Effects of \( \beta \)- and \( \kappa \)-casein, and \( \beta \)-lactoglobulin single and composite genotypes on milk composition and milk coagulation properties of Italian Holsteins assessed by FT-MIR

Filippo Cendon \(^a\), Marco Franzoi \(^d\), Mauro Pena \(^a\), Massimo De Marchi \(^a\) and Martino Cassandro \(^a,b,c\)

\(^a\)Dipartimento di Agronomia, Animali, Alimenti, Risorse naturali e Ambiente, University of Padova, Legnaro, Italy; \(^b\)Federazione delle Associazioni Nazionali di Razza e di Specie, Roma, Italy; \(^c\)Associazione Nazionale Allevatori di Razza Frisona Bruna Jersey Italiana, Cremona, Italy

ABSTRACT
The present study aimed to investigate the effect of \( \beta \)- and \( \kappa \)-casein, and \( \beta \)-lactoglobulin genotypes on milk composition and milk coagulation properties predicted using Fourier-transform mid-infrared spectroscopy (FT-MIR). The final dataset provided 74,721 observations from 5316 Holstein–Friesian cows reared in 122 herds in the Veneto region. Individual cow milk composition was obtained from Dairy Herd Improvement (DHI) testing between the years 2012 and 2021. Genotypes were analysed through customised bovine genotyping chip. The effect of both separate and combined milk protein genotype on major milk components, SCS, urea, and predicted milk coagulation properties, were investigated applying a linear mixed model. \( \beta \)-lactoglobulin BB resulted in the higher fat content, whereas protein and casein content was mostly enhanced by the \( \kappa \)-casein BB haplotype. The urea content was significantly lower in \( \beta \)-lactoglobulin AA and \( \kappa \)-casein EE genotypes and SCS was significantly lower in \( \beta \)-casein A1A1, \( \beta \)-lactoglobulin AA, and \( \kappa \)-casein BB genotypes. Milk coagulation properties performance was significantly improved by \( \beta \)-casein A1A1, \( \beta \)-LG BB and \( \kappa \)-casein BB, both as single variants and as composite genotype. The effect of genotypes on milk composition and milk coagulation properties showed five and four principal hierarchical clusters, respectively. Clustering highlighted unexpected relationships among genotypes in terms of final milk quality, suggesting it could be possible to select favourable genotypes maintaining sufficient diversity.

HIGHLIGHTS
- Milk protein variants have an impact on milk technological properties, but phenotype collection is laborious and expensive.
- Data collected using mid-infrared spectroscopy can be used to estimate milk coagulation properties and perform population studies.
- Milk protein variants revealed unexpected relationships with milk composition and technological traits.

Introduction
Among milk proteins, caseins represent the main component of the entire milk cheese-making process (Mariani and Pecorari 1991). These are a family of phosphoproteins, synthesised in the mammary gland, responsible for many of the technological properties of milk. \( \beta \)-casein and \( \kappa \)-casein (\( \beta \)-CN and \( \kappa \)-CN, respectively) differ in the amino acid sequences and the distribution of the electrical charges, given both by the different amino acids and the phosphorylation on serine and threonine residues (Walstra and Jenness 1984; Rollema 1992; Holt et al. 2013). \( \beta \)-CN is the most hydrophobic casein that represents 34% of total caseins, and it is constituted by a polypeptide chain of 209 amino acids (Holt et al. 2013). \( \kappa \)-CN represents about 13% of caseins, and it is composed of a polypeptide chain of 169 amino acids of which valine represents the carboxy-terminus (Holt et al. 2013).

CONTACT Dr Marco Franzoi marco.franzoi@unipd.it Dipartimento di Agronomia, Animali, Alimenti, Risorse naturali e Ambiente, University of Padova, Viale dell’Università 16, Legnaro 35020, Italy

© 2021 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.
Main whey proteins include β-lactoglobulin (β-LG), synthesised within the udder; moreover, in smaller quantities, there are also serum albumin, immunoglobulins, proteose-peptones, lactoferrin, and transferrin (Kuzdzal-Savoie 1980; Davies et al. 1983; Jenness 1985). β-LG is a globular protein formed by a peptide chain of 162 amino acids and is the most represented among ruminant whey proteins (Pérez and Calvo 1995). The milk caseins β and κ, and also the whey proteins as β-LG have several genetic polymorphisms (Farrell et al. 2004). The protein variants, indeed, are distinguished from each other by the substitution or deletion of amino acids within the polypeptide chain (Martin et al. 1999).

Over the years, the interest in the genetic polymorphisms of milk proteins mainly concerned the influence that these can exert on the characteristics of milk and on human health (Caroli et al. 2009; Massella et al. 2017). In addition rennet clotting time (RCT), curd firming rate (k20), and curd firmness (a30) are affected by milk protein composition (Wedholm et al. 2006; Penasa et al. 2010). Several studies investigated the effects of milk protein polymorphisms on MCP (Buchberger and Dovc 2000) and protein composition (Heck et al. 2009); indeed and significant association between allele B of κ-CN and enhanced MCP has been reported (Lodes et al. 1996; Walsh, Guinee, Reville, et al. 1998; Buchberger and Dovc 2000; Hallén et al. 2008; Heck et al. 2009). The genetic variant B of β-CN improves the total content of casein in milk, promotes the formation of smaller casein micelles, and improves the coagulation properties of milk and cheese yield (Poulsen et al. 2013). The study of these polymorphisms turns out to be a valid practical application to improve the efficiency of dairy production, considering it could be possible to carry out animal selection for a specific variant within a protein. Nevertheless, it is not fully understood whether the interaction among protein genotypes and MCP is only determined by the protein fraction content or also by different three-dimensional conformation and chemical properties of protein variants. Moreover, cheese-making properties of individual cow milk are influenced by different factors such as herd management, parity, stage of lactation, and season, raising difficulties in disentangling co-variance of such effects with milk protein polymorphisms (Wedholm et al. 2006; Jensen et al. 2012; Penasa et al. 2016).

