ and heritabilities (h^2) for milk protein fraction proportions of total nitrogen (% N) using animal models not considering (base model) or considering the β-CN (*CSN2*), κ-CN (*CSN3*), and β-LG (*BLG*) polymorphisms as fixed effects one at a time or all together (*CSN2-CSN3-BLG*)

		Genetic parameter		Variation[2]	
Trait[1]	Animal model	σ_a^2	h^2	$\Delta\%\sigma_a^2$	$\Delta\%h^2$
α_{S1}-CN					
	Base model	1.33 (0.057)	0.59 (0.018)		
	+ CSN2	0.88 (0.064)	0.45 (0.018)	−34	−24
	+ CSN3	0.98 (0.078)	0.47 (0.039)	−26	−20
	+ BLG	1.03 (0.050)	0.49 (0.033)	−23	−17
	+ CSN2-CSN3-BLG	0.23 (0.043)	0.23 (0.045)	−83	−61
α_{S2}-CN					
	Base model	0.30 (0.037)	0.31 (0.019)		
	+ CSN2	0.22 (0.033)	0.24 (0.026)	−27	−23
	+ CSN3	0.22 (0.014)	0.27 (0.037)	−25	−13
	+ BLG	0.24 (0.020)	0.28 (0.052)	−21	−10
	+ CSN2-CSN3-BLG	0.08 (0.026)	0.17 (0.010)	−73	−45
β-CN					
	Base model	3.43 (0.102)	0.79 (0.017)		
	+ CSN2	1.62 (0.102)	0.49 (0.044)	−53	−38
	+ CSN3	2.63 (0.096)	0.63 (0.055)	−23	−20
	+ BLG	2.92 (0.031)	0.71 (0.013)	−15	−10
	+ CSN2-CSN3-BLG	0.31 (0.037)	0.25 (0.018)	−91	−68
κ-CN					
	Base model	0.67 (0.102)	0.60 (0.020)		
	+ CSN2	0.49 (0.083)	0.53 (0.023)	−33	−12
	+ CSN3	0.41 (0.018)	0.41 (0.030)	−39	−32
	+ BLG	0.59 (0.043)	0.59 (0.030)	−12	−2
	+ CSN2-CSN3-BLG	0.11 (0.021)	0.33 (0.031)	−84	−45
α-LA					
	Base model	0.03 (0.001)	0.25 (0.008)		
	+ CSN2	0.02 (0.005)	0.25 (0.015)	−33	−2
	+ CSN3	0.03 (0.003)	0.24 (0.021)	−10	−3
	+ BLG	0.02 (0.011)	0.24 (0.041)	−23	−3
	+ CSN2-CSN3-BLG	0.01 (0.003)	0.23 (0.025)	−67	−8
β-LG					
	Base model	0.68 (0.080)	0.53 (0.014)		
	+ CSN2	0.55 (0.081)	0.46 (0.021)	−19	−13
	+ CSN3	0.55 (0.043)	0.49 (0.049)	−19	−8
	+ BLG	0.48 (0.040)	0.42 (0.037)	−29	−21
	+ CSN2-CSN3-BLG	0.22 (0.026)	0.31 (0.036)	−68	−42

[1]Milk protein fractions are expressed as percentage of the total N content in milk.

[2]$\Delta\%\sigma_a^2$ = proportion of the genetic variance lost from the base model after the inclusion of the 3 major genes one at a time or all together; $\Delta\%h^2$ = proportion of the heritability lost from the base model after the inclusion of the 3 major genes one at a time or all together.

Table 3. Posterior means (SD) of additive genetic variances $\left(\sigma_a^2\right)$ and heritabilities (h^2) for total protein and protein fraction contents in milk (g/L) using animal models not considering (base model) or considering the β-CN (*CSN2*), κ-CN (*CSN3*), and β-LG (*BLG*) polymorphisms as fixed effects one at a time or all together (*CSN2-CSN3-BLG*)

		Genetic parameter		Variation[2]	
Trait[1]	Animal model	σ_a^2	h^2	$\Delta\%\sigma_a^2$	$\Delta\%h^2$
Total protein					
	Base model	0.43 (0.075)	0.44 (0.068)		
	+ CSN2	0.41 (0.079)	0.42 (0.074)	−5	−5
	+ CSN3	0.40 (0.081)	0.43 (0.074)	−6	−2
	+ BLG	0.39 (0.083)	0.41 (0.080)	−9	−7
	+ CSN2-CSN3-BLG	0.34 (0.083)	0.38 (0.071)	−22	−14
α_{S1}-CN					
	Base model	0.52 (0.013)	0.57 (0.021)		
	+ CSN2	0.40 (0.026)	0.51 (0.013)	−23	−11
	+ CSN3	0.45 (0.034)	0.53 (0.021)	−13	−7
	+ BLG	0.46 (0.032)	0.51 (0.018)	−12	−11
	+ CSN2-CSN3-BLG	0.27 (0.014)	0.41 (0.032)	−48	−28
α_{S2}-CN					
	Base model	0.06 (0.008)	0.29 (0.011)		
	+ CSN2	0.04 (0.003)	0.26 (0.025)	−33	−10
	+ CSN3	0.06 (0.024)	0.26 (0.017)	−8	−10
	+ BLG	0.06 (0.034)	0.27 (0.026)	−8	−7
	+ CSN2-CSN3-BLG	0.03 (0.011)	0.21 (0.029)	−50	−28
β-CN					
	Base model	0.84 (0.007)	0.59 (0.022)		
	+ CSN2	0.49 (0.018)	0.47 (0.046)	−42	−20
	+ CSN3	0.70 (0.029)	0.56 (0.012)	−17	−5
	+ BLG	0.75 (0.024)	0.53 (0.039)	−11	−10
	+ CSN2-CSN3-BLG	0.26 (0.025)	0.38 (0.021)	−69	−36
κ-CN					
	Base model	0.36 (0.009)	0.56 (0.012)		
	+ CSN2	0.26 (0.007)	0.47 (0.011)	−28	−16
	+ CSN3	0.22 (0.014)	0.44 (0.036)	−39	−22
	+ BLG	0.29 (0.047)	0.52 (0.018)	−19	−7
	+ CSN2-CSN3-BLG	0.05 (0.013)	0.31 (0.045)	−86	−45
α-LA					
	Base model	0.004 (0.0003)	0.23 (0.009)		
	+ CSN2	0.004 (0.0003)	0.22 (0.021)	−14	−2
	+ CSN3	0.004 (0.0007)	0.23 (0.035)	−16	−2
	+ BLG	0.003 (0.0007)	0.21 (0.034)	−25	−9
	+ CSN2-CSN3-BLG	0.002 (0.0010)	0.20 (0.042)	−55	−13
β-LG					
	Base model	0.18 (0.003)	0.53 (0.011)		
	+ CSN2	0.16 (0.017)	0.48 (0.014)	−11	−9
	+ CSN3	0.16 (0.026)	0.49 (0.011)	−11	−8
	+ BLG	0.14 (0.025)	0.47 (0.023)	−22	−11
	+ CSN2-CSN3-BLG	0.10 (0.008)	0.38 (0.031)	−44	−28

[1]Milk protein fractions are expressed as a percentage of the total N content in milk.

[2]$\Delta\%\sigma_a^2$ = proportion of the genetic variance lost from the base model after the inclusion of the 3 major genes one at a time or all together; $\Delta\%h^2$ = proportion of the heritability lost from the base model after the inclusion of the 3 major genes one at a time or all together.

between the protein fractions in g/L and milk yield were all negative, ranging from −0.25 to −0.41, except for α_{S1}-CN, which had a correlation of 0.28 (Figure 1A). As shown in Figure 1B, total protein correlated negatively with almost all the protein fractions in % N, except κ-CN and β-LG, ranging from −0.23 for α_{S2}-CN to −0.48 for α-LA; the positive correlations were around 0.20. Correction for the genotypes of the genes had a small effect on the magnitude of genetic correlations, in particular increasing or decreasing the entity of the correlation but not changing their initial sign. In the case of the correlation between total protein and κ-CN in % N, correction for the major genes almost canceled out the estimate, making the proportion of κ-CN independent of the total protein content in milk.

Table 5 shows the additive genetic correlations between the protein fractions before (above the diagonal) and after (below the diagonal) correction for the geno-

types of the *CSN2*, *CSN3,* and *BLG* genes. Regarding the relative proportions of the protein fractions, we can see that almost all the proteins were negatively correlated with β-CN, except α-LA (0.28). All the others correlated positively with each other. Only κ-CN was negatively correlated with β-LG (−0.39). In terms of milk contents, likewise almost all the protein fractions were negatively correlated with β-CN, except β-LG (0.18). For daily yields, however, all the correlations were positive, except between κ-CN and β-CN (−0.36). After correction for the major genes, the matrix of the protein fraction correlations did not change substantially, with only a few exceptions, for example, the correlation between β-CN and α_{S1}-CN in g/L changed from −0.31 to 0.21. There were other potential differences, but the standard deviations were often high.

