DOI: 10.5965/223811712312024117

Revista de Ciências Agroveterinárias 23 (1): 2024
Universidade do Estado de Santa Catarina

Milk protein polymorphisms and casein haplotypes in Blanco Orejinegro cattle of Colombia

Polimorfismos de proteína do leite e haplótipos de caseína em gado Blanco Orejinegro da Colômbia

Juan C Rincon-Florez (ORCID 0000-0002-6769-6407), Maryi N Pulido-Hoyos M (ORCID 0009-0004-9572-0576), Darwin Y Hernandez-Herrera* (ORCID 0000-0002-1758-604X)

Universidad Nacional de Colombia, Palmira, Colombia. *Author for correspondence: dyhernandezh@unal.edu.co

Submission: 05/09/2023 | Acceptance: 08/11/2023

ABSTRACT

The aim was to determine the genetic variation in the *CSN1S1, CSN2, CSN1S2, CSN3* and *LGB* genes in Blanco Orejinegro cattle. 419 animals from 15 herds were genotyped with GGP Bovine 150 K (n= 70) and 50 K (n= 349) chips. Information was obtained from 43 SNPs in the mentioned genes and protein variants **B, *C* and **D* of *α_{S1}-CN*; **A^1*, **A^2*, **B*, **H^2* and **F* of *β-CN*; **A* and **D* of *α_{S2}-CN, *A*, **A^1*, **B*, **I* and **H* of *κ-CN* and **A*, **B*, **C*, **D*, **E*, **F* and **H* of *β-LG* were reconstructed. Allele and genotypic frequencies were estimated for SNPs and for protein variants; Hardy-Weinberg equilibrium and F_{ST} values were evaluated for each of the SNPs under different structuring criteria. LD values and haplotypic frequencies were estimated for caseins. The most frequent variants were *CSN1S1*B* (0.804), *CSN2*A^2* (0.509), *CSN1S2*A* (0.997), CSN3*A (0.679) and *β-LG*B* (0.657). None of the variants showed deviations from HWE, but the *CSN2*A^2* allele showed a slight increasing trend over time. The F_{ST} values were low (0.035) regardless of the structuring criteria. Twenty-eight *CSN1S1-CSN2-CSN1S2-CSN3* haplotypes were found, 22 of them with frequencies <5%; the three most frequent were *BB-A^1A^2-AA-AA-AA-AA* (16.6%), *BB-A^1A^2-AA-AA-AA-AB* (14.1%) and *BB-A^2A^2-AA-AA-AA* (10.1%). A good potential of BON cattle to produce high quality milk with functional value was reported.

KEYWORDS: beta-lactoglobulin; cow´s milk; CSN gene variants; milk proteins.

RESUMO

O objetivo foi determinar a variação genética nos genes *CSN1S1, CSN2, CSN1S2, CSN3* e *LGB* em bovinos Blanco Orejinegro. 419 animais de 15 rebanhos foram genotipados com chips GGP Bovine 150 K (n= 70) e 50 K (n= 349). Foram obtidas informações de 43 SNPs nos genes mencionados e as variantes proteicas **B, *C* e **D* de *αS1-CN*; **A1, *A2, *B, *H2* e **F* de *β-CN*; **A* e **D* de *αS2-CN*, **A, *A1, *B, *I* e **H* de *κ-CN* e **A, *B, *C, *D, *E, *F* e **H* de *β-LG* foram reconstruídas. As frequências alélicas e genotípicas foram estimadas para SNPs e para variantes de proteínas; o equilíbrio de Hardy-Weinberg e os valores de F_{ST} foram avaliados para cada um dos SNPs sob diferentes critérios de estruturação. Os valores de LD e as frequências haplotípicas foram estimados para as caseínas. As variantes mais frequentes foram *CSN1S1*B* (0.804), *CSN2*A2* (0.509), *CSN1S2*A* (0.997), *CSN3*A* (0.679) e *β-LG*B* (0.657). Nenhuma das variantes apresentou desvios da HWE, mas o alelo *CSN2*A2* apresentou uma leve tendência de aumento ao longo do tempo. Os valores de F_{ST} foram baixos (0.035) independentemente dos critérios de estruturação. Foram encontrados 28 haplótipos *CSN1S1-CSN2-CSN1S2-CSN3*, 22 deles com frequências <5%; os três mais frequentes foram *BB-A^1A^2-AA-AA-AA-AA* (16.6%), *BB-A^1A^2-AA-AA-AA-AB* (14.1%) e *BB-A^2A^2-AA-AA-AA* (10.1%). Foi relatado um bom potencial do gado BON para produzir leite de alta qualidade com valor funcional.

PALAVRAS-CHAVE: beta-lactoglobulina; leite de vaca; variantes do gene CSN; proteínas do leite.