The common limitation of previous studies dealing with the relationship between milk protein genotypes and MCP was the limited number of observations for each genotype, due to the labor-intensive collection of MCP phenotypes. For the same reason, environmental factors such as the herd effect or seasonality were difficult to be assessed. The present work took advantage of the possibility to predict milk coagulation properties from milk mid-infra-red spectra within the frame of Dairy Herd Improvement programs (Penasa et al. 2015; Tiezzi et al. 2015). Therefore, the aim of the present study was to estimate the effects of β-CN, κ-CN, and β-LG single and composite genotype on milk quality traits and MCP of Italian Holstein–Friesians cows, using mid-infra-red spectroscopy (MIRS) data and chemometrics predictions.

**Material and methods**

**Data origin**

Information on individual cow milk samples from Dairy Herd Improvement (DHI) testing was provided by the Breeders Association of Veneto Region (Vicenza, Italy). Information about cow genotype was provided by the National Association of Friesian, Brown, and Jersey Breeders (Cremona, Italy). The genotypes of milk protein polymorphisms were carried out through Custom Bovine SNP50K Bead-Chip (Illumina Inc., USA). All sampling activities were carried out within routine DHI activities by Breeders Association of Veneto Region (Vicenza, Italy) and Italian Holstein, Brown and Jersey Association (Cremona, Italy) in full compliance with current legislation, in particular referring to animal welfare.

Samples were collected in the Veneto region between 2012 and 2021. The distribution of sampled cows over Veneto provinces is reported in Figure 1. Farms in the Veneto region are characterised by medium herd size (43 col), added with 20 μl of preservative (Bronysolv; AN.A.LI.TIK Austria, Vienna, Austria), were analysed in the laboratory of the Breeders Association of Veneto Region (Padova, Italy) according to guidelines of International Committee for Animal Recording (ICAR) for milk quality analyses. Somatic cell count (SCC) was assessed by Fossomatic (FOSS Electric A/S, Hillerød, Denmark) and transformed to somatic cell score (SCS) through the equation $SCS = \log_2(SCC/100) + 3$ (Wiggans and Shook 1987). Milk fat, protein, and casein percentage, milk urea (mg/dL), and predicted milk coagulation properties were assessed using different versions of Fourier Transformed – Mid Infra-red MilkoScan instruments (Foss, Hillerød, Denmark). According to the manufacturer instructions and ICAR committee, MilkoScan instruments were standardised to ensure the reliability of predicted traits (Feudale et al. 2002; Juhl 2017). Predictive equations...
for milk fat, protein, casein and lactose percentage, urea, and coagulation properties were purchased from the instrument producer and continuously updated at the newest available version.

Data editing
The original dataset \((n=132,835)\), including observations from 7783 genotyped animals, was edited to retain records from purebred HF, between 5 and 305 days in milk (DIM) reared in the same herd during the complete period of the study (2012 to 2021). Contemporary groups were defined as cows tested in the same herd and date (herd-test-date, HTD) and HTD with less than 3 animals, and cows with less than 3 observations within lactation were removed. Observation traits deviating more than 3 SD from the overall mean of the trait were set as missing. After editing, 74,721 observations from 5316 cows and 122 herds were available for further analyses. Parity and DIM averaged 1.96 ± 1.77 and 148.15 ± 84.47 days, respectively. Herd size ranged from 6 to 198 cows.

Statistical analysis
Milk composition, SCS, urea, and predicted milk coagulation properties were analysed through a linear mixed model using PROC HPMIXED of SAS ver. 9.4 (SAS Institute Inc., Cary, NC). The model was:

\[
y_{ijklm} = \mu + S_i + P_j + \beta - CN_k + \kappa - CN_l + \beta - LG_m + (S \times P)_ij + HTD_n + C_0(\beta - CN_k \times \kappa - CN_l \times \beta - LG_m) + e_{ijklmno}
\]

where \(y_{ijklm}\) is the dependent variable; \(\mu\) is the overall intercept of the model; \(S_i\) is the fixed effect of the \(i\)th class of stage of lactation (\(l = 1\) to 30, being all classes of 10 DIM from 6 to 305 DIM); \(P_j\) is the fixed effect of the \(j\)th parity of the cow (\(j = 1\) or 2, with 1 being primiparous cows and 2 being multiparous cows); \(\beta - CN_k\) is the fixed effect for cow \(\beta - CN\) genotype (\(k = A1A1, A1A2, A2A2\)); \(\kappa - CN_l\) is the fixed effect for cow \(\kappa - CN\) genotype (\(l = AA, AB, AE, BB, BE, EE\)); \(\beta - LG_m\) is the fixed effect for cow \(\beta - LG\) genotype (\(m = AA, AB, BB\)); \((S \times P)_ij\) is the fixed interaction effect between stage of
lactation and parity; HTD, is the random effect of the nth contemporary group (n = 1–3542); C孙(β-CNk × κ-CNl × β-LGm,n) is the random effect of the nth cow nested within the composite genotype (a = 1–5316) ~ N(0, σ2(β-CN)) over Cow, where σ2(β-CN) is the cow variance tested in within composite genotype variance; and eijklmno is the random residual ~ N(0, σ2), where σ2 is the residual variance. Moreover, the combined effect of genotypes was investigated separately substituting the fixed effect of single genotypes, that is, β-CN, κ-CN, and β-LG, with Gp fixed effect for composite genotype (p = 1–40). Composite genotypes with less than 150 observations or present in less than five cows were excluded. Multiple comparisons of least squares means were performed using Bonferroni’s adjustment. Significance was set at p < .05 unless otherwise stated. Heatmap and graphical representations were created using the package ggplot2 and Hierarchical Clustering Algorithms through the statistical software R (R Core Team 2013).