Table 4. Posterior means (SD) of additive genetic variances $\left(\sigma_a^2\right)$ and heritabilities (h^2) for milk yield (kg/d) and milk protein fraction daily yields (g/d) using animal models not considering (base model) or considering the β-CN (*CSN2*), κ-CN (*CSN3*), and β-LG (*BLG*) polymorphisms as fixed effects one at a time or all together (*CSN2-CSN3-BLG*)

		Genetic parameter		Variation[2]	
Trait[1]	Animal model	σ_a^2	h^2	$\Delta\%\sigma_a^2$	$\Delta\%h^2$
Milk yield					
	Base model	2.99 (0.36)	0.14 (0.019)		
	+ CSN2	2.71 (0.41)	0.13 (0.054)	−9	−7
	+ CSN3	2.74 (0.33)	0.13 (0.051)	−8	−7
	+ BLG	2.72 (0.28)	0.13 (0.041)	−9	−12
	+ CSN2-CSN3-BLG	2.18 (0.59)	0.12 (0.027)	−27	−26
α_{S1}-CN					
	Base model	325 (11.3)	0.26 (0.016)		
	+ CSN2	235 (8.3)	0.24 (0.096)	−28	−8
	+ CSN3	297 (15.3)	0.25 (0.057)	−9	−6
	+ BLG	277 (11.1)	0.24 (0.067)	−15	−10
	+ CSN2-CSN3-BLG	159 (11.7)	0.20 (0.054)	−51	−23
α_{S2}-CN					
	Base model	53 (1.3)	0.21 (0.013)		
	+ CSN2	31 (0.2)	0.20 (0.009)	−41	−5
	+ CSN3	43 (1.4)	0.20 (0.069)	−18	−5
	+ BLG	49 (1.7)	0.20 (0.057)	−4	−5
	+ CSN2-CSN3-BLG	19 (1.4)	0.18 (0.048)	−63	−14
β-CN					
	Base model	590 (12.4)	0.22 (0.009)		
	+ CSN2	359 (12.6)	0.20 (0.055)	−39	−9
	+ CSN3	565 (14.9)	0.22 (0.084)	−4	−2
	+ BLG	490 (15.8)	0.21 (0.088)	−17	−3
	+ CSN2-CSN3-BLG	235 (17.4)	0.19 (0.052)	−60	−14
κ-CN					
	Base model	148 (9.7)	0.29 (0.017)		
	+ CSN2	124 (2.2)	0.27 (0.036)	−16	−7
	+ CSN3	107 (2.0)	0.26 (0.081)	−27	−12
	+ BLG	105 (2.5)	0.28 (0.061)	−29	−2
	+ CSN2-CSN3-BLG	41 (2.4)	0.23 (0.061)	−72	−21
α-LA					
	Base model	6 (1.10)	0.24 (0.039)		
	+ CSN2	5 (0.01)	0.22 (0.014)	−15	−8
	+ CSN3	5 (0.15)	0.22 (0.053)	−13	−10
	+ BLG	4 (0.16)	0.22 (0.047)	−24	−7
	+ CSN2-CSN3-BLG	3 (0.60)	0.18 (0.052)	−52	−25
β-LG					
	Base model	100 (2.0)	0.39 (0.017)		
	+ CSN2	81 (1.2)	0.37 (0.031)	−19	−5
	+ CSN3	93 (1.8)	0.36 (0.089)	−6	−7
	+ BLG	77 (2.0)	0.35 (0.067)	−23	−11
	+ CSN2-CSN3-BLG	52 (2.6)	0.30 (0.056)	−48	−23

[1]Milk protein fractions are expressed as percentage of the total N content in milk.

[2]$\Delta\%\sigma_a^2$ = proportion of the genetic variance lost from the base model after the inclusion of the 3 major genes one at a time or all together; $\Delta\%h^2$ = proportion of the heritability lost from the base model after the inclusion of the 3 major genes one at a time or all together.

Effects of the β-CN Genotypes on Milk Yield and the Protein Profile

Figure 2 shows the contrasts between the estimated effects of the *CSN2* genotypes on milk protein fractions expressed as % N (A), g/L (B), and g/d (C). In particular, these contrasts reflect the differences between all the genotypes found in the present population (BB, A^1B, and A^1A^2) and the A^2A^2 genotype. This genotype was taken as the reference in the contrasts because A^2 had been identified as the ancestral variant of the *CSN2* gene. Negative and positive differences are shown in red and green, respectively. The asterisks indicate that almost all the contrasts differed significantly from zero.

As Figure 2 shows, the BB genotype of *CSN2* significantly increased the β-CN content compared with the reference A^2A^2 in terms of relative proportions, milk contents, and daily yields. Milk yield was also higher with the BB genotype (+1.0 kg/d). Correspondingly, the BB genotype reduced the proportions, milk contents, and daily productions of all the other protein fractions and the total protein content of the milk. Their heterozygote A^2B was generally intermediate between the 2 homozygotes, indicating an additive effect of the alleles. However, some of the values of the heterozygote also indicated nonadditive effects, as in the daily yields of α_{S1}-CN, where there was a significant overdominance of the B allele linked to an overdominance on daily milk yield.

Unfortunately, in the population studied only 4 animals had the A^1A^1 genotype for the *CSN2* gene, and they were therefore not included in the statistical analysis. However, the heterozygotes A^1B and A^1A^2 can help in comparing the A^1 variant with the A^2 and B variants. The contrasts between A^2B or A^1A^2 and the

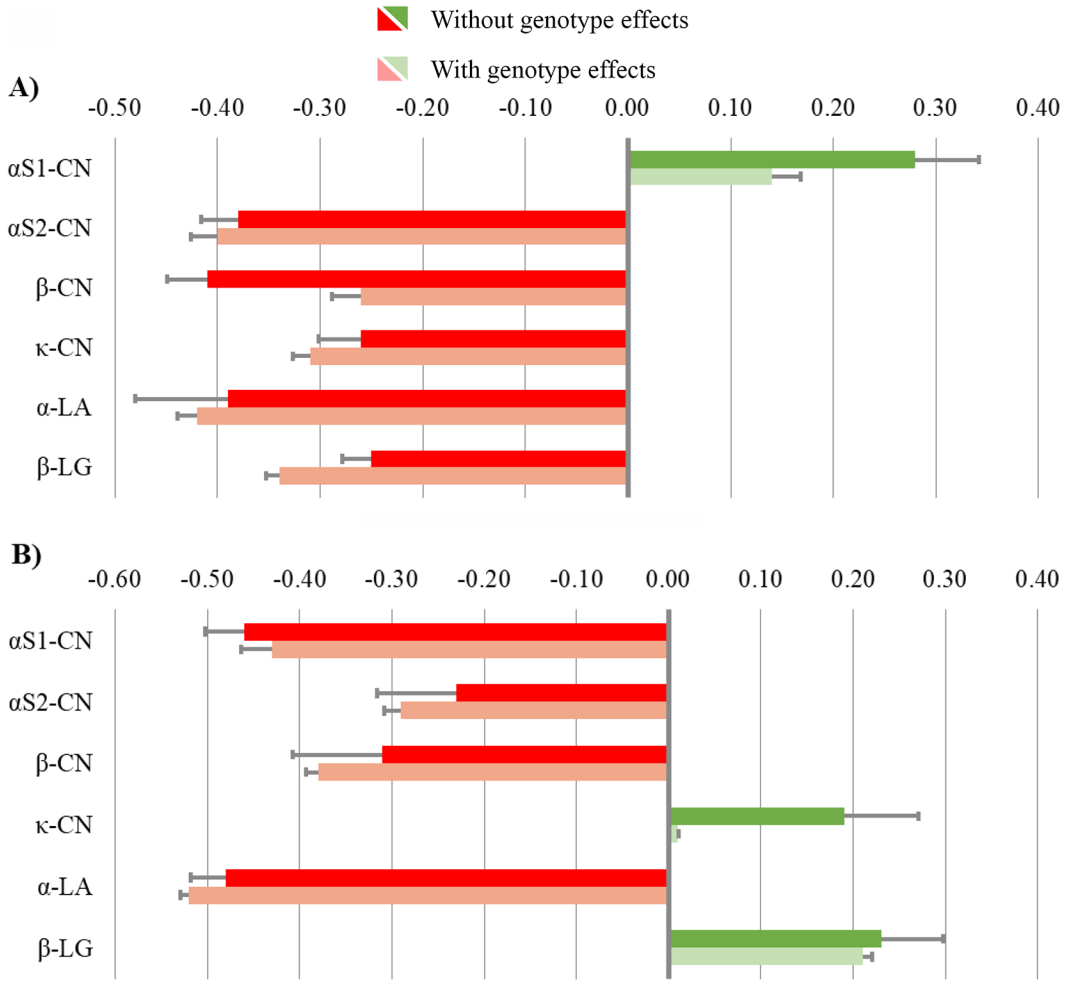

Figure 1. Posterior means and SD of additive genetic correlations (A) between the daily milk yield (kg/d) and the milk protein fractions (g/L) and (B) between the total milk protein content (g/L) and the proportions of milk protein fractions on total nitrogen (% N) estimated without (dark color) and considering (light color) genotype effects for *CSN2*, *CSN3*, and *BLG* genes in Brown Swiss.

reference A^2A^2 reveal the effect of substituting the B allele with the A^1. Starting with the proportion of α_{S1}-CN, shown in Figure 2A, substituting the B allele with the A^1 shifted the contrast from negative to slightly positive, and put the A^1 variant closer to the A^2 variant. Moreover, the substitution in both the A^2 alleles in the A^1B contrast yielded a value intermediate between the other 2 contrasts, indicating an additive effect between the alleles. This analysis shows the A^1 variant to behave similarly to the A^2 in decreasing the relative proportion and content of β-CN in milk and increasing all the other fractions. When the protein fractions are expressed as daily yield, the A^1 variant seemed always to have the same effect as the A^2 variant. However, in this case A^1B was often the lowest genotype, probably because it was characterized by a lower milk yield than the others.