INTRODUCTION

The major proteins in bovine milk are the caseins *α_{S1}-CN*, *β-CN*, *α_{S2}-CN* and *κ-CN*, encoded by *CSN1S1*, *CSN2*, *CSN1S2* and *CSN3* genes, respectively, whereas the whey proteins are *α-LA* and *β-LG*, encoded by the genes *LAA* and *LGB* respectively (FARRELL et al. 2004). Genetic variation at these loci is high and they have been studied in several dairy breeds in countries such as Benin (VANVANHOSSOU et al. 2021), France (SANCHEZ et al. 2020), Italy (CHESSA et al. 2020), Czech Republic (KYSELOVÁ et al. 2019), Panama (VILLALOBOS-CORTÉS et al. 2023) and Sudan (AHMED et al. 2017); and in *Bos taurus* (KOLENDA & SITKOWSKA 2021), *Bos indicus* (MUKESH et al. 2022) and their crosses (MOHAN et al. 2021). The casein genes are located in tandem between positions 85.4 and 85.7 Mb on chromosome 6 (ARS-UCD1.2 genome assembly), suggesting that they may be inherited as a haplotype rather than independently (VILLALOBOS-CORTÉS et al. 2023). In contrast, the *LAA* and *LGB* genes are located on chromosomes 5 (31.2 Mb) and 11 (103.2 Mb), respectively (SANCHEZ et al. 2020).

Caseins are important proteins because they are associated with high dairy yield, milk quality and technological performance. For example, a higher amount of caseins and whey proteins has been associated with the *CSN1S1* gene (MOHAN et al. 2021); high milk production has been associated with the *CSN1S2* gene (ARDICLI et al. 2018); the *CSN2* gene is not only associated with high production and quality (MOHAN et al. 2021, VIGOLO et al. 2022), but also with a healthier product, specifically the animal milk with the *$*A^2$* variant, as the *$*A^1$* variant during its digestion digests the bioactive pet *β-casomorphin-7* properties have opioid properties, and has been postulated as a risk factor in the development of diseases such as atherosclerosis, type I diabetes, sudden infant death syndrome, neurological diseases and milk allergy (CHESSA et al. 2020, THIRUVENGADAM et al. 2021, VILLALOBOS-CORTÉS et al. 2023).

On the other hand, the CSN3 gene has been linked to the structural stability of casein micelles, which may affect cheese quality by reducing coagulation time and improving curd firmness (POULSEN et al. 2017, ZEPEDA-BATISTA et al. 2017). Similarly, the *BB-A^2A^2-BB* haplotype of *CSN1S1-CSN2-CSN3* has been associated with higher milk protein content, better coagulation, better ability to retain fat and protein in the curd, and consequently higher cheese yield compared to the *BB-A^1A^1-AA* haplotype (PERNA et al. 2016a, b). Likewise, variations in the *LAA* and *LGB* genes, which account for about 90% of whey proteins, can influence milk protein composition, curd formation and cheese yield (MACEDO et al. 2020).

The Blanco Orejinegro (BON) breed of cattle is one of the most important Colombian Creole breeds, adapted to high tropical conditions. It is mainly used in dual-purpose production systems (milk and meat), with a birth weight of 28.39±3.77 kg for males and 26.68±3.59 kg for females. Adjusted weaning weight at 240 days was 165±29.16 kg and adjusted weight at 480 days was 215.38±30.57 kg. Mean age at first calving 1104±141 days and calving interval 487±147 days. In terms of milk production, it averages 3.6 l/cow per day and 1439 l/lactation corrected to 270 days, with compositional values of fat 4.39%, protein 3.92%, solids non-fat 9.69%, solids total 14.08% and lactose 5.04% (MARTÍNEZ et al. 2012c).

Unlike the other American Creole breeds (brown in color), it has pigmented epidermis and mucous membranes, with a predominantly white, grey, and sarda-colored coat (MARTÍNEZ et al. 2012c). It has Iberian and British ancestry (MARTÍNEZ et al. 2012a) and more than 500 years of adaptation to tropical conditions (CAIVIO-NASNER et al. 2021b), which has allowed it to develop traits such as hardiness (ARRIETA et al. 2021), resistance to some diseases (LÓPEZ-HERRERA et al. 2001) and parasites (ROCHA et al. 2019), adaptation to climate (LEÓN et al. 2019), ability to digest low quality forage (MARTÍNEZ et al. 2012b) and good productive (CAIVIO-NASNER et al. 2021c) and reproductive efficiency (RIVEROS et al. 2022). This performance, together with its resistance and adaptive traits, confirms its value as an important genetic resource for the tropics, not only in Colombia.

Genomic studies on genetic diversity (CAIVIO-NASNER et al. 2021a), allele frequencies associated with genetic diseases (CAIVIO-NASNER et al. 2021b), selection indices (AMAYA et al. 2022) and genome-wide association studies on growth traits (LONDOÑO-GIL et al. 2021) have also been carried out. However, the study of genes related to milk production and quality has not been advanced in this breed. In Colombia, studies have only been carried out in Holsteins from the departments of Antioquia (PADILLA-DOVAL et al. 2021) and Nariño (SOLARTE et al. 2009), while only one study reports on the variation of the *LAA* and *LGB* genes in local breeds (ROSERO-ALPALA et al. 2011) and one on the *CSN3* gene in the Hartón del Valle breed (NARANJO et al. 2007). Therefore, the aim of this study was to determine the genetic variation in the *CSN1S1, CSN2, CSN1S2, CSN3* and *LGB* genes in Blanco Orejinegro cattle.

MATERIAL AND METHODS

Animals and DNA extraction

The study population consisted of 419 animals (72% females and 28% males) from 15 herds in the departments of Antioquia (n= 92), Caldas (n= 60), Cundinamarca (n= 48), Meta (n= 36), Risaralda (n= 131) and Tolima (n= 52). Between 11 and 60 animals per herd were sampled. The animals were born between 2004 and 2018, with the lowest number of individuals sampled in 2005 (n= 13) and the highest in 2017 (n= 56).