Results and discussion
Genotype and allele frequencies

Genotype and allele frequencies of analysed samples are reported in Table 1. Notably, even if in the last years the attention towards A2 milk increased globally due to potential positive effects on human health compared to other variants (Thiruvengadam et al. 2021), 39.13% of β-CN alleles in analysed samples was A1. Accordingly, only 36.49% of analysed cows were homozygous for A2 β-CN, whereas the remaining animals had at least one A1 allele. The correspondence between actual genotype frequency and the one expected from alleles percentage suggested no segregation between A1 and A2 cows is ongoing in Italian Holsteins. Previously, Chessa et al. (2020) investigated the temporal trend of allele frequencies in the Italian Holstein population, highlighting small differences between 2000 and 2010 and the period 2010 and 2017. During the latter period, A1 and A2 represented 36.7% and 55.8% of population alleles, respectively, in accordance with the results of the present study. The same authors reported the remaining alleles being mostly β-CN I, B, and A3. Such minor alleles were not investigated in the present study. Sebastiani et al. (2020) examined the frequencies of β-CN alleles in Holstein cows reared in 17 herds of central Italy and reported very similar A2 allele frequency compared to the present study (60.65%) and slightly lower A1 frequency (30.39%). Lower frequencies compared to the present study were expected, considering the aforementioned authors took in account also minor alleles such as B and I. Similar investigations in Holstein population, from several countries in Europe and Asia, reported A1 and A2 to represent about 36.9% (range from 26% to 50%), and about 57.1% (range from 26% to 50%) of β-CN alleles, respectively (Dai et al. 2016; Ciesliriska et al. 2019).

Frequencies of A allele of β-LG in analysed samples was 54.45%, in accordance with previous studies which reported frequencies between 27% and 57% in different Holstein populations, between late 90s and early 2000s (Heidari et al. 2012). Results differed from a previous investigation in Italian Holstein, which reported a frequency of the A allele of 43.6% (Caroli et al. 2004).

The most represented κ-CN allele was A (56.81%), followed by B (33.31%) and E (9.88%). Similar to Chessa et al. (2020), no individual carried κ-CN C variant. Even if the positive effect of B variant of κ-CN on milk technological properties was well-known, and tests for quantification of B protein variant of κ-CN are commercially available (Rossini et al. 2011), only 10.89% of tested cows were homozygous for B allele. The frequency of the B allele from the present study was confirmed by previous investigations in Holstein–Friesian cows predominantly from US, Italy, and Canada (Chessa et al. 2020).

Effect of β-CN genotype on milk composition and predicted MCP

The fat content of milk was not significantly influenced by β-CN genotype (Table 2). However, the statistical analysis highlighted significant, even if minor, differences in protein and casein content among β-CN genotypes, in agreement with findings of Lundén et al. (1997). Similarly, a recent investigation on Agerolese cattle breed found significantly higher protein and casein content in β-CN A2A2 milk compared to other genotypes. Previously, Ng-Kwai-Hang et al. (1986) found opposite results in Holstein–Friesian, even if the statistical significance of the differences

| Table 1. Genotypes and allele frequencies of β-casein, β-lactoglobulin, and κ-casein in milk of Holstein–Friesian cows. |
|-------------------------------|------------------------|-------------------------|
| Protein* | Frequency of genotypes, % | Frequency of alleles, % |
| β-CN | A1A1 | 14.75 | 48.76 | 36.49 | 39.13 | 60.87 |
| | A1A2 | 14.75 | 48.76 | 36.49 | 39.13 | 60.87 |
| | A2A2 | 14.75 | 48.76 | 36.49 | 39.13 | 60.87 |
| | A1 | 14.75 | 48.76 | 36.49 | 39.13 | 60.87 |
| | A2 | 14.75 | 48.76 | 36.49 | 39.13 | 60.87 |
| β-LG | AA | 29.59 | 49.74 | 20.67 | 54.45 | 45.55 |
| | AB | 29.59 | 49.74 | 20.67 | 54.45 | 45.55 |
| | BB | 29.59 | 49.74 | 20.67 | 54.45 | 45.55 |
| | A | 29.59 | 49.74 | 20.67 | 54.45 | 45.55 |
| | B | 29.59 | 49.74 | 20.67 | 54.45 | 45.55 |
| κ-CN | AA | 32.06 | 37.77 | 11.73 | 10.89 | 7.06 | 0.49 | 56.81 | 33.31 | 9.88 |
| | AB | 32.06 | 37.77 | 11.73 | 10.89 | 7.06 | 0.49 | 56.81 | 33.31 | 9.88 |
| | AE | 32.06 | 37.77 | 11.73 | 10.89 | 7.06 | 0.49 | 56.81 | 33.31 | 9.88 |
| | BB | 32.06 | 37.77 | 11.73 | 10.89 | 7.06 | 0.49 | 56.81 | 33.31 | 9.88 |
| | BE | 32.06 | 37.77 | 11.73 | 10.89 | 7.06 | 0.49 | 56.81 | 33.31 | 9.88 |
| | E | 32.06 | 37.77 | 11.73 | 10.89 | 7.06 | 0.49 | 56.81 | 33.31 | 9.88 |
between β-CN A2A2 and A1A1 genotypes was not reported. Moreover, GWAS meta-analysis highlighted the chromosomal region associated with caseins genes had a significant effect on milk protein content (van den Berg et al. 2020). Among β-CN genotypes, A1A1 showed the lowest SCS content, followed by A1A2 and A2A2. Previously, Vallas et al. (2012) did not find a significant association between SCS and β-κ-CN composite genotype.

Considering predicted milk coagulation properties, β-CN genotype did not significantly affect RCT, while β-CN A1A1 individuals showed significantly improved k20 and a30 compared to A1A2 and A2A2 (Table 2). Previous studies described dissimilar and not significant differences between β-CN A1 and A2 variants on RCT. Marziali and Ng-Kwai-Hang (1986) reported shorter RCT for A1A2 milk compared to A1A1, meanwhile, Ketto et al. (2017) findings suggested A2A2 milk having the less favourable RCT and k20 compared to A1A2. The favourable effect of A1 allele on RCT, compared to A2 allele, was supported by studies on Holstein and Jersey cows (Jensen et al. 2012; Poulsen et al. 2013), and on Finnish Ayrshire and Finnish Friesian cows (Ikonen et al. 1999). Moreover, Lodes et al. (1996) demonstrated that β-CN A1A1 had shorter RCT and k20 than β-CN A1A2 and β-CN A2A2 genotypes. The favourable effect of β-CN A1 on a30 was reported by Ketto et al. (2017), who demonstrated significantly superior performance in terms of a30 for A1A2 compared to A2A2 milk. Different behaviour among β-CN variants on MCP have been attributed to the different net charge of protein molecules, with A1 and B variants having one and two positively charged amino acids more than A2, respectively (McLean et al. 1987).