Effects of the κ-CN Genotypes on Milk Yield and the Protein Profile

Figure 3 shows the contrasts of the estimated effects of the *CSN3* genotypes on the milk protein fractions expressed as % N (A), g/L (B), and g/d (C). In particular, these contrasts show the differences between the AA or AB genotypes and the BB genotype, taken as the reference. The asterisks indicate that almost all the contrasts were significantly different from zero.

Looking at Figure 3, it is clear that the κ-CN relative proportion and content in milk, and daily yield, and also the total protein content were lower in the AA genotype of κ-CN than in the BB genotype. Conversely, all the other protein fractions, particularly β-CN, were higher in the AA genotype, and there was no difference in terms of milk yield. Analysis of the contrasts between the AB and BB genotypes revealed that the difference between the heterozygote and the homozygote was often close to half the difference between the homozygotes, placing the AB genotype in an intermediate position between the homozygotes, suggesting that the alleles show additive effects.

Effects of the β-LG Genotypes on Milk Yield and the Protein Profile

Figure 4 presents the contrasts of the estimated effects of the *BLG* genotypes on the milk protein fractions expressed as % N (A), g/L (B), and g/d (C). In particular, these contrasts show the differences between the AA or AB genotypes and the BB genotype, taken as the reference. The asterisks indicate that most of the contrasts were significantly different from zero.

Looking at Figure 4, we can see that the genotypes influence many protein fractions in different ways depending on how they are expressed. In general, the relative proportions, contents in milk, and daily yields of β-LG and α_{S2}-CN were higher in the AA genotype than in the BB genotype. Conversely, the proportion of all the other protein fractions, except β-CN, were lower. As the BB genotype was associated with an increase in the total protein content of milk, the β-CN content was also higher, and only the α-LA content was signifi-

Table 5. Posterior mean (SD) for the highest density posterior interval of additive genetic correlations for the milk protein fractions estimated without (above diagonal) and considering (below diagonal) genotype effects for *CSN2*, *CSN3*, and *BLG* genes in Brown Swiss

Trait	α_{S1}-CN	α_{S2}-CN	β-CN	κ-CN	α-LA	β-LG
Protein fraction proportion (% N)						
α_{S1}-CN	—	0.62 (0.063)	−0.73 (0.005)	0.59 (0.043)	0.41 (0.024)	0.26 (0.017)
α_{S2}-CN	0.37 (0.097)	—	−0.49 (0.047)	0.39 (0.064)	0.69 (0.022)	0.25 (0.019)
β-CN	−0.58 (0.021)	−0.41 (0.176)	—	−0.79 (0.027)	0.28 (0.017)	−0.25 (0.009)
κ-CN	0.57 (0.016)	0.42 (0.102)	−0.62 (0.016)	—	0.16 (0.026)	−0.39 (0.015)
α-LA	0.48 (0.091)	0.57 (0.147)	0.08 (0.024)	−0.09 (0.013)	—	0.41 (0.022)
β-LG	0.11 (0.015)	0.26 (0.065)	−0.29 (0.051)	−0.44 (0.077)	0.48 (0.023)	—
Milk contents (g/L)						
α_{S1}-CN	—	0.74 (0.051)	−0.31 (0.009)	0.55 (0.025)	0.45 (0.044)	0.57 (0.009)
α_{S2}-CN	0.78 (0.072)	—	−0.25 (0.029)	0.36 (0.043)	0.73 (0.201)	0.35 (0.017)
β-CN	0.21 (0.009)	−0.37 (0.09)	—	−0.49 (0.034)	−0.17 (0.03)	0.18 (0.006)
κ-CN	0.69 (0.059)	0.48 (0.065)	−0.45 (0.047)	—	0.23 (0.045)	0.34 (0.009)
α-LA	0.59 (0.075)	0.79 (0.072)	−0.11 (0.037)	0.33 (0.021)	—	0.44 (0.064)
β-LG	0.69 (0.094)	0.46 (0.054)	0.40 (0.053)	0.61 (0.041)	0.57 (0.013)	—
Daily yield (g/d)						
α_{S1}-CN	—	0.81 (0.037)	0.44 (0.019)	0.52 (0.013)	0.64 (0.033)	0.61 (0.013)
α_{S2}-CN	0.70 (0.005)	—	0.38 (0.019)	0.41 (0.093)	0.71 (0.0653)	0.50 (0.055)
β-CN	0.41 (0.007)	0.36 (0.012)	—	−0.36 (0.0219)	0.46 (0.023)	0.31 (0.038)
κ-CN	0.50 (0.031)	0.40 (0.029)	−0.35 (0.017)	—	0.11 (0.015)	0.38 (0.029)
α-LA	0.59 (0.049)	0.69 (0.031)	0.42 (0.038)	0.10 (0.024)	—	0.59 (0.057)
β-LG	0.58 (0.051)	0.45 (0.05)	0.28 (0.022)	0.34 (0.026)	0.52 (0.058)	—

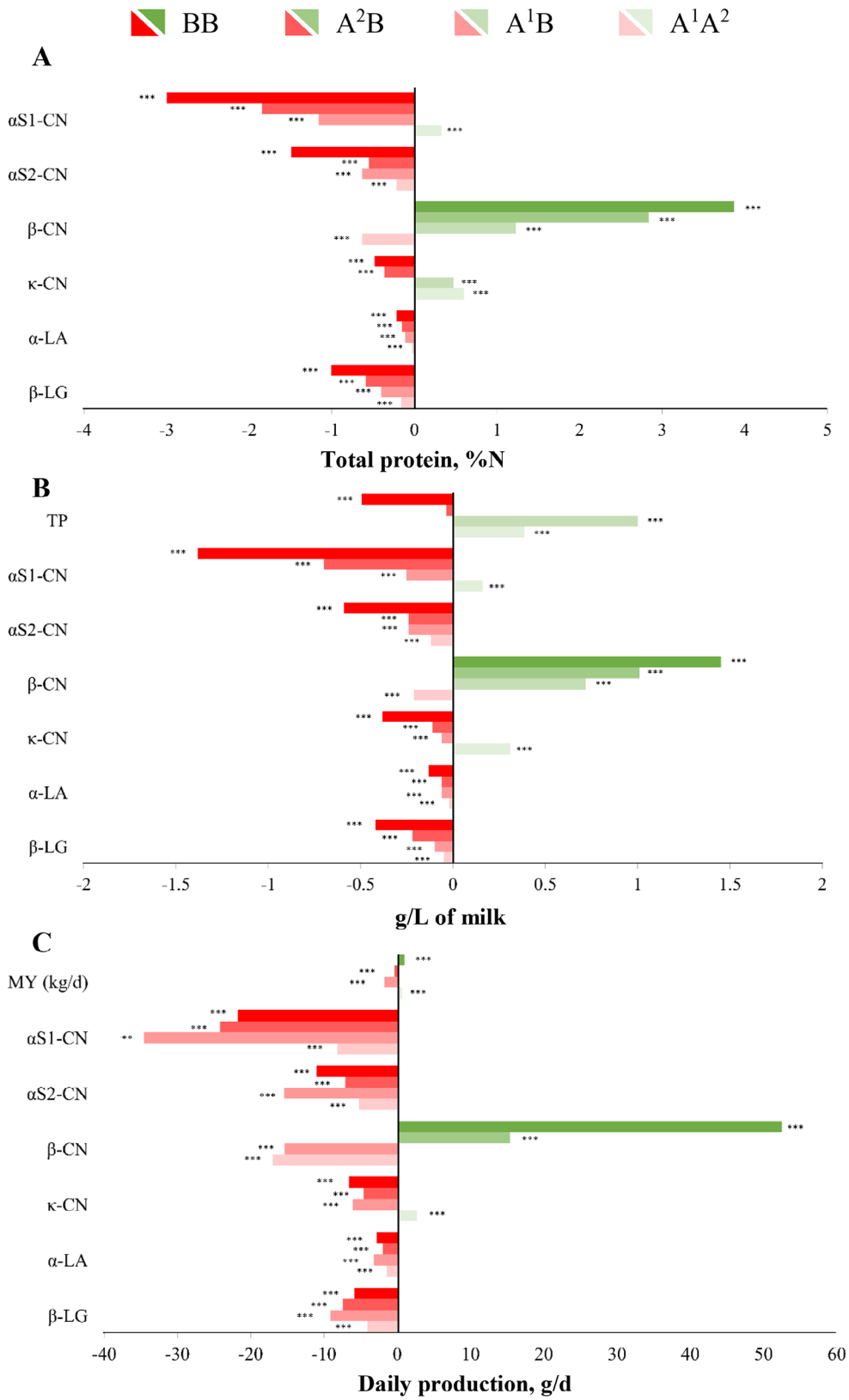

Figure 2. Contrasts of estimated effects for *CSN2* (β-CN) genotype on milk protein fraction content expressed in % N (A), g/L (B), and g/d (C) and on total milk protein (TP), and daily milk yield (MY). Homozygote genotype A2A2 is taken as reference. Asterisks represent the significance test. ***$P < 0.001$ for 95% of the highest posterior density (HPD95) of each genotype with respect to reference genotype.

cantly lower. Lastly, as the AA genotype was associated with a decrease in daily milk yield (−0.4 kg/d), the daily production of all protein fractions was lower, with the exception of β-LG and α_{S2}-CN. Unlike the *CSN3* gene, the alleles of the *BLG* gene showed an additive effect only in the case of their effects on β-LG, whereas all the other fractions were characterized by significant nonadditive effects.