Using the GeneJET Genomic DNA Purification Kit (Thermo Scientific™), DNA was extracted from 404 whole blood samples obtained by venipuncture using the BD Vacutainer® system with 21-gauge needles and tubes containing EDTA (7.2 mg) as anticoagulant, and from 15 semen samples stored at -196 °C in thermos flasks containing liquid nitrogen (CAIVIO-NASNER et al. 2021b).

DNA quality was assessed qualitatively on 1% agarose gels stained with GelRed™ (Biotium, Fremont, CA, USA) and quantitatively using the NanoDrop™ 2000/2000c spectrophotometer (Thermo Scientific™, Madison, WI, USA). The samples had an average concentration of 50.1±5.4 ng/µl and an absorbance ratio A260/A280 between 1.75 and 1.85. Samples were then lyophilized in the Console FreeZone™ kit (Labconco™, Kansas City, MO, USA) and stored for genotyping.

Genotyping

The lyophilized DNA was sent to Neogen Genomics (https://genomics.neogen.com/en) where 70 animals were genotyped using the GGP Bovine 150 K chip (140,668 SNPs) and the remaining animals using the GGP Bovine 50 K chip (48,668 SNPs) (CAIVIO-NASNER et al. 2021a). Forty-three SNPs were identified in the genotyping chips, four in the *CSN1S1* gene, 14 in the *CSN2* gene, two in the *CSN1S2* gene, 12 in the *CSN3* gene and 11 in the *LGB* gene. According to the method proposed by SANCHEZ et al. (2020), these polymorphisms were used to reconstruct the protein variants **B*, **C* and **D* of *α_{S1}-CN*; **A^1*, **A^2*, **B*, **H^2* and **F* of *β-CN*; **A* and **D* of *α_{S2}-CN, *A*, **A^1*, **B*, **I* and **H* of *κ-CN* and **A*, **B*, **C*, **D*, **E*, **F* and **H* of *β-LG*.

Analysis of data

Allele and genotypic frequencies for the 43 SNPs were estimated using R software (R core team 2020). Intragenic haplotypes or protein variants (CHESSA et al. 2020) were then reconstructed for all genes, and genotypic and allele frequencies were estimated and plotted for each gene, taking into account sex, year of birth (2004 to 2018) and herd (1 to 15), using R (http://www.r-project.org/) and PLINK v1.9 (PURCELL et al. 2007).

Using PLINK v1.9 software, Hardy-Weinberg equilibrium (HWE) and F_{ST} values were estimated for each of the evaluated SNPs, using herd, sex and year as population structuring factors (only for casein genes due to their economic importance). Then, using the LD function of the Gaston package v1.5.9 (DANDINE-ROULLAND & PERDRY 2018) for R, the data were trimmed using sliding windows and a criterion r^2= 0.2, with which the global F_{ST} was estimated from the informative markers. Then, the LD values obtained among all variants were plotted using the LD.plot function of the same package. Finally, the haplotype frequencies for the caseins (*CSN1S1-CSN2-CSN1S2-CSN3*) were estimated from the reconstructed intragenic haplotypes (CHESSA et al. 2020), using the pegas package v1.0-1 (PARADIS 2010) with R software.

RESULTS

In the *CSN1S1* gene, only the rs132656458 variant was monomorphic, in the others the reference genotype and allele were most frequent (Table 1). The allele and genotypic frequencies of the SNPs present in this gene did not vary by the effect of sex, year and herd, so they were in HWE. The four variables of the *CSN1S1* gene allowed the reconstruction of the *CSN1S1*B*, **C* and **D* variants of the *α_{S1}-CN* protein, of which only the last one has a low frequency (0.006). In the BON population studied, 4 genotypes were found, of which the homozygous **B/*B* (0.646) and heterozygous **B/*C* (0.311) were the most frequent (Table 1).

Of 10 SNPs reconstructing the *CSN2*A1* variant of the *β-CN* protein, six were fixed in the population, and the frequency of the alternative allele in the polymorphic variants did not exceed 0.003 (Table 2). Thus, the frequency of the *CSN2*A^1* allele was 0.470. The SNP rs43703011, which discriminates the *CSN2*A^2* allele from the *β-CN* protein, was the most frequent (0.509) in the BON cattle studied. Although the variant did not deviate from the theoretical HWE, the frequency of this allele varied according to sex, herd and year. Thus, this allele was higher in males (Figure 1A), with frequencies ranging from 0.20 in herd 11 (n=11) to fixed in herd 13 (n=48) (Figure 1B), and frequencies above 50% were found in at least nine of the 13 years, with a slightly increasing trend over time (Figure 1C).

Table 1. Description of the variants present in the genotyping chip located in the *CSN1S1* gene and their frequencies. Allelic and genotypic frequencies of the reconstructed variants in the α_{S1}-*CN* protein found in the BON breed.