\textbf{Effect of β-LG genotype on milk composition and predicted MCP}

Milk fat content was higher for β-LG genotype BB compared to AA (Table 3), similarly to findings of McLean et al. (1984) on Friesian and Jersey cows, and from Bovenhuis et al. (1992) on Dutch-Friesians and Holstein–Friesians crosses. Nevertheless, the same results were not confirmed by Mayer et al. (1997) investigating 1742 Brown Swiss cows, suggesting possible differences across breeds. Also, Heck et al. (2009) did not find significant differences among β-LG genotypes in fat content testing 1732 Holstein–Friesian cows. Similarly, van den Berg et al. (2020) did not include β-LG variants among significant genes having an effect on fat content. Presented results, together with reported literature, suggested an indirect effect of β-LG variants on milk fat percentage, likely connected to linkage disequilibrium with other genes in the same chromosomal region.

Lundén et al. (1997) observed a significant effect of β-LG genotype on protein and casein content, in contrast with findings of the present study. Similarly, Hallén et al. (2008) found significant differences among b-LG genotypes when CN was expressed as a fraction of the major proteins (quantified using HPLC). Such differences were essentially linked to a lower level of total protein in b-LG BB samples and a slightly higher concentration of κ-CN compared to other genotypes.

The milk content of urea is usually referred to as an index of efficient transformation of feed nitrogen in the protein of milk (Bobbo et al. 2020). Cows carrying β-LG AA genotype resulted in the lower urea content in milk. Botaro et al. (2008) found similar, but not statistically significant, results in Girolando but not in Holstein breeds. Singh et al. (2015) highlighted a significantly lower SCS in Holstein cows (n = 126) carrying β-LG AA genotype compared to BB, in accordance with the results of the present study. Such results could be linked to different susceptibility to mastitis or to a different response of the immune system to udder infections among β-LG genotypes.
Table 4. Least squares means of milk composition and coagulation properties of κ-CN genotypes.

| κ-casein | AA | AB | AE | BB | BE | EE |
|----------|----|----|----|----|----|----|
| Fat, %   | 3.78 (0.01) | 3.80 (0.01) | 3.78 (0.01) | 3.79 (0.01) | 3.74 (0.02) | 3.82 (0.06) |
| Protein, % | 3.36 (0.01)b | 3.40 (0.01)a | 3.36 (0.01)b | 3.42 (0.01)a | 3.36 (0.01)b | 3.34 (0.03)b |
| Casein, % | 2.63 (0.01)c | 2.67 (0.01)b | 2.63 (0.01)c | 2.68 (0.01)c | 2.63 (0.01)c | 2.58 (0.03)c |
| Urea, mg/dl | 22.19 (0.10)c | 22.81 (0.10)b | 22.24 (0.13)c | 23.14 (0.13)a | 22.91 (0.15)b | 22.11 (0.37)c |
| RCT, min | 25.00 (0.14)a | 24.46 (0.13)bc | 24.90 (0.18)ab | 24.08 (0.18)c | 24.42 (0.21)bc | 24.71 (0.53)abc |
| k₂₀, min | 7.62 (0.02)ab | 7.21 (0.02)c | 7.59 (0.04)b | 6.85 (0.04)d | 7.36 (0.06)c | 7.70 (0.17)d |
| a₃₀, min | 18.06 (0.15)c | 19.83 (0.15)b | 18.47 (0.22)c | 21.47 (0.23)a | 19.79 (0.27)b | 19.09 (0.74)bc |

Note. Means with different superscript letters are statistically different (p < .05) according to Bonferroni post-hoc test. Standard errors of least squares means were reported in brackets.

SCS: Somatic cell score; RCT: rennet coagulation time; k₂₀: curd firming rate; a₃₀: curd firmness.

Beta-LG AA showed a less favourable RCT, k₂₀ and a₃₀ compared with the other investigated genotypes (Table 3). Previous research is not in accordance with the effect of β-LG on RCT. Reporting only studies with statistically significant differences among β-LG genotypes, Kübarsepp et al. (2005), Ng-Kwai-Hang et al. (2002), and Ketto et al. (2017) reported that the most favourable genotypes were AA, AB, and BB using different statistical approaches, respectively. Celik (2003) supported AA being the most favourable genotype analysing milk from 50 Brown Swiss cows. In the present study β-LG BB had shorter k₂₀ (7.19 min) than AB and AA genotypes. Such a result is in accordance with Kübarsepp et al. (2005).

The stability of the complex between κ-CN and β-LG has been reported to be influenced by protein variants (Imafidon et al. 1991). In this view, a deeper understanding of the effect of such interaction on the coagulation process from a molecular perspective and under controlled conditions could be useful to fully determine which is the β-LG variant having the most favourable effect on MCP.

**Effect of κ-CN genotype on milk composition and predicted MCP**

Even if previous researches dealt with the effect of κ-CN on fat content, most data focussed only on A and B variants (Bovenhuis et al. 1992; Mayer et al. 1997). According to findings presented in Table 4, Heck et al. (2009) did not find significant differences in milk fat percentage among κ-CN genotypes.

Protein content, as well as casein content of milk, were influenced by κ-CN genotypes, with BB being the most favourable one (Table 4). Similar results were found by Hallén et al. (2008), Heck et al. (2009), and Mackle et al. (1999) using different computational approaches. In particular, Heck et al. (2009) reported a 0.19% difference in milk protein content between BB and EE genotypes, similar to the 0.18% of the present study. On the other hand, Albarella et al. (2020) did not find significant differences in protein and casein contents among Agerolese breed individuals carrying κ-CN AB and BB genotypes.