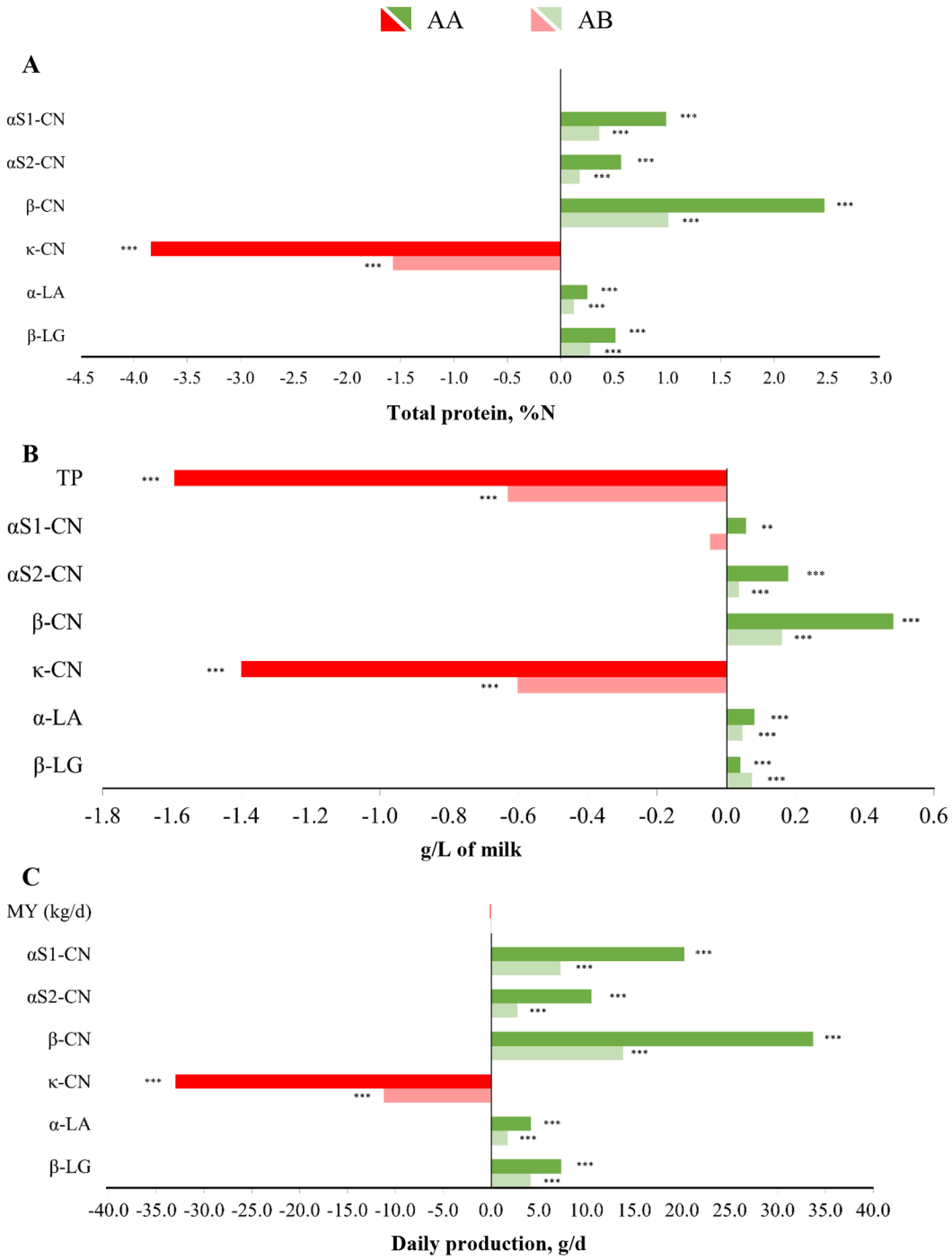

Figure 3. Contrasts of estimated effects for *CSN3* (κ-CN) genotype on milk protein fraction content expressed in % N (A), g/L (B), and g/d (C) and on total milk protein (TP), and daily milk yield (MY). Homozygote genotype BB is taken as reference. Asterisks represent the significance test. $^{**}P < 0.01$ and $^{***}P < 0.001$ for 95% of the highest posterior density (HPD95) of each genotype with respect to reference genotype.

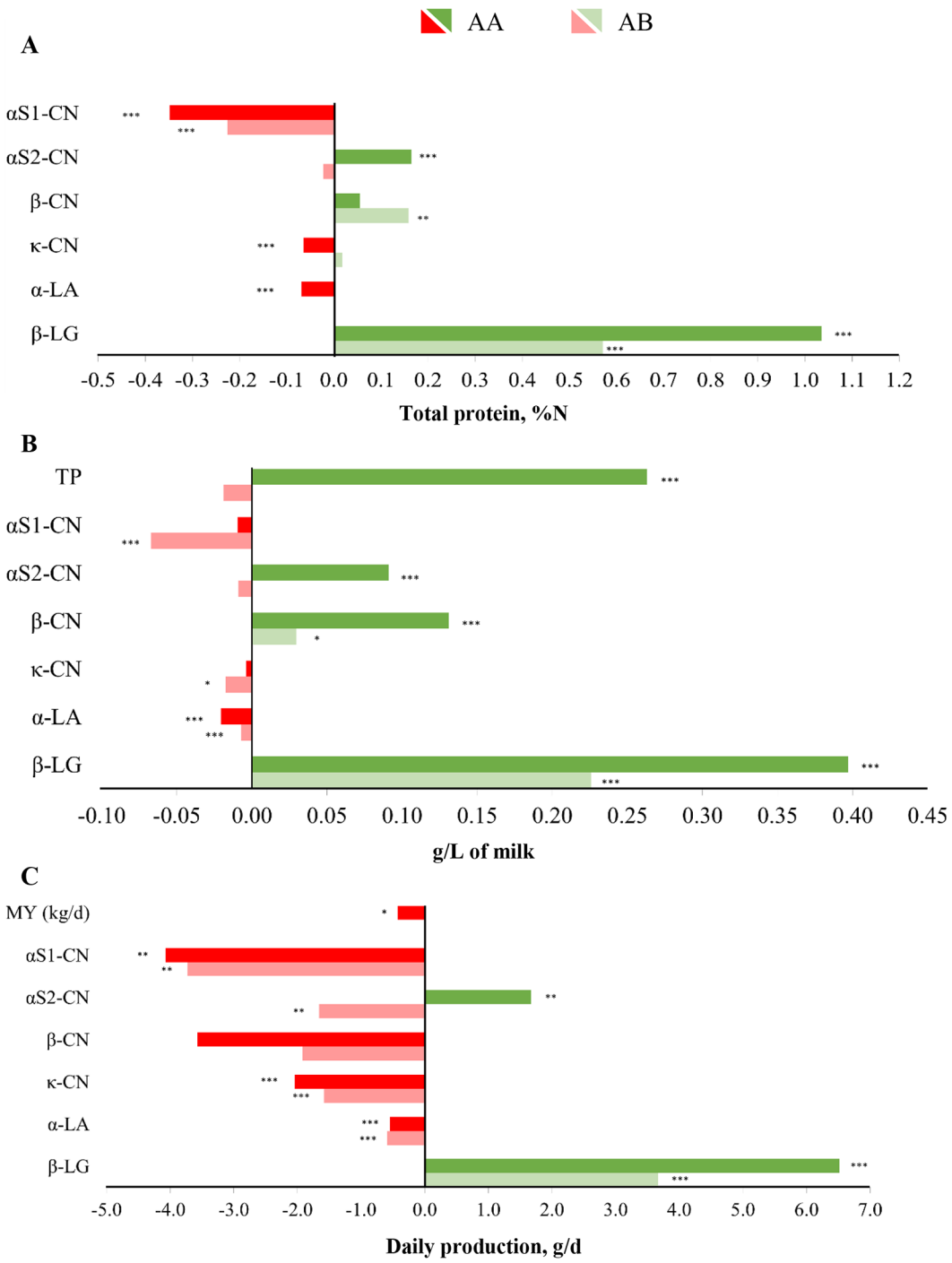

Figure 4. Contrasts of estimated effects for *BLG* (β-LG) genotype on milk protein fraction content expressed in % N (A), g/L (B), and g/d (C) and on total milk protein (TP), and daily milk yield (MY). Homozygote genotype BB is taken as reference. Asterisks represent the significance test. $*P < 0.05$, $**P < 0.01$, and $***P < 0.001$ for 95% of the highest posterior density (HPD95) of each genotype with respect to reference genotype.

DISCUSSION

It is well known that genetic factors play an important role in the phenotypic variability of the protein profile of bovine milk (Ng-Kwai-Hang et al., 1987; Heck et al., 2009). Unlike other production traits, milk protein fractions are controlled not only by many genes with smaller effects, but also by the major genes codifying for the genetic variants of milk proteins. Many researchers have studied the effects of casein and whey

protein genetic variants on the production and quality traits of milk (Cecchinato et al., 2012; Vallas et al., 2012; Albarella et al., 2020), and some also the effects of these variants on the protein fraction profile of milk (Heck et al., 2009; Bonfatti et al., 2010; Gustavsson et al., 2014). However, the effects of genetic variants of milk protein fractions have seldom been analyzed while simultaneously estimating the size of the genetic variance due to background polygenic effects. In addition, as far as the authors are aware, such analyses have never considered simultaneously milk protein fractions expressed qualitatively (% N), quantitatively (g/L) or as daily yield (g/d).