Gene/ Allele frequency	Variant	Position UCD 1.2	SNP ID	Type of variant	Genotype frequency			Allele frequency		Genotipic frequency
*CSN1S1*B*										
0.804	CSN1S1_2 (Val91Ala)	6: 85420971	rs132656458	Missense	TT	TA	AA	T	A	BB=0.646
					1.000	0.000	0.000	1.000	0.000	
	CSN1S1_3 (617G>A)	6: 85428068	rs133474041	3 UTR Variant	GG	GA	AA	G	A	
					0.729	0.253	0.019	0.855	0.145	BC= 0.311
*CSN1S1*C*										
0.193	alphaS1Casein26181 (Glu207Gly)	6: 85427427	rs43703010	Missense	AA	AG	GG	A	G	BD= 0.006
					0.651	0.311	0.037	0.807	0.193	
										CC= 0.037
*CSN1S1*D*										
0.006	CSN1S1_1 (Ala68Thr)	6: 85418630	rs433385179	Missense	GG	AG	AA	G	A	
					0.994	0.006	0.000	0.997	0.003	

Table 2. Description of the variants present in the genotyping chip located in the *CSN2* gene and their frequencies. Allelic and genotypic frequencies of the reconstructed variants in the *β-CN* protein found in the BON breed.

Gene/ Allele frequency	Variant	Position UCD 1.2	SNP ID	Type of variant	Genotype frequency			Allele frequency		Genotipic frequency
*CSN2*A^1*										
0.470	BCN_8491 (Val247Ala)	6: 85450908	rs715383373	Missense	TT	TC	CC	T	C	
					0.994	0.006	0.000	0.997	0.003	
	BCN_8463	6: 85450976	NR	NR	CC	CT	TT	C	T	
					1.000	0.000	0.00	1.000	0.000	
	CSN2_2 (His198Pro)	6: 85451055	rs454083280	Missense	TT	TG	GG	T	G	
					1.000	0.000	0.000	1.000	0.000	
	CSN2_X1471 1_8219 (His156Gln)	6: 85451180	rs43703012	Missense	CC	CA	AA	C	A	
					1.000	0.000	0.000	1.000	0.000	
	CSN2_5 (Met93Leu)	6: 85451236	NR	NR	GG	GT	TT	G	T	
					1.000	0.000	0.000	1.000	0.000	
	CSN2_6 (Leu88Ile)	6: 85451248	NR	NR	CC	CA	AA	C	A	A^1A^1= 0.203
					1.000	0.000	0.000	1.000	0.000	
	CSN2_X1471 1_8115 (Gln72Glu)	6: 85451284	NR	NR	CC	CG	GG	C	G	A^1A^2= 0.497
					1.000	0.000	0.000	1.000	0.000	A^1B= 0.029
	CSN2_8 (Glu87Lys)	6: 85452710	rs721259074	Missense	CC	CT	TT	C	T	
					0.994	0.006	0	0.997	0.003	A^1F= 0.006
	CSN2_X1471 1_6687 (Glu86Lys)	6: 85452713	rs3423226649	Missense	CC	CT	TT	C	T	A^2A^2= 0.259
					0.997	0.003	0.000	0.999	0.001	
	CSN2_X1471 1_6562 (Arg75Cys)	6: 85452837	rs799602970	Missense	CC	CT	TT	C	T	A^2H^2= 0.003
					0.994	0.006	0.000	0.997	0.003	BF= 0.003
*CSN2*A^2*										
0.509	CSN2_7 (His67Pro)	6: 85451298	rs43703011	Missense	AA	AC	CC	A	C	
					0.241	0.499	0.261	0.490	0.510	
*CSN2*B*										
0.016	CSN2_3 (Arg172Ser)	6: 85451132	rs43703013	Missense	CC	CG	GG	C	G	
					0.917	0.083	0.000	0.958	0.042	
*CSN2*F*										
0.004	CSN2_1 (Pro202Leu)	6: 87181364	rs433954503	Missense	GG	GA	AA	G	A	
					0.971	0.029	0.000	0.986	0.014	
*CSN2*H^2*										
0.001	CSN2_4 (Met143Leu)	6: 85451221	rs109299401	Missense	AA	AC	CC	A	C	
					0.994	0.006	0.000	0.997	0.003	

The genotyping chips used in this study only assessed two variants in the *CSN1S2* gene (rs441966828 and rs463985801). For both SNPs, the reference genotype and allele were the most common (Table 3), with no significant variation for sex, year of birth and herd effects, and no significant deviations from HWE. The SNP rs441966828 discriminates the *CSN1S2*A* from the α_{S2}*-CN* protein, which was the most common in BON cattle, whereas the SNP rs463985801 discriminates the *CSN1S2*D* variant, which was the least common in the study population.