κ-CN EE genotype demonstrated significantly lower urea compared to AB, BB, and BE genotypes (Table 4). To the authors’ knowledge, this is the first time κ-CN protein variants are associated with changes in milk urea content. SCS in the analysed individuals was lower for κ-CN BB genotype compared to AA. Previously, Vallas et al. (2012) did not find a significant association between SCS and β-κ-CN composite genotype.

A significant effect of κ-CN on MCP was found, with the BB genotype performing significantly better in terms of RCT compared to AA (Jakob and Puhan 1992; Walsh, Guinee, Harrington, et al. 1998) (Table 4). Accordingly, Vallas et al. (2012) findings demonstrated allele B of κ-CN having the best coagulation characteristics compared to A variant. Moreover, Ikonen et al. (1999), Kübarsepp et al. (2005), and Jensen et al. (2012) supported present findings suggesting κ-CN BB genotype was linked to improved RCT. A possible explanation for such difference is that κ-CN B determines smaller casein micelles compared to κ-CN A (Walsh, Guinee, Harrington, et al. 1998). This possibly leads to a lower critical proportion of κ-CN to be proteolysed in order to initiate micelles aggregation, essentially due to increased diffusion of micelles (McMahon and Brown 1984; Dalgleish 1993).

Individuals carrying κ-CN BB genotypes were characterised by a significantly shorter k₂₀ compared to all other κ-CN genotypes. Walsh, Guinee, Harrington, et al. (1998) investigated the effect of κ-CN variants on low-moisture mozzarella cheese manufacturing and reported a significantly improved rate of curd firming in milk with κ-CN BB compared to AA. Finally, the effect of κ-CN genotype on a₃₀ was in accordance with previous studies (Lodes et al. 1996; Walsh, Guinee, Harrington, et al. 1998).
**Effect of composite genotype on milk composition**

The effect of genotypes on milk composition is represented as a heatmap in Figure 2. To generate the heatmap, standardised least squares means and standard errors were imputed. The map highlighted five principal clusters. The first cluster is characterised by average SCS, high urea levels, and the highest percentages of fat, protein, and casein. Such cluster is mostly represented by genotypes composed of $\beta$-CN A2A2, $\beta$-LG BB and $\kappa$-CN BB. Such results are in agreement with the findings of Mayer et al. (1997), even if a direct comparison is not possible considering cited authors did not include $\beta$-CN A1 variant among investigated alleles. The second cluster was characterised by low SCS levels, average urea, and good protein and casein levels. Such characteristics suggested genotypes belonging to this cluster to be favourable for udder health, efficiency, and production. Genotypes of the cluster are represented by $\beta$-CN A1A2 followed by A1A1, $\beta$-LG AA and AB, and $\kappa$-CN AB followed by BB. In this context, $\beta$-CN A1, $\beta$-LG A, and $\kappa$-CN B alleles are the most represented. The third cluster is composed of a single genotype, A1A1, BB, and BB for $\beta$-CN, $\beta$-LG, and $\kappa$-CN, respectively. This genotype is characterised by an average level of SCS, extremely low urea, and high levels of protein, casein and fat. Nevertheless, such results could be biased by the low number of animals carrying such genotype ($n = 17$, Supplementary Table 1). The fourth cluster is characterised by medium-high SCS and average to low protein and casein content. $\beta$-CN and $\beta$-LG genotypes are almost equally represented, while the presence of $\kappa$-CN E allele is particularly high in this cluster. Finally, the fifth cluster is represented by genotypes with low protein, casein, and fat contents. All genotypes carry at least one copy of $\beta$-CN A1 allele, and most of them also have a $\kappa$-CN E copy. To the authors’ knowledge, this is the first time a clustering approach was applied...
to milk compositional characteristics in order to highlight composite genotypes with similar milk compositions. Perna et al. (2016) studied the effect of caseins haplotype on milk composition and suggested BB-A2A2-BB genotype for \( \beta \)-CN-\( \beta \)-CN-\( \kappa \)-CN having good compositional characteristics and average SCS levels, in accordance to present findings. The same authors reported the highest levels of fat for BB-A1A1-AA casein haplotype despite average levels of protein, and the highest SCS. Accordingly, the same haplotype for \( \beta \)-CN-\( \kappa \)-CN in the present study clustered in the fourth group, characterised by similar composition except for the genotype carrying \( \beta \)-LG AB variants, having lower SCS levels.

**Effect of composite genotype on MCP**

The effect of milk protein variants on MCP was represented as a heatmap in Figure 3. Four principal clusters were observed. The first cluster is mainly represented by genotypes with poor coagulation properties in terms of RCT, \( k_{20} \), and \( a_{30} \). Four genotypes belong to this cluster, all presenting at least one \( \kappa \)-CN A allele. The second and largest cluster presented lower \( a_{30} \) and average to poor RCT and \( k_{20} \). Notably, most genotypes carrying \( \kappa \)-CN A and E alleles belong to cluster, while no \( \kappa \)-CN BB genotype is clustered in this group. The third cluster gathered genotypes with good \( a_{30} \) and average to good RCT and \( k_{20} \). In this group, most of the genotypes carrying \( \kappa \)-CN BB and AB are present, while \( \kappa \)-CN E allele was not present. Accordingly, the favourable effect of \( \beta \)-CN-\( \kappa \)-CN A1A1-AB genotype on \( a_{30} \) was previously demonstrated in Finnish Ayrshire cows, even if the same results were not confirmed in Finnish Friesian. Considering \( k_{20} \), Finnish Friesian \( \beta \)-CN-\( \kappa \)-CN genotypes A1A1-AB and A1A2-AB were proposed as the most favourable. However, the same study did not take into