This work aimed to distinguish between the genetic variance of the major milk protein fractions controlled by the *CSN2*, *CSN3*, and *BLG* genotypes and the remaining polygenic background to determine the quota of the genetic variance controlled by the 3 genes separately, and to estimate the effects of their genotypes on the individual protein fractions. The genetic parameters of these traits and the different ways they can be expressed were previously analyzed with the base model and presented in a recent study within the same project on genomic selection (Macedo Mota et al., 2020), and therefore the descriptive statistics of the population and the results obtained from the base model will not be discussed further here. Instead, they will be used as our point of departure for discussion of the variations caused by the inclusion of the genotypes of the major genes in the model.

Genetic Variance of the Background Polygenic Effects on the Milk Protein Fractions

In the present work, the data were corrected for the effect of the genotypes of the *CSN2*, *CSN3*, and *BLG* genes. Our results indicating that most of the additive genetic variance was controlled by all 3 major genes, even where they are not the encoding gene for the fraction in question, are reflected in the previous literature. However, these findings are not entirely in agreement with those of Bonfatti et al. (2011) based on a Simmental population, in part because the protein proportions were expressed differently (caseins as proportions of the total casein content, and whey proteins as proportions of the total whey protein), and also because the authors used the *CSN2-CSN3* haplotypes instead of the genotypes of the 2 genes. In both studies, the α_{S2}-CN proportion was the casein fraction least affected by the major genes. In particular, Bonfatti et al. (2011) found the size of the genetic variance explained by the major genes to be only 16% compared with 73% in the present work. The size of the genetic variance in the α_{S1}-CN and κ-CN proportions explained by the 3 genes in both studies was between 80 and 90%. The casein most affected by the genotypes was β-CN, which explained more than 90% of the additive genetic variance. The influence of the 3 genotypes on the proportions of the whey proteins was slightly higher in the present work (66 and 67%), but still lower than their influence on the casein proportions.

Unfortunately, Bonfatti et al. (2011) did not make comparative estimates of the effects of the individual genes on the traits, but Schopen et al. (2009) reported reductions in the additive genetic variance after inclusion in the model of the genotypes of the 3 major genes individually. Unlike Bobe et al. (1999), who found a significant effect of the *CSN3* and *BLG* genes on the α_{S1}-CN proportion, Schopen et al. (2009) did not observe any strong effect of the major gene polymorphisms on this fraction. Our results, however, show that both the *CSN3* and *BLG* genes explained around 20% of the additive genetic variance, whereas the *CSN2* gene, which was not included in the Bobe et al. (1999) study, explained more than 30%. Bijl et al. (2014) also found a strong association between the *BLG* gene and the proportion of α_{S1}-CN-8P, which in our study is subsumed in the total α_{S1}-CN. Similar to Schopen et al. (2009), we found the *CSN2* gene to have a greater influence on the α_{S2}-CN and β-CN proportions, whereas the strongest influence of the *CSN3* and the *BLG* genes were on the κ-CN and β-LG proportions, respectively. However, unlike in the cited work, α-LA was more affected by the *CSN2* gene. The fact that the α-LA encoding gene (*LAA*) is on another chromosome (the fifth instead of the sixth for all caseins) suggests a sort of specific co-regulation of these genes when the proteins are expressed as relative proportions.

In analyzing the contents of the protein fractions in milk quantitatively, in g/L, the fact that the content of a given protein fraction is the result of multiplying the proportion of the fraction by the total protein in milk (also in g/L) should be considered. The genetic parameters of a given fraction content depend therefore on the genetic parameters of both the fraction proportion and the total protein content (co-variances included). In fact, only 20% of the genetic variance in total protein was controlled by the major genes. This reduction is much smaller than the reduction seen for the proportions of the different protein fractions (Table 2), but is also less than the reduction in the genetic variance in the protein fraction contents (Table 3). The small effect of the major genes on total protein suggests that their expression is down- and upregulated by the polygene responsible for the total protein content to compensate for the fluctuation in the expression of the protein fractions (Leroux et al., 2003). However, it also reflects compensations between the different protein

fractions caused by some negative genetic correlations between some of them, as will be discussed later. Of course, the major genes still have a notable effect on the protein fractions, as indicated by their influence on the genetic variance in their contents in milk, which ranges from 44 to 86%. Here, too, our results were similar to those of Bonfatti et al. (2011). In both studies, the fraction on which the effect of the major genes was most reduced in favor of the effect of the polygene after multiplication by total milk protein was α_{S1}-CN. The κ-CN fraction, on the other hand, was affected to the same or a slightly higher extent by the major genes, also in terms of g/L. In Bonfatti et al. (2011), the κ-CN showed a limited reduction of the proportion of genetic variance after the conversion to g/L, but it was still over the 80% as in the present study. Taking the major genes individually, *CSN2* controlled all the caseins except κ-CN, which was mostly influenced by its own gene (*CSN3*). The *BLG* gene controlled both the whey proteins. Bonfatti et al. (2011) considered a mixed inheritance model including together the effects of *CSN2-CSN3* haplotypes and *BLG* genotypes, therefore a comparison is not possible. When the protein fractions were expressed as daily yield (g/d), traits not considered by Bonfatti et al. (2011), the proportion of genetic variance controlled by the genotypes of the 3 major genes was still very similar to that of the protein fraction contents in milk.

In a recent phenotypic study on a different data set (cows of 6 different breeds reared in multibreed herds), Amalfitano et al. (2020) found that the genetic variants of the β-CN, κ-CN, and β-LG protein fractions represented the major sources of variation in the proportions of all the milk protein fractions. Inclusion of the genetic variants in the model greatly reduced the variations due to breed and the residual variance. In this genetic study, too, inclusion of the genetic variants in the statistical model not only reduced the additive genetic variance, but also the residual variance. So, in both this study and Amalfitano et al. (2020), including the genetic variants of the milk proteins seems to absorb not only part of the genetic variance but also part of the residual variance.

Effect of Genetic Variants on the Heritabilities of the Milk Protein Fractions

The genetic variance remaining after inclusion of the genetic variants of the 3 major genes is due to the background polygenic effects of the cow's genome. So, the heritability *strictu sensu* calculated from the new residual variance quantifies the genetic improvement of a population obtainable beyond that due to modification of the allelic frequencies of the major genes.

The reduction in the genetic variance after inclusion of the genotypes of the major genes in the model, if not accompanied by a reduction in the residual variance, is expected to reduce the heritability of the trait proportionally less than the reduction in the additive genetic variance, because the genetic variance affects both numerator and denominator of the heritability equation. For example, if a trait has a heritability of 0.40, and after inclusion of the major genes in the model the genetic variance decreases by 50%, the new heritability decreases to 0.25 (i.e., by 37.5%). On the other hand, if the residual variance decreases proportionally equally to the reduction in genetic variance, the heritability should not change. In the present study, the decrease in heritability is always lower than the decrease in genetic variance, but often less than expected because of a parallel decrease in the residual variance that is less severe than the decrease in genetic variance.

When the protein fractions were expressed qualitatively, in % N (Table 2), the heritability obtained from the base model was moderate to high (h^2: 0.25–0.79) and differed little to that found by Bonfatti et al. (2011; h^2: 0.28–0.69). Similar to our results, except for α-LA, which was not accounted for in the cited work, correction for the major genes explained more than 40% of the heritabilities of all the fractions. The only difference was in α_{S2}-CN, which was less affected by the major genes in Bonfatti et al. (2011).

When the protein fractions were analyzed quantitatively, in g/L, the estimates of heritability obtained from the base model (Table 3) were almost in the same range as the protein fraction proportions, with the notable exception of β-CN (h^2: 0.79 for the relative proportion, 0.59 for the content in milk). However, the reduction in heritability caused by inclusion in the statistical model of the genotypes of the major genes was much lower for the content of the protein fractions than for their nitrogen proportions, indicating the presence of other genes or polymorphisms able to control the genetic variability of these traits. The same effect was observed by Bonfatti et al. (2011), and in both studies this was particularly true for α_{S1}-CN in g/L, where the $\Delta\%h^2$ was half that of α_{S1}-CN in % N.

When the protein fractions were expressed as daily yields, the heritability estimates were lower (h^2: 0.14–0.39, Table 4) due to the influence of the single test-day milk yield ($h^2 = 0.14$), from which the daily yields of all protein fractions are derived by multiplication. Unlike the genetic variance controlled by the genotypes, the quota of the heritability explained by them was much lower than that of the protein fraction contents in milk. To understand this pattern, we should consider that daily milk yield was also associated with the genotypes of the 3 major genes (Table 4), and that the decrease

in heritability was identical to the decrease in genetic variance. This is probably due also to the negative genetic correlations between protein contents and milk yield, and to the much larger variability of the latter. In fact, differences at protein level tend to cancel out at milk level.

Genetic Correlations

Genetic correlations are important for understanding the response of the protein fractions to selective breeding. In the present study, the aim was to see how the major genes affected the correlations of total protein, milk yield, and protein fractions expressed as relative proportions, milk contents, and daily yields. Casein-encoding genes are known to be situated in a short genetic linkage group on a 250-kb region of chromosome 6 (Caroli et al., 2009), but there are also other genes known to be responsible for the regulation of casein and whey protein expression (Schopen et al., 2011; Pegolo et al., 2018).