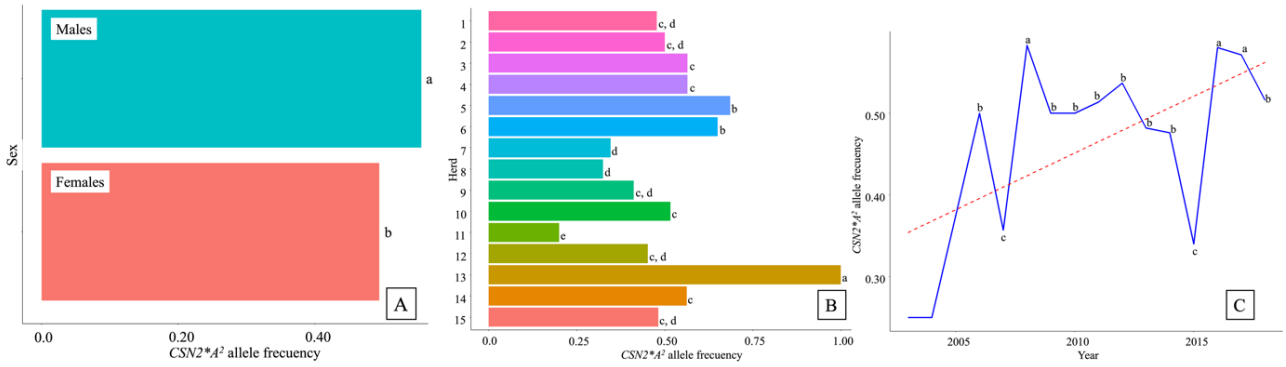

Figure 1. Frequency of the *CSN2*A²* allelic variant according to sex (A), herd (B) and year (C) in the BON breed. Different letters indicate statistically significant differences (p< 0.05) between allele frequencies by sex, herd and year.

Table 3. Description of the variants present in the genotyping chip located in the *CSN1S2* gene and their frequencies. Allelic and genotypic frequencies of the reconstructed variants in the α_{S2}*-CN* protein found in the BON breed.

Gene/ Allele frequency	Variant	Position UCD 1.2	SNP ID	Type of variant	Genotype frequency			Allele frequency		Genotipic frequency
*CSN1S2*A*										
0.997	CSN2S2_1 (Ser23Phe)	6: 85533780	rs441966828	Missense	CC	CT	TT	C	T	AA= 0.994
					1.000	0.000	0.000	1.000	0.000	
*CSN1S2*D*										AD= 0.006
0.003	CSN1S2_2 (Glu74Asp)	6: 85536434	rs463985801	Missense	GG	GT	TT	G	T	
					0.994	0.006	0.000	0.997	0.003	

Only two of the variants evaluated in the *CSN3* gene were fixed in the BON (Table 4), both associated with the *κ-CN* variant *A*, which has the highest frequency (0.679). Similarly, the variant rs43703015, which determines the *κ-CN*B* variant, had the second highest frequency (0.293). This variant was significantly higher in females (Figure 2A), in flocks 2, 5 and 7 (0.5, 0.55 and 0.51, respectively) (Figure 2B), and with a slight tendency to decrease over time (Figure 2C), since the frequency was higher than 40% between the years 2005 and 2010, whereas this frequency did not exceed 35% in the last three years studied.

On the other hand, seven of the 11 variants analyzed in the *LGB* gene were fixed in the study population (Table 5). All variants were in HWE and did not vary by sex, year or herd effects. Reconstruction of the *β-LG* variants identified 7 alleles and 10 genotypes, with the *β-LG*B* allele (0.657) and the **B/*B* genotype (0.389) being the most common, followed by the *β-LG*A* allele and the heterozygous **A/*B* genotype. Of the other alleles, only the *β-LG*H* allele had a frequency greater than 5%.

After pruning, 14 of the 32 variants identified in the *CSN1S1, CSN2, CSN1S2* and *CSN3* genes were considered informative. There were no differences in the F_{ST} values before and after pruning. The coefficient of population structure considering all SNPs was higher when herd (F_{ST}= 0.035) was considered as structuring factor compared to sex (F_{ST}= 0.003) and year (F_{ST}= 0.001). The population structure value increased when only *CSN3* gene and herd were considered (F_{ST}= 0.054) compared to *CSN2* gene and herd (F_{ST}= 0.035). Clustering by sex was higher with *CSN2* gene (F_{ST}= 0.003) compared to *CSN3* gene (F_{ST}= 0.001). Clustering by year was similar for *CSN2* and *CSN3* genes (F_{ST}= 0.001).

Table 4. Description of the variants present in the genotyping chip located in the *CSN3* gene and their frequencies. Allelic and genotypic frequencies of the reconstructed variants in the *κ-CN* protein found in the BON breed.

Gene/ Allele frequency	Variant	Position UCD 1.2	SNP ID	Type of variant	Genotype frequency			Allele frequency		Genotipic frequency
*CSN3*A*										
	CSN3_AY380228_12690 (Arg/His)	6: 85656358	NR	Missense	GG	GA	AA	G	A	
					0.997	0.003	0.000	0.999	0.001	
	CSN3_AY380228_12940 (Thr=)	6: 85656608	NR	Synonymous	TT	TC	CC	T	C	
					1.000	0.000	0.000	1.000	0.000	
	CSN3_AY380228_12950 (Arg117Ser)	6: 85656618	rs110870535	Missense	CC	CA	AA	C	A	
					0.997	0.003	0.000	0.999	0.001	
	CSN3_AY380228_12951 (Arg118His)	6: 85656619	rs716557965	Missense	GG	AG	AA	G	A	
					0.997	0.003	0.000	0.999	0.001	
0.679	CSN3_AY380228_13096	6: 85656764	NR	Intron	TT	TG	GG	T	GG	AA= 0.439
					1.000	0.000	0.000	1.000	0.000	
	CSN3_AY380228_13104_1 (Ala169Asp)	6: 85656772	rs43703016	Missense	AA	AC	CC	A	C	AB= 0.430
					0.496	0.419	0.085	0.705	0.295	AH= 0.048
	CSN3_AY380228_13124 (Ser176Gly)	6: 5656792	rs43703017	Missense	AA	AG	GG	A	G	
					0.997	0.003	0.000	0.999	0.001	AI= 0.006
	CSN3_AY380228_13165 (Ala189=)	6: 85656833	rs110014544	Synonymous	AA	AG	GG	A	G	
					0.496	0.418	0.086	0.705	0.295	A^1B= 0.003
*CSN3*A*[1]										
0.001	CSN3_AY380228_13111 (Pro171=)	6: 85656779	rs439304887	Synonymous	AA	AG	GG	A	G	
					0.991	0.009	0.000	0.996	0.004	BB= 0.077
*CSN3*B*										
0.293	CSN3_AY380228_13068 (Ile157Thr)	6: 85656736	rs43703015	Missense	TT	TC	CC	T	C	
					0.509	0.403	0.088	0.71	0.29	
*CSN3*H*										
0.024	CSN3_AY380228_13065 (Thr156Ile)	6: 85656733	rs450402006	Missense	CC	CT	TT	C	T	
					0.925	0.075	0.000	0.963	0.037	
*CSN3*I*										
0.003	CSN3_AY380228_12971 (Ser125Ala)	6: 85656639	rs43706475	Missense	TT	TG	GG	T	G	
					0.989	0.011	0.000	0.994	0.006	