---

**Figure 3.** Hierarchical clustering of standardised milk coagulation properties least squares means among the genotypes of \( \beta \)-casein (\( \beta \)-CN), \( \beta \)-lactoglobulin (\( \beta \)-LG), and \( \kappa \)-casein (\( \kappa \)-CN). \( a_{30} \): curd firmness; \( k_{20} \): curd firming rate; RCT: rennet coagulation time.
consideration homozygous animals for κ-CN B (Ikonen et al. 1997). Finally, the fourth cluster is characterised by very good MCP. All genotypes belonging to such cluster were uncommon in Holstein population, being present in 0.32%, 1.8%, and 0.49% of analysed cows for genotypes A1A1-BB-BB, A2A2-AB-BB, and A1A1-BB-BE of β-CN, β-LG, κ-CN, respectively. Notably, one of the genotypes represented in such cluster is A1A1-BB-BB, the same demonstrating extremely good compositional traits. It is important to notice that in this cluster is present a composite genotype with A2A2 β-CN. The good performance in terms of technological traits of A2A2-AB-BB composite genotype, similar to results reported by Perna et al. (2016) for BB-A2A2-BB (CSN1S1-CSN2-CSN3) genotype, could be attributed to a synergistic effect of β-CN A2 and κ-CN B on micelle size (Walsh, Guinee, Reville, et al. 1998; Ketto et al. 2017).

**Conclusions**

The present study demonstrated the effect of single and composite milk protein genotypes on milk composition and milk coagulation properties, predicted using milk infra-red spectroscopy, from a large database of Holstein cows. β-CN A2A2 cows produced milk with higher protein and casein content compared to other genotypes, while β-CN A1A1 reported the most favourable urea level, SCS, κ20, and a30, while no significant differences among genotypes were found for fat content and RCT. The most favourable β-LG genotype was BB for fat content and milk coagulation properties, while the AA genotype resulted in lower urea and SCS. Finally, κ-CN BB was the best genotype for most of the analysed traits, except for fat content, resulting in no significant differences among genotypes, and urea, being EE and AE and AA the most favourable genotypes. Moreover, the relationships in terms of milk composition and milk coagulation properties among milk protein composite genotypes were investigated using hierarchical clustering, suggesting the existence of genotypes sharing common characteristics. This aspect could be taken into consideration for future selection plans. Finally, considering the increasing interest in β-CN A2A2 milk from a health point of view, the detrimental effect of the A2 allele on milk technological traits should be taken into account. The present study highlighted the existence of composite genotypes carrying β-CN A2A2 and having average to good MCP. Such genotypes should be considered to counteract the possible detrimental effects of A2A2 selection.

**Ethical approval**

Present study was developed in the framework of national dairy herd improvement program. All activities involving animals were conducted by state-recognized Dairy Herd Improvement Associations. All applicable international, national, and/or institutional guidelines for the care of animals were followed.

**Acknowledgements**

We thank the Breeders Association of Veneto Region – ARAV (Vicenza, Italy) and the National Association of Friesian, Brown and Jersey Breeders – ANAFIBJ (Cremona, Italy) for the support and for the dataset.

**Disclosure statement**

This article has been accepted for publication in the *Italian Journal of Animal Science*, published by Taylor & Francis. The authors are fully responsible for the content of the manuscript and declare that there is no conflict of interest associated with the paper.

**ORCID**

Filippo Cendron [http://orcid.org/0000-0002-8221-7566](http://orcid.org/0000-0002-8221-7566)

Marco Franzoi [http://orcid.org/0000-0001-8701-5632](http://orcid.org/0000-0001-8701-5632)

Mauro Penasa [http://orcid.org/0000-0001-9984-8738](http://orcid.org/0000-0001-9984-8738)

Massimo De Marchi [http://orcid.org/0000-0001-7814-2525](http://orcid.org/0000-0001-7814-2525)

Martino Cassandro [http://orcid.org/0000-0002-8709-2870](http://orcid.org/0000-0002-8709-2870)

**Data availability statement**

The data that support the findings of this study are available from the corresponding author, MF, upon reasonable request.

**References**

Albarella S, Selvaggi M, D’Anza E, Cosenza G, Caira S, Scaloni A, Fontana A, Peretti V, Ciottola F. 2020. Influence of the casein composite genotype on milk quality and coagulation properties in the endangered agerolese cattle breed. Animals. 10(5):892.

Botaro BG, Lima YVR, Aquino AA, Fernandes RHR, Garcia JF, Santos MV. 2008. Effect of beta-lactoglobulin polymorphism and seasonality on bovine milk composition. J Dairy Res. 75(2):176–181.

Bovenhuis H, Van Arendonk JAM, Korver S. 1992. Associations between milk protein polymorphisms and milk production traits. J Dairy Sci. 75(9):2549–2559.

Buchberger J, Dovc P. 2000. Lactoprotein genetic variants in cattle and cheese making ability. Food Technol Biotechnol. 38(2):91–98.
Caroli A, Chessa S, Bolla P, Budelli E, Gandini GC. 2004. Genetic structure of milk protein polymorphisms and effects on milk production traits in a local dairy cattle. J Anim Breed Genet. 121(2):119–127.

Caroli AM, Chessa S, Erhardt GJ. 2009. Invited review: milk protein polymorphisms in cattle: effect on animal breeding and human nutrition. J Dairy Sci. 92(11):5335–5352.

Celik S. 2003. β-Lactoglobulin genetic variants in Brown Swiss breed and its association with compositional properties and rennet clotting time of milk. Int Dairy J. 13(9):727–731.

Chessa S, Gattolin S, Cremonesi P, Soglia D, Finocchiaro R, Kaam J-TV, Marusi M, Civati G. 2020. The effect of selection on casein genetic polymorphisms and haplotypes in Italian Holstein cattle. Ital J Anim Sci. 19(1):833–839.

Cieślińska A, Fiedorowicz E, Zwierzchowska G, Kordulewska N, Jarmolowska B, Kostyra E. 2019. Genetic polymorphism of β-casein gene in polish red cattle-preliminary study of A1 and A2 frequency in genetic conservation herd. Animals. 9(6):377.

Dai R, Fang Y, Zhao W, Liu S, Ding J, Xu K, Yang L, He C, Ding F, Meng H. 2016. Identification of alleles and genotypes of beta-casein with DNA sequencing analysis in Chinese Holstein cow. J Dairy Res. 83(3):312–316.