Figure 2 presents the correlations of total protein (A) and milk yield (B) with the protein fractions. Improvement in these 2 traits is the focus of most dairy cattle breeding programs, but modification of the protein profile is still not a breeding goal (Miglior et al., 2005). Selection for increased milk yield is known to result in genetic worsening of the protein and fat contents of milk. This means that the increased secretion of lactose, water, and minerals in the udder is accompanied by a proportionally smaller increase in protein and fat secretion (dilution effect). It is not known, however, whether this unfavorable genetic correlation between milk yield and total protein extends to all the protein fractions. It is clear from Figure 2B that all the protein fractions are negatively correlated with milk yield (from −0.25 for β-LG to −0.41 for β-CN) with the notable exception of the α_{S1}-CN content, which appears to be positively correlated (0.28). It can also be seen that after inclusion of the genotypes of the 3 major genes the correlation between milk yield and α_{S1}-CN content remains favorable although much smaller. The unfavorable correlation with the β-CN content also decreases after inclusion of the major genes, whereas the other correlations were only negligibly affected.

It is not known whether selection in favor of the total protein content in milk would change its composition, but it can be seen from Figure 2A that total protein content exhibited a modest, favorable correlation with the relative proportions of κ-CN (0.19) and β-LG (0.23), and a negative correlation with all the other fractions (from −0.23 with α_{S2}-CN to −0.48 with α-LA). The major genes played an evident role only with respect to κ-CN. In fact, including the genotypes of the 3 major genes in the statistical model canceled the favorable correlation between κ-CN and the total milk protein content, so the 2 traits seem to be genetically independent. On the other hand, including the genotypes of the 3 major genes did not substantially modify the sign and entity of the other genetic correlations. Conflicting results are reported in the literature. Schopen et al. (2009) found the protein percentage in milk to have a strong positive correlation with the proportion of κ-CN, but correlated negatively with α_{S1}-CN and α-LA; Gebreyesus et al. (2016) found only small positive or almost null correlations between protein composition and total protein; finally, Sanchez et al. (2017) found that total milk protein correlated negatively with α-LA in all 3 breeds they studied, negatively with β-CN in the Montbéliarde breed, and positively with α_{S1}-CN in the Normande breed.

Table 5 shows the genetic correlations between the protein fractions. When these are considered in terms of their relative proportions, their sum therefore being constant, it is expected that favorable correlations between some of them will be compensated for by unfavorable correlations with others. Table 5 shows that almost all the proportions of the protein fractions were positively correlated with each other, but negatively correlated with the β-CN proportion (from −0.25 for β-LG to −0.79 for κ-CN, except α-LA). There was also a negative correlation between the proportions of β-LG and κ-CN (−0.39). These results agree with the findings of Bonfatti et al. (2010, 2011), who also found strong negative correlations between the proportion of β-CN and the other caseins. They suggest that the negative correlations between β-CN and the proportions of the other fractions could be associated with an effect of competitive synthesis. In fact, the synthesis of some fractions can influence the synthesis of others due to a limited pool of transcription factors or AA, or both (Bobe et al., 1999). Schopen et al. (2009) found that the protein fraction proportions were more negatively correlated with α_{S1}-CN, the other major casein in milk. When the fractions are considered in terms of their contents in milk, we would expect positive correlations between them if their proportions remained constant, but, as was observed, this is not true. The negative correlations between the proportions of β-CN and the other protein fractions are also confirmed at level of their contents in milk (g/L). This is only partially in agreement with results presented in Bonfatti et al. (2011). The increase in daily milk yield tends, of course, to increase the daily production of all the milk protein fractions, and this is confirmed by the positive correlations observed between the daily yields of all the protein fractions. The only exception was κ-CN, which was still negatively correlated with β-CN (−0.36).

Therefore, β-CN and κ-CN were the only fractions that were always negatively correlated with each other in qualitative and quantitative terms, and as daily production. The mechanisms regulating the synthesis of proteins in the udder, and hence their daily yields, and their milk concentrations and proportions, are probably controlled more by the background polygene, because correction for the genotypes of the 3 major genes did not substantially change the correlation matrix. Only in the case of the correlation between β-CN and α_{S1}-CN in g/L did correction for the major genes have a greater effect, changing the correlation from −0.31 to 0.21. This suggests that selecting for favorable genetic variants for β-CN content can negatively affect the α_{S1}-CN content, as will be shown later. Discrepancies between the cited studies could be ascribed to the very different allelic frequencies of different breeds, and the different statistical methods used by the authors.

Effects of the β-CN Genotypes on Milk Yield and the Protein Profile

The results of the statistical models with and without inclusion of the genetic variants of the 3 major genes considered clearly revealed that the different alleles codifying for different genetic variants are responsible for not inconsiderable effects on the milk protein profile, whether in qualitative, quantitative, or daily production terms.

In particular, compared with the ancestral β-CN A^2 variant, the presence of the B allele promotes the synthesis of the β-CN proportion and content in milk, and its yield, at the expense of all the other fractions. Unfortunately, the very small numbers of A^1A^1 animals in the population prevented us from decisively establishing the differences between the A^1 and A^2 variants, but the heterozygotes A^1B and A^1A^2 helped us place the protein profile of the A^1 variant closer to the A^2 than the B variant. It is also true that the A^1 variant is the result of an SNP in position 8101 on the *CSN2* gene encoding for the A^2 variant, whereas the B has one more SNP than the A^1 variant in position 8267 (Caroli et al., 2009). In terms of the relative proportions and contents of α_{S1}-CN and κ-CN, the A^1 variant seemed to be superior to the A^2. Most previous studies used the haplotypes of β-κ-CN, but all of them agree that the presence of the *CSN2* B allele is associated with a higher proportion and content of β-CN, and lower proportions and contents of α_{S1}-CN in particular, and also α_{S2}-CN (Hallén et al., 2008; Heck et al., 2009; Bonfatti et al., 2010). Moreover, Heck et al. (2009) and Bonfatti et al. (2010) showed that the presence of the *CSN2* A^1 allele was associated with higher proportions and contents of α_{S1}-CN and κ-CN compared with the *CSN2* A^2 allele. This effect is only marginal in our study. The increase in the B allele of *CSN2* could be of interest to the dairy industry, provided that its effect on human health is favorable. This is not so much for the increase in β-CN, which is compensated for by the decrease in α_{S1}-CN, as for the decrease in α_{S2}-CN and β-LG, both of which are associated with a worsening of the coagulation properties of milk and a lower cheese yield (Cipolat-Gotet et al., 2018; Amalfitano et al., 2019).

Effects of the κ-CN Genotypes on Milk Yield and the Protein Profile

In terms of contents in milk, compared with its strong negative effect on κ-CN the A allele of the *CSN3* gene had a much less favorable effect on the other protein fractions, so that overall there was a negative effect on the total protein content of milk. As daily milk yield was not affected by this gene, its effect on the daily productions of individual protein fractions was very similar to its effect on their proportions. The higher content and relative proportion of the κ-CN in the presence of the B variant is a consistent finding in the literature. In fact, it is known that the SNP in the promoter region of the *CSN3* gene encoding the B variant is responsible for overexpression of the protein fraction in milk (Ikonen et al., 1997; Bobe et al., 1999; Hallén et al., 2008). This overexpression is probably compensated for by the effect of downregulation of the genes encoding for the other fractions, which reduces their relative proportions (Bobe et al., 1999; Leroux et al., 2003). However, the literature is inconsistent regarding the effect of *CSN3* on the fractions other than κ-CN. Bobe et al. (1999) found that an increase in the κ-CN proportion in the presence of *CSN3* B corresponded with a decrease in the α_{S1}-CN and β-LG proportions. Heck et al. (2009) partially agree with this, but also found a decrease in the α-LA and an increase in the α_{S2}-CN proportions. Similar to us, Bonfatti et al. (2010) attributed the increase in the total protein content of milk in the presence of *CSN3* B to the increase in the κ-CN content alone. They also confirmed the increase in the κ-CN proportion to be mainly at the expense of α_{S1}-CN, supporting the hypothesis of McClenaghan et al. (1995) that there is specific expression completion in the mammary gland. None of these studies found great variations in the contents of the other protein fractions due to the *CSN3* genotype. The B variant of κ-CN is currently considered a potential selection goal for breeders because of its beneficial effects on milk coagulation properties, principally due to the higher expression of κ-CN. However, more studies need to be conducted on the effect of this allele on the other protein fractions.

Effects of the β-LG Genotypes on Milk Yield and the Protein Profile

Although the literature agrees about the favorable effect of the A allele on the proportion, content in milk, and daily yield of β-LG, the effect of the *BLG* gene on the other protein fractions is not well known. As for κ-CN, the increase in β-LG in the presence of the A variant is attributed to polymorphism in the promoter region of the *BLG* gene (Folch et al., 1999). Heck et al. (2009), instead, pointed to the differences in the stability of the mRNA transcribed from the LGB A and B alleles as an explanation for the variation in β-LG content. However, all the studies agree on attributing to the A variant the higher proportion and content of β-LG. About the other protein fractions, compared with the β-LG B, in our study the presence of the A variant was associated, to a lesser degree than β-LG, with an increase of proportion and content in milk, and daily yield also of α_{S2}-CN. It was also associated with a higher content of β-CN and, consequently, of total protein. The influence of A allele on the other fractions was instead unfavorable, including also a slight negative effect on milk production. Bobe et al. (1999) partially agree with our results, reporting a decrease in the proportions of α_{S1}-CN, β-CN, and κ-CN with the increase in β-LG. Heck et al. (2009) found that the increase in β-LG in the presence of the A variant was compensated for by decreases in all the other fractions. At the same time, Bonfatti et al. (2010) showed that the effect of the *BLG* B was to increase the total protein content and the contents of all the protein fractions in the milk, except β-LG, which decreased. It seems from our results that milk carrying the A variant of β-LG is less suitable for cheese making because of the higher proportions and contents of β-LG and α_{S2}-CN, which are detrimental to the coagulation and syneresis processes (Amalfitano et al., 2019).