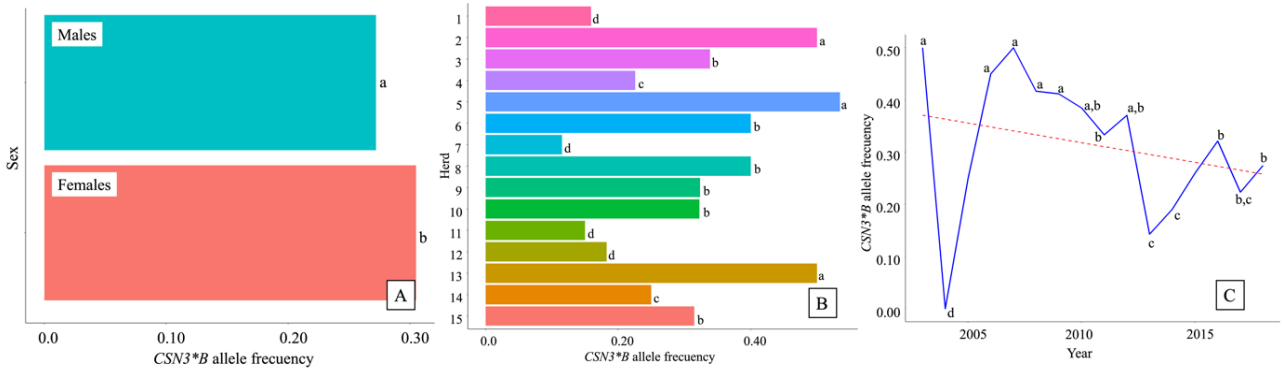

Figure 2. Frequency of the *CSN3*B* allelic variant according to sex (A), herd (B) and year (C) in the BON breed. Different letters indicate statistically significant differences ($p< 0.05$) between allele frequencies by sex, herd and year.

Table 5. Description of the variants present in the genotyping chip located in the *LGB* gene and their frequencies. Allelic and genotypic frequencies of the reconstructed variants in the *β-LG* protein found in the BON breed.

Gene/ Allele frequency	Variant	Position UCD 1.2	SNP ID	Type of variant	Genotype frequency			Allele frequency		Genotipic frequency
*LGB*B*										
0.657	LGB_X14710_3982 (Asn79=)	11: 103257948	rs110180463	Splice region	CC	TC	TT	C	T	AA= 0.054
					0.461	0.413	0.126	0.668	0.332	
	LGB_X14710_5174 (Asn104=)	11: 103259143	rs110641366	Synonymous	TT	TC	CC	T	C	AB= 0.374
					0.391	0.443	0.166	0.613	0.387	
	LGB_X14710_5962 (Pro142Leu)	11: 103259931	rs3423321021	Missense	CC	TC	TT	C	T	AC= 0.014
					1.000	0.000	0.000	1.000	0.000	
	LGB_X14710_5970 (Asp145Tyr)	11: 103259939	rs475846954	Missense	GG	GT	TT	G	T	AF= 0.003
					1.000	0.000	0.000	1.000	0.000	
	DB-2053-seq-rs209836063	11: 103280858	rs209836063	Upstream Gene	TT	TC	CC	T	C	AH= 0.003
					1.000	0.000	0.000	1.000	0.000	
*LGB*A*										BB=0.389
0.251	LGB_X14710_5263 (Ala134Val)	11: 103259232	rs109625649	Missense	CC	CT	TT	C	T	BC= 0.037
					0.433	0.415	0.151	0.641	0.359	
*LGB*C*										BD= 0.003
0.026	LGB_X14710_4027 (Ile94Met)	11: 103257993	rs3423333959	Missense	CC	CG	CC	C	G	BE= 0.023
					1.000	0.000	0.000	1.000	0.000	
*LGB*F*										BH= 0.100
0.001	LGB_X14710_3080 (Pro66Ser)	11: 103257043	rs3423333957	Missense	CC	CT	TT	C	T	
					0.997	0.003	0.000	0.999	0.001	
*LGB*H*										
0.051	LGB_X14710_4003_1 (Lys70Asn)	11: 103257969	NR	Missense	GG	GC	CC	G	C	
					1.000	0.000	0.000	1.000	0.000	
*LGB*E*										
0.011	LGB_X14710_5233 (Glu124Gly)	11: 103259202	rs3423334663	Missense	AA	AG	GG	A	G	
					1.000	0.000	0.000	1.000	0.000	
*LGB*D*										
0.001	LGB_X14710_3065 (Glu61Gln)	11: 103257028	rs211077340	Missense	GG	GC	CC	G	C	
					1.000	0.000	0.000	1.000	0.000	