Dalglish DG. 1993. The enzymatic coagulation of milk. In: Fox PF, editor. Cheese: chemistry, physics and microbiology; volume 1 general aspects. New York (NY): Springer; p. 69–100.

Davies DT, Holt C, Christie WW. 1983. Biochemistry of lactation. In: Mepham TB, editor. The composition of milk. Amsterdam (The Netherlands): Elsevier; p. 71–117.

Farrell HM, Jimenez-Flores R, Bleck GT, Brown EM, Butler JE, Creamer LK, Hicks CL, Hollar CM, Ng-Kwai-Hang KF, Swaisgood HE. 2004. Nomenclature of the proteins of the genetic variants of k-casein in the cheese production. J Dairy Sci. 87(6):1641–1674.

Feudale RN, Woody NA, Tan H, Myles AJ, Brown SD, Ferré J. 2002. Transfer of multivariate calibration models: a review. Chemometr Intell Lab Syst. 64(2):181–192.

Hallén E, Wedholm A, Andrén A, Lundén A. 2008. Effect of beta-casein, kappa-casein and beta-lactoglobulin genotypes on concentration of milk protein variants. J Anim Breed Genet. 125(2):119–129.

Heck JML, Schennink A, van Valenberg HJF, Boovenhuis H, Visker MHPW, van Arendonk JAM, van Hooijdonk ACM. Effects of bovine milk protein and milk composition, casein micelle and salt properties and composition of bovine milk, Journal of Dairy Science, Volume 92, Issue 3, 2009, Pages 1192–1202, ISSN 0022-0302, https://doi.org/10.3168/jds.2008-1208. (https://www.sciencedirect.com/science/article/pii/S002203020907428X)

Heidari M, Azari MA, Hasani S, Khanahmadi A, Zerehadaran S. 2012. Effect of polymorphic variants of GH, Pit-1, and β-LG genes on milk production of Holstein cows. Russ J Genet. 48(4):417–421.

Holt C, Carver JA, Ecroyd H, Thorn DC. 2013. Invited review: caseins and the casein micelle: their biological functions, structures, and behavior in foods. J Dairy Sci. 96(10):6127–6146.

Ikonen T, Ahlfors K, Kempe R, Ojala M, Ruottinen O. 1999. Genetic parameters for the milk coagulation properties and prevalence of noncoagulating milk in Finnish dairy cows. J Dairy Sci. 82(1):205–214.

Ikonen T, Ojala M, Syväöja E-L. 1997. Effects of composite casein and β-lactoglobulin genotypes on renneting properties and composition of bovine milk by assuming an animal model. AFSci. 6(4):283–294.

Imafidon GI, Ng-Kwai-Hang HF, Harwalkar VR, Ma C-Y. 1991. Effect of genetic polymorphism on the thermal stability of β-Lactoglobulin and κ-Casein mixture. J Dairy Sci. 74(6):1791–1802.

Jakob E, Puhan Z. 1992. Technological properties of milk as influenced by genetic polymorphism of milk proteins—a review. Int Dairy J. 2(3):157–178.

Jenness R. 1985. Biochemical and nutritional aspects of milk and colostrum. In: Larson BL, editor. Lactation. Ames (IA): The Iowa State University Press; p. 164–197.

Jensen HB, Poulsen NA, Andersen KK, Hammershøj M, Poulsen HD, Larsen LB. 2012. Distinct composition of bovine milk from Jersey and Holstein-Friesian cows with good, poor, or noncoagulation properties as reflected in protein genetic variants and isoforms. J Dairy Sci. 95(12):6905–6917.

Juhl HV, inventor; Foss Analytical AS, assignee. 2017 March 28. Method for compensating amplitude drift in a spectrometer and spectrometer performing said method. United States Patent US 20140114601A1.

Ketto IA, Knutsen TM, Øyaas J, Heringstad B, Ådnøy T, Devold TG, Skieie SB. 2017. Effects of milk protein polymorphism and composition, casein micelle size and salt distribution on the milk coagulation properties in Norwegian Red cattle. Int Dairy J. 70:55–64.

Kübarsepp I, Henno M, Vinalas H, Sabre D. 2005. Effect of κ-casein and β-lactoglobulin genotypes on the milk rennet coagulation properties. Agron Res. 3(1):55–64.

Kuzdzal-Savoie S. 1980. Determination of free fatty acids in milk and milk products. In: Moore JH, editor. Flavor impairment of milk and milk products due to lipolysis. Brussels (Belgium): FIL-IDF.

Lodes A, Buchberger J, Krause I, Aumann J, Kloster-Meyer H. 1996. The influence of genetic variants of milk proteins on the compositional and technological properties of milk. 2. Rennet coagulation time and firmness of the rennet cured. Milchwissenschaft. 51(10):543–548.

Lundén A, Nilsson M, Janson L. 1997. Marked effect of β-lactoglobulin polymorphism on the ratio of casein to total protein in milk. J Dairy Sci. 80(11):2996–3005.

Mackle TR, Bryant AM, Petch SF, Hill JP, Auldist MJ. 1999. Nutritional influences on the composition of milk from cows of different protein phenotypes in New Zealand. J Dairy Sci. 82(1):172–180.

Mariani P, Pecorari M. 1991. Il ruolo delle varianti genetiche della k-caseina nella produzione del formaggio [The role of the genetic variants of k-casein in the cheese production]. Scienza e Tecnica Lattiero Casearia. 42(4):255–285. Italian.

Martin P, Ollivier-Bousquet M, Grosclaude F. 1999. Genetic polymorphism of caseins: a tool to investigate casein micelle organization. Int Dairy J. 9(3–6):163–171.

Marziali AS, Ng-Kwai-HangKF. 1986. Effects of milk composition and genetic polymorphism on coagulation properties of milk. J Dairy Sci. 69(7):1793–1798.

Massella E, Piva S, Giacometti F, Liuizzo G, Zambrini AV, Serrano A. 2017. Evaluation of bovine beta casein
polymerase in two dairy farms located in northern Italy. Ital J Food Saf. 6(3):6904.