CONCLUSIONS

In this study we showed that the proportions (% N) and contents (g/L) of protein fractions are highly heritable; more heritable than those expressed as daily yields (g/d). The genotypes of 3 major genes (*CSN2*, *CSN3*, and *BLG*) absorb the major part of the genetic variance of all protein fractions and a small part of the residual variance, whereas the remaining amount of variance is explained by the polygenic effect. The heritability estimates, although reduced by the inclusion in the statistical model of the gene casein cluster, exhibited larger values with respect to that of milk yield, therefore selective breeding for protein fractions seems to be feasible. At the population level, beyond a genomic array including the genotypes of the *CSN2*, *CSN3,* and *BLG,* a phenotypic assessment through infrared tools might be a viable option along with the integration of genomic information.

ACKNOWLEDGMENTS

The authors thank the Autonomous Province of Trento (Italy), who provided the funds for the project, and Valentina Bonfatti (Department of Comparative Biomedicine and Food Science, University of Padova, Legnaro, Italy) for assistance in the HPLC analyses. The authors have not stated any conflicts of interest.

REFERENCES

Albarella, S., M. Selvaggi, E. D'anza, G. Cosenza, S. Caira, A. Scaloni, A. Fontana, V. Peretti, and F. Ciotola. 2020. Influence of the casein composite genotype on milk quality and coagulation properties in the endangered agerolese cattle breed. Animals (Basel) 10:892. https://doi.org/10.3390/ani10050892.

Amalfitano, N., C. Cipolat-Gotet, A. Cecchinato, M. Malacarne, A. Summer, and G. Bittante. 2019. Milk protein fractions strongly affect the patterns of coagulation, curd firming, and syneresis. J. Dairy Sci. 102:2903–2917. https://doi.org/10.3168/jds.2018-15524.

Amalfitano, N., G. Stocco, A. Maurmayr, S. Pegolo, A. Cecchinato, and G. Bittante. 2020. Quantitative and qualitative detailed milk protein profiles of 6 cattle breeds: Sources of variation and contribution of protein genetic variants. J. Dairy Sci. 103:11190–11208. https://doi.org/10.3168/jds.2020-18497.

Atsbeha, D. M., D. Kristofersson, and K. Rickertsen. 2016. Component supply responses in dairy production. Eur. Rev. Agric. Econ. 43:193–215. https://doi.org/10.1093/erae/jbv019.

Bailey, K. W., C. M. Jones, and A. J. Heinrichs. 2005. Economic returns to Holstein and Jersey herds under multiple component pricing. J. Dairy Sci. 88:2269–2280. https://doi.org/10.3168/jds.S0022-0302(05)72903-9.

Barnett, M. P. G., W. C. Mcnabb, N. C. Roy, K. B. Woodford, and A. J. Clarke. 2014. Dietary A1 β-casein affects gastrointestinal transit time, dipeptidyl peptidase-4 activity, and inflammatory status relative to A2 β-casein in Wistar rats. Int. J. Food Sci. Nutr. 65:720–727. https://doi.org/10.3109/09637486.2014.898260.

Bijl, E., H. J. F. van Valenberg, T. Huppertz, A. C. M. van Hooijdonk, and H. Bovenhuis. 2014. Phosphorylation of α_{S1}-casein is regulated by different genes. J. Dairy Sci. 97:7240–7246. https://doi.org/10.3168/jds.2014-8061.

Birgisdottir, B. E., J. P. Hill, A. V. Thorsson, and I. Thorsdottir. 2006. Lower consumption of cow milk protein A1 β-casein at 2 years of age, rather than consumption among 11-to 14-year-old adolescents, may explain the lower incidence of type 1 diabetes in Iceland than in Scandinavia. Ann. Nutr. Metab. 50:177–183. https://doi.org/10.1159/000090738.

Bittante, G., C. Cipolat-Gotet, and A. Cecchinato. 2013. Genetic parameters of different measures of cheese yield and milk nutrient recovery from an individual model cheese-manufacturing process. J. Dairy Sci. 96:7966–7979. https://doi.org/10.3168/jds.2012-6517.

Bittante, G., C. Cipolat-Gotet, F. Malchiodi, E. Sturaro, F. Tagliapietra, S. Schiavon, and A. Cecchinato. 2015. Effect of dairy farming system, herd, season, parity, and days in milk on modeling of the coagulation, curd firming, and syneresis of bovine milk. J. Dairy Sci. 98:2759–2774. https://doi.org/10.3168/jds.2014-8909.

Bobe, G., D. C. Beitz, A. E. Freeman, and G. L. Lindberg. 1999. Effect of milk protein genotypes on milk protein composition and its genetic parameter estimates. J. Dairy Sci. 82:2797–2804. https://doi.org/10.3168/jds.S0022-0302(99)75537-2.

Bonfatti, V., A. Cecchinato, L. Gallo, A. Blasco, and P. Carnier. 2011. Genetic analysis of detailed milk protein composition and coagulation properties in Simmental cattle. J. Dairy Sci. 94:5183–5193. https://doi.org/10.3168/jds.2011-4297.

Bonfatti, V., G. Di Martino, A. Cecchinato, D. Vicario, and P. Carnier. 2010. Effects of β-κ-casein (CSN2–CSN3) haplotypes and β-lactoglobulin (BLG) genotypes on milk production traits and detailed protein composition of individual milk of Simmental cows. J. Dairy Sci. 93:3797–3808. https://doi.org/10.3168/jds.2009-2778.

Bonfatti, V., L. Grigoletto, A. Cecchinato, L. Gallo, and P. Carnier. 2008. Validation of a new reversed-phase high-performance liquid chromatography method for separation and quantification of bovine milk protein genetic variants. J. Chromatogr. A 1195:101–106. https://doi.org/10.1016/j.chroma.2008.04.075.

Caroli, A. M., S. Chessa, and G. J. Erhardt. 2009. Invited review: Milk protein polymorphisms in cattle: Effect on animal breeding and human nutrition. J. Dairy Sci. 92:5335–5352. https://doi.org/10.3168/jds.2009-2461.

Cecchinato, A., A. Albera, C. Cipolat-Gotet, A. Ferragina, and G. Bittante. 2015. Genetic parameters of cheese yield and curd nutrient recovery or whey loss traits predicted using Fourier-transform infrared spectroscopy of samples collected during milk recording on Holstein, Brown Swiss, and Simmental dairy cows. J. Dairy Sci. 98:4914–4927. https://doi.org/10.3168/jds.2014-8599.

Cecchinato, A., C. Ribeca, A. Maurmayr, M. Penasa, M. De Marchi, N. P. P. Macciotta, M. Mele, P. Secchiari, G. Pagnacco, and G. Bittante. 2012. Short communication: Effects of β-lactoglobulin, stearoyl-coenzyme A desaturase 1, and sterol regulatory element binding protein gene allelic variants on milk production, composition, acidity, and coagulation properties of Brown Swiss cows. J. Dairy Sci. 95:450–454. https://doi.org/10.3168/jds.2011-4581.

Cecchinato, A. C., C. Cipolat-Gotet, J. Casellas, M. Penasa, A. Rossoni, and G. Bittante. 2013. Genetic analysis of rennet coagulation time, curd-firming rate, and curd firmness assessed over an extended testing period using mechanical and near-infrared instruments. J. Dairy Sci. 96:50–62. https://doi.org/10.3168/jds.2012-5784.

Cipolat-Gotet, C., A. Cecchinato, M. Malacarne, G. Bittante, and A. Summer. 2018. Variations in milk protein fractions affect the efficiency of the cheese-making process. J. Dairy Sci. 101:8788–8804. https://doi.org/10.3168/jds.2018-14503.

De Noni, I., R. J. Fitzgerald, H. J. T. Korhonen, Y. Le Roux, C. T. Livesey, I. Thorsdottir, D. Tomé, and R. Witkamp. 2009. Review of the potential health impact of β-casomorphins and related peptides. EFSA Sci. Rep. 231:1–107. https://doi.org/10.2903/j.efsa.2009.231r.

Elliott, R. B., D. P. Harris, J. P. Hill, N. J. Bibby, and H. E. Wasmuth. 1999. Type I (insulin-dependent) diabetes mellitus and cow milk: Casein variant consumption. Diabetologia 42:292–296. https://doi.org/10.1007/s001250051153.

Farrell, H. M. Jr., R. Jimenez-Flores, G. T. Bleck, E. M. Brown, J. E. Butler, L. K. Creamer, C. L. Hicks, C. M. Hollar, K. F. Ng-Kwai-Hang, and H. E. Swaisgood. 2004. Nomenclature of the proteins of cows' milk - Sixth revision. J. Dairy Sci. 87:1641–1674. https://doi.org/10.3168/jds.S0022-0302(04)73319-6.

Fiedorowicz, E., L. H. Markiewicz, K. Sidor, D. Świątecka, A. Cieślińska, M. Matysiewicz, K. Piskorz-Ogórek, E. Sienkiewicz-Szłapka, M. Teodorowicz, A. Świątecka, and E. Kostyra. 2016. The influence of breast milk and infant formulae hydrolysates on bacterial adhesion and Caco-2 cells functioning. Food Res. Int. 89:679–688. https://doi.org/10.1016/j.foodres.2016.09.022.

Folch, J. M., P. Dovč, and J. F. Medrano. 1999. Differential expression of bovine β-lactoglobulin A and B promoter variants in transiently transfected HC11 cells. J. Dairy Res. 66:537–544. https://doi.org/10.1017/S0022029999003787.