Figure 3A, shows the LD values among the 14 SNPs after pruning in the genes determining milk proteins. The highest LD value was found between two variants determining the *κ-CN* variant *A* (CSN3_AY380228_13104_1 and CSN3_AY380228_13165), also the linkage between the SNP CSN3_AY380228_13068 determining the protein variant *κ-CN*B* with those mentioned above in *κ-CN*A* showed values of LD=0.99. Other high LD values were found among the SNPs classifying the *CSN1S1*B* and *CSN1S1*C* variants of the α_{S1}*-CN*, followed by the linkage between the $*A^1$ variants of the *CSN2* and *CSN3* genes (LD= 0.66). In particular, the CSN2_7 SNP, which determines the $CSN2*A^2$ variant, showed a higher linkage with the variants analyzed in the *CSN1S1* gene than with other SNPs in the *CSN2* gene and only 0.06 with the *CSN3*B* variant.

On the other hand, after data pruning, only 4 of the 11 analyzed *β-LG* variants were considered informative. The LD values of these are shown in Figure 3B. The variant LGB_X14710_3080 (*β-LG*F* allele) showed no linkage with the other SNPs. While the LGB_X14710_3982 and LGB_X14710_5174 variants, which reconstruct the *β-LG*B* allele, showed high linkage with the LGB_X14710_5263 variant, which determines the **A* allele of *β-LG*.

Finally, 28 *CSN1S1-CSN2-CSN1S2-CSN3* haplotypes were found (Table 6). Six haplotypes showed a frequency higher than 8.6%. While 11 haplotypes were in the range of 0.6% to 4.6% and 11 had frequencies equal to 0.3%. The most frequent *CSN1S1-CSN2-CSN1S2-CSN3* haplotypes were *BB-A^1A^2-AA-AA-AA* (16.6%), *BB-A^1A^2-AA-AA-AA-AB* (14.1%) and *BB-A^2A^2-AA-AA-AA* (10.1%).

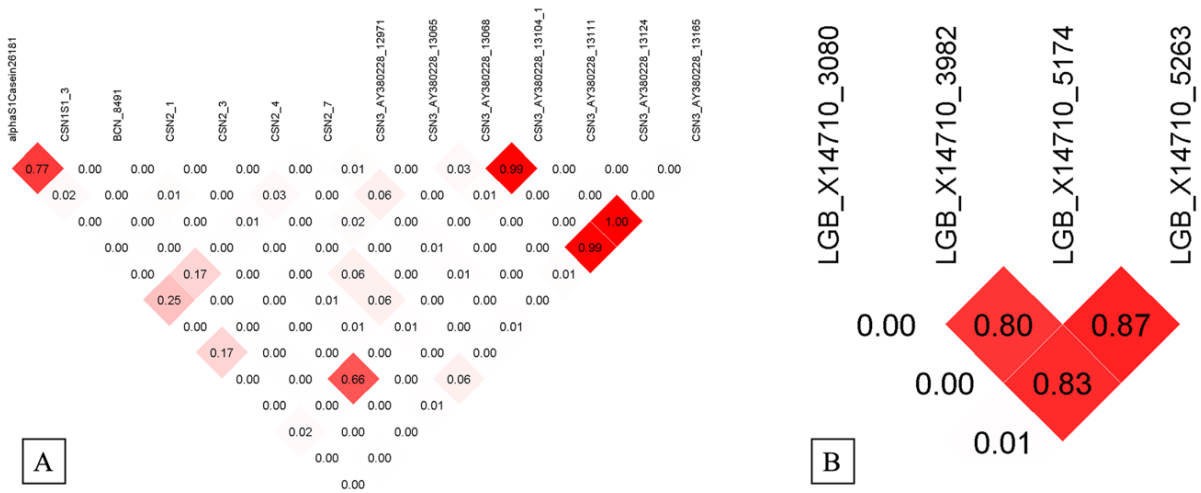

Figure 3. Linkage disequilibrium heat map of informative variants in the *κ-Casein* (A) and *β-Lactoglobulin* (B) genes in BON breed.

Table 6. Estimated haplotype frequencies from the reconstructed protein variants in the *CSN1S1-CSN2-CSN1S2-CSN3* genes found in the BON breed.