Mayer HK, Ortner M, Tschager E, Ginzinger W. 1997. Composite milk protein phenotypes in relation to composition and cheesemaking properties of milk. Int Dairy J. 7(5):305–310.

McLean DM, Graham ERB, Ponzoni RW, McKenzie HA. 1984. Effects of milk protein genetic variants on milk yield and composition. J Dairy Res. 51(4):531–546.

McLean DM, Graham ERB, Ponzoni RW, Mckenzie HA. 1987. Effects of milk protein genetic variants and composition on heat stability of milk. J Dairy Res. 54(2):219–235.

McMahon DJ, Brown RJ. 1984. Enzymic coagulation of casein micelles: a review. J Dairy Sci. 67(5):919–929.

Ng-Kwai-Hang KF, Dodds C, Boland M, Auldist MJ. 2002. The influence of genetic variants of beta-lactoglobulin on gelation speed and firmness of rennet curd. Milchwissenschaft. 57:267–269.

Ng-Kwai-Hang KF, Hayes JF, Moxley JE, Monardes HG. 1986. Relationships between milk protein polymorphisms and major milk constituents in Holstein–Friesian cows. J Dairy Sci. 69(1):22–26.

Penasa M, Cassandro M, Pretto D, Marchi MD, Comin A, Chessa S, Zotto RD, Bittante G. 2010. Short communication: influence of composite casein genotypes on additive genetic variation of milk production traits and coagulation properties in Holstein–Friesian cows. J Dairy Sci. 93(7):3346–3349.

Penasa M, De Marchi M, Ton S, Ancilotto L, Cassandro M. 2015. Reproducibility and repeatability of milk coagulation properties predicted by mid-infrared spectroscopy. Int Dairy J. 47:1–5.

Penasa M, Toffanin V, Cologna N, Cassandro M, Marchi MD. 2016. Effects of dairy factory, milk casein content and titratable acidity on coagulation properties in Trentingrana dairy industry. J Dairy Res. 83(2):242–248.

Pérez MD, Calvo M. 1995. Interaction of beta-lactoglobulin with retinol and fatty acids and its role as a possible biological function for this protein: a review. J Dairy Sci. 78(5):978–988.

Perna A, Intaglietta I, Gambacorta E, Simonetti A. 2016. The influence of casein haplotype on quality, coagulation, and yield traits of milk from Italian Holstein cows. J Dairy Sci. 99(5):3288–3294.

Poulsen NA, Bertelsen HP, Jensen HB, Gustavsson F, Glantz M, Lindmark Månsson H, Andrén A, Paulsson M, Bendixen C, Buitenhuis AJ, et al. 2013. The occurrence of noncoagulating milk and the association of milk coagulation properties with genetic variants of the caseins in 3 Scandinavian dairy breeds. J Dairy Sci. 96(8):4830–4842.

R Core Team. 2013. R: a language and environment for statistical computing. Vienna (Austria): R Foundation for Statistical Computing.

Rollemo HS. 1992. Casein association and micelle formation. In: Fox PF, editor. Advanced dairy chemistry. 2nd ed. Vol 1. Essex (GB): Elsevier Science Publisher, Ltd; p. 111–140.

Rossoni A, Nicoletti C, Santos E, Malacarne M. 2011. Analysis of the contents of k-casein B in bulk milk samples. Scienza e Tecnica Lattiero-Casearia. 62(4):267–273.

Sebastiani C, Arcangeli C, Ciullo M, Torricelli M, Cinti G, Fischella S, Biagetti M. 2020. Frequencies evaluation of β-Casein gene polymorphisms in dairy cows reared in central Italy. Animals. 10(2):252.

Singh U, Deb R, Kumar S, Singh R, Sengar G, Sharma A. 2015. Association of prolactin and beta-lactoglobulin genes with milk production traits and somatic cell count among Indian Frieswal (HF × Sahiwal) cows. Biomark Genom Med. 7(1):38–42.

Thiruvengadam M, Venkidasamy B, Thirupathi P, Chung I-M, Subramanian U. 2021. β-Casomorphin: a complete health perspective. Food Chem. 337:127765.

Tiezzi F, Valente BD, Cassandro M, Maltecca C. 2015. Causal relationships between milk quality and coagulation properties in Italian Holstein–Friesian dairy cattle. Genet Sel Evol. 47(1):45.

Vallas M, Kaart T, Värvi S, Pärna K, Jõedu I, Viinalass H, Pärna E. 2012. Composite k-casein genotypes and their effect on composition and coagulation of milk from Estonian Holstein cows. J Dairy Sci. 95(11):6760–6769.

van den Berg I, Xiang R, Jenko J, Pausch H, Boussaha M, Schrooten C, Tribout T, Gjusvland AB, Boichard D, Nordbro Ø, et al. 2020. Meta-analysis for milk fat and protein percentage using imputed sequence variant genotypes in 94,321 cattle from eight cattle breeds. Genet Sel Evol. 52(1):37.

Walsh CD, Guinee TP, Harrington D, Mehra R, Murphy JJ, Fitzgerald RJ. 1998. Cheesemaking, compositional and functional characteristics of low-moisture part-skim Mozzarella cheese from bovine milks containing k-casein AA, AB or BB genetic variants. J Dairy Res. 65(2):307–315.

Walsh CD, Guinee TP, Reville WD, Harrington D, Murphy JJ, O’Kennedy BT, Fitzgerald RJ. 1998. Influence of k-Casein genetic variant on rennet gel microstructure, cheddar cheesemaking properties and casein micelle size. Int Dairy J. 8(8):707–714.

Walstra P, Jenness R. 1984. Dairy chemistry and physics. New York (NY): Wiley.

Wedholm A, Larsen LB, Lindmark-Månsson H, Karlsson AH, Andrén A. 2006. Effect of protein composition on the cheese-making properties of milk from individual dairy cows. J Dairy Sci. 89(9):3296–3305.

Wiggans GR, Shook GE. 1987. A lactation measure of somatic cell count. J Dairy Sci. 70(12):2666–2672.