Gebreyesus, G., M. S. Lund, L. Janss, N. A. Poulsen, L. B. Larsen, H. Bovenhuis, and A. J. Buitenhuis. 2016. Short communication: Multi-trait estimation of genetic parameters for milk protein composition in the Danish Holstein. J. Dairy Sci. 99:2863–2866. https://doi.org/10.3168/jds.2015-10501.

Geweke, J. 1992. Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments. Bayesian Stat. 4:169–193.

Gustavsson, F., A. J. Buitenhuis, M. Johansson, H. P. Bertelsen, M. Glantz, N. A. Poulsen, H. Lindmark Månsson, H. Stålhammar, L. B. Larsen, C. Bendixen, M. Paulsson, and A. Andrén. 2014. Effects of breed and casein genetic variants on protein profile in milk from Swedish Red, Danish Holstein, and Danish Jersey cows. J. Dairy Sci. 97:3866–3877. https://doi.org/10.3168/jds.2013-7312.

Hallén, E., A. Wedholm, A. Andrén, and A. Lundén. 2008. Effect of β-casein, κ-casein and β-lactoglobulin genotypes on concentration of milk protein variants. J. Anim. Breed. Genet. 125:119–129. https://doi.org/10.1111/j.1439-0388.2007.00706.x.

Heck, J. M. L., A. Schennink, H. J. F. Van Valenberg, H. Bovenhuis, M. H. P. W. Visker, J. A. M. van Arendonk, and A. C. M. van Hooijdonk. 2009. Effects of milk protein variants on the protein composition of bovine milk. J. Dairy Sci. 92:1192–1202. https://doi.org/10.3168/jds.2008-1208.

Ikonen, T., M. Ojala, and E. L. Syväoja. 1997. Effects of composite casein and β-lactoglobulin genotypes on renneting properties and composition of bovine milk by assuming an animal model. Agric. Food Sci. 6:283–294. https://doi.org/10.23986/afsci.72791.

Kamiński, S., A. Cieślińska, and E. Kostyra. 2007. Polymorphism of bovine beta-casein and its potential effect on human health. J. Appl. Genet. 48:189–198. https://doi.org/10.1007/BF03195213.

Küllenberg de Gaudry, D. K., S. Lohner, C. Schmucker, P. Kapp, E. Motschall, S. Horrlein, C. Roger, and J. J. Meerpohl. 2019. Milk A1 β-casein and health-related outcomes in humans: A systematic review. Nutr. Rev. 77:278–306. https://doi.org/10.1093/nutrit/nuy063.

Laugesen, M., and R. Elliott. 2003. Ischaemic heart disease, type 1 diabetes, and cow milk A1 beta-casein. N. Z. Med. J. 116:U295.

Leroux, C., F. Le Provost, E. Petit, L. Bernard, Y. Chilliard, and P. Martin. 2003. Real-time RT-PCR and cDNA macroarray to study the impact of the genetic polymorphism at the α_{s1} -Casein locus on the expression of genes in the goat mammary gland during lactation. Reprod. Nutr. Dev. 43:459–469. https://doi.org/10.1051/rnd:2003032.

Macedo Mota, L. F. M., S. Pegolo, V. Bisutti, G. Bittante, and A. Cecchinato. 2020. Genomic analysis of milk protein fractions in brown swiss cattle. Animals (Basel) 10:336. https://doi.org/10.3390/ani10020336.

Martin, P., M. Szymanowska, L. Zwierzchowski, and C. Leroux. 2002. The impact of genetic polymorphisms on the protein composition of ruminant milks. Reprod. Nutr. Dev. 42:433–459. https://doi.org/10.1051/rnd:2002036.

McClenaghan, M., A. Springbett, R. M. Wallace, C. J. Wilde, and A. J. Clark. 1995. Secretory proteins compete for production in the mammary gland of transgenic mice. Biochem. J. 310:637–641. https://doi.org/10.1042/bj3100637.

McLean, D. M., E. R. Graham, R. W. Ponzoni, and H. A. McKenzie. 1984. Effects of milk protein genetic variants on milk yield and composition. J. Dairy Res. 51:531–546. https://doi.org/10.1017/S0022029900032854.

Miglior, F., B. L. Muir, and B. J. Van Doormaal. 2005. Selection indices in Holstein cattle of various countries. J. Dairy Sci. 88:1255–1263. https://doi.org/10.3168/jds.S0022-0302(05)72792-2.

Milan, A. M., A. Shrestha, H. J. Karlström, J. A. Martinsson, N. J. Nilsson, J. K. Perry, L. Day, M. P. G. Barnett, and D. Cameron-Smith. 2020. Comparison of the impact of bovine milk β-casein variants on digestive comfort in females self-reporting dairy intolerance: A randomized controlled trial. Am. J. Clin. Nutr. 111:149–160. https://doi.org/10.1093/ajcn/nqz279.

Misztal, I., S. Tsuruta, D. Lourenco, I. Aguilar, A. Legarra, and Z. Vitezica. 2018. Manual for BLUPF90 family of programs. University of Georgia.

Ng-Kwai-Hang, K. F., J. F. Hayes, J. E. Moxley, and H. G. Monardes. 1987. Variation in milk protein concentrations associated with genetic polymorphism and environmental factors. J. Dairy Sci. 70:563–570. https://doi.org/10.3168/jds.S0022-0302(87)80042-5.

Pegolo, S., N. Mach, Y. Ramayo-Caldas, S. Schiavon, G. Bittante, and A. Cecchinato. 2018. Integration of GWAS, pathway and network analyses reveals novel mechanistic insights into the synthesis of milk proteins in dairy cows. Sci. Rep. 8:566. https://doi.org/10.1038/s41598-017-18916-4.

Perdry, H., and L. Dandine-Roulland. 2018. 46th European Mathematical Genetics Meeting (EMGM) 2018, Cagliari, Italy, April 18–20, 2018: Abstracts. Hum. Hered. 83:1–29. https://doi.org/10.1159/000488519.

Sanchez, M. P., M. Ferrand, M. Gelé, D. Pourchet, G. Miranda, P. Martin, M. Brochard, and D. Boichard. 2017. Short communication: Genetic parameters for milk protein composition predicted using mid-infrared spectroscopy in the French Montbéliarde, Normande, and Holstein dairy cattle breeds. J. Dairy Sci. 100:6371–6375. https://doi.org/10.3168/jds.2017-12663.

Schopen, G. C. B., J. M. L. Heck, H. Bovenhuis, M. H. P. W. Visker, H. J. F. van Valenberg, and J. A. M. van Arendonk. 2009. Genetic parameters for major milk proteins in Dutch Holstein-Friesians. J. Dairy Sci. 92:1182–1191. https://doi.org/10.3168/jds.2008-1281.

Schopen, G. C. B., M. H. P. W. Visker, P. D. Koks, E. Mullaart, J. A. M. van Arendonk, and H. Bovenhuis. 2011. Whole-genome association study for milk protein composition in dairy cattle. J. Dairy Sci. 94:3148–3158. https://doi.org/10.3168/jds.2010-4030.

Sebastiani, C., C. Arcangeli, M. Ciullo, M. Torricelli, G. Cinti, S. Fisichella, and M. Biagetti. 2020. Frequencies evaluation of β-casein gene polymorphisms in dairy cows reared in central Italy. Animals (Basel) 10:252. https://doi.org/10.3390/ani10020252.

Smith, B. J. 2007. boa: An R package for MCMC output convergence assessment and posterior inference. J. Stat. Softw. 21:1–37. https://doi.org/10.18637/jss.v021.i11.

Summer, A., F. Di Frangia, P. Ajmone Marsan, I. De Noni, and M. Malacarne. 2020. Occurrence, biological properties and potential effects on human health of β-casomorphin 7: Current knowledge and concerns. Crit. Rev. Food Sci. Nutr. 60:3705–3723. https://doi.org/10.1080/10408398.2019.1707157.

Vallas, M., T. Kaart, S. Värv, K. Pärna, I. Jõudu, H. Viinalass, and E. Pärna. 2012. Composite β-κ-casein genotypes and their effect on composition and coagulation of milk from Estonian Holstein cows. J. Dairy Sci. 95:6760–6769. https://doi.org/10.3168/jds.2012-5495.

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

Wedholm, A., L. B. Larsen, H. Lindmark-Månsson, A. H. Karlsson, and A. Andrén. 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.

Yin, H., J. Miao, and Y. Zhang. 2010. Protective effect of β-casomorphin-7 on type 1 diabetes rats induced with streptozotocin. Peptides 31:1725–1729. https://doi.org/10.1016/j.peptides.2010.05.016.

Zhang, W., J. Miao, S. Wang, and Y. Zhang. 2013. The protective effects of beta-casomorphin-7 against glucose-induced renal oxidative stress in vivo and vitro. PLoS One 8:e63472. https://doi.org/10.1371/journal.pone.0063472.

Zhu, L., J. Li, D. Wu, and B. Li. 2018. The protective effect of beta-casomorphin-7 via promoting Foxo1 activity and nuclear translocation in human lens epithelial cells. Cutan. Ocul. Toxicol. 37:267–274. https://doi.org/10.1080/15569527.2018.1445095.

ORCIDS

Nicolò Amalfitano https://orcid.org/0000-0001-6030-1620
Lucio Flavio Macedo Mota https://orcid.org/0000-0001-9983-1784
Guilherme J. M. Rosa https://orcid.org/0000-0001-9172-6461
Alessio Cecchinato https://orcid.org/0000-0003-3518-720X
Giovanni Bittante https://orcid.org/0000-0001-7137-7049