Haplotype *CSN1S1-CSN2- CSN1S2-CSN3*	Frequency	Haplotype *CSN1S1-CSN2- CSN1S2-CSN3*	Frequency
BB-A^1A^1-AA-AB	0.098	*BC-A^1A^1-AA-AB*	0.086
BB-A^1A^2-AA-A^1B	0.003	*BC-A^1A^2-AA-AA*	0.095
BB-A^1A^2-AA-AA	0.196	*BC-A^1A^2-AA-AB*	0.046
BB-A^1A^2-AA-AB	0.141	*BC-A^1B-AA-AB*	0.003
BB-A^1B-AA-AB	0.026	*BC-A^1F-AA-AB*	0.003
BB-A^1F-AA-AB	0.003	*BC-A^2A^2-AA-AA*	0.032
BB-A^2A^2-AA-AA	0.101	*BC-A^2A^2-AA-AH*	0.017
BB-A^2A^2-AA-AE	0.003	*BC-A^2A^2-AA-BB*	0.029
BB-A^2A^2-AA-AH	0.026	*BD-A^2A^2-AD-BB*	0.006
BB-A^2A^2-AA-AI	0.006	*CC-A^1A^1-AA-AB*	0.017
BB-A^2A^2-AA-BB	0.035	*CC-A^1A^2-AA-AA*	0.012
BB-A^2H^2-AA-BB	0.003	*CC-A^1A^2-AA-AB*	0.003
BB-BF-AA-AB	0.003	*CC-A^2A^2-AA-AH*	0.003
BC-A^1A^1-AA-AA	0.003	*CC-A^2A^2-AA-BB*	0.003

DISCUSSION

In the Butana breed (*Bos indicus*), 17 SNPs in the *CSN1S1* gene were found by sequencing using the **B* variant (ENSBTAT00000010119.3) as a reference (AHMED et al. 2017), the genotyping chips used in this research contain only four SNPs in this gene, which allow distinguishing the **B*, **C* and **D* variants from the nine variants *(*A, *B, *C, *D, *E, *F, *G, *H* and **I*) of the α_{S1}-*CN* protein reported (GALLINAT et al. 2013). This could be related to the breeds used in the chip design (AL KALALDEH et al. 2023) or to the ease of reconstructing variants from DNA sequencing, as all possible polymorphisms are available.

The *CSN1S1*B* allele was the most common in 13 breeds of *Bos taurus* studied, as in the present investigation, with values ranging from 0.53 in the Sarabi breed to being fixed in the Angler, German Yellow, Highland Cattle, Limpurger and Shorthorn breeds (GALLINAT et al. 2013). In Turkish Holstein, the **B* allele frequency was 0.95 (ARDICLI et al. 2018) and in Czech Simmental it was 0.90 (ČÍTEK et al. 2023). On the other hand, the *CSN1S1*C* allele was most frequent in *Bos indicus* Butana (0.66, n= 50), Gir (0.50, n= 7), Golpayegani (0.48, n= 20) and Sistani (0.88, n= 21) (GALLINAT et al. 2013). Several authors report that

animals with **B/*B* genotype have higher milk production and higher milk protein and fat concentration (ARDICLI et al. 2018, MOHAN et al. 2021, ČÍTEK et al. 2023), the frequency of this genotype in the studied BON cattle population was 0.646 (Table 1).

From the 14 SNPs present in the genotyping chips used in this work (GGP Bovine 150 K chip, n= 70 and GGP Bovine 50 K chip, n= 349), located in the *CSN2* gene, it is possible to reconstruct five of the 12 variants of the *β-CN* protein reported (CAROLI et al. 2009) *(*A1, *A2, *A3, *B, *C, *D, *E, *F, *G, *H1, *H2* and **I*). Consistent with what GALLINAT et al. (2013) reported in 13 breeds of *Bos taurus*, in Holstein and Jersey from Italy (CHESSA et al. 2020) and in the synthetic breed Karan Fries from India (*B. taurus* x *B. indicus*) (MOHAN et al. 2021), the *CSN2*A2* allele was the most frequent, with values above 0.5 in all breeds, values similar to those reported in this work. Similarly, this variant was even more common in the *Bos indicus* breeds evaluated by AHMED et al. (2017) and GALLINAT et al. (2013). Notably, the **H²* variant was not reported by the above authors. Whereas the **B* allele was the most common allele in Colombian Holstein (PADILLA-DOVAL et al. 2021).

As mentioned above, intestinal digestion of *β-CN*A¹* protein releases the bioactive peptide *β-casomorphin-7*, which is considered a risk factor for the development of several diseases (CHESSA et al. 2020, THIRUVENGADAM et al. 2021, VILLALOBOS-CORTÉS et al. 2023), in particular milk allergy and lactose intolerance (MUNTEAN et al. 2022), compared to *β-CN*A²* protein. Lactose intolerance and cow's milk allergy can be confused. The former is essentially due to lactase deficiency in the intestine (JAISWAL & WORKU 2022), whereas the latter may be a humoral, cellular, or mixed process mediated by B and T lymphocytes to epitopes resulting from intestinal digestion of milk, which bind mainly to immunoglobulin E (MUNTEAN et al. 2022). In this regard, at least five IgE-binding epitopes, three common to the *β-CN*A¹, *A²* and **B* variants and two to the *β-CN*A¹* and **A²* variants, have been reported following in vitro gastrointestinal digestion of cow's milk (LISSON et al. 2013).

Several studies have reported the possible effect of genetic variation in the *CSN2* gene on the immunomodulatory, antihypertensive, antimicrobial and antithrombotic function of bioactive peptides derived from the digestion (*in vitro*, *in vivo* and *in situ*) of the *β-CN* protein (FITZGERALD et al. 2020, JAISWAL & WORKU 2022, MUNTEAN et al. 2022), with particular focus on the *β-CN*A²* variant for the production of A2-type milk. Therefore, selection in favor of this allele has started in several countries (CHESSA et al. 2020). For example, in the Italian Holstein breed, the frequency of this allele has increased by 10% between 1990 and 2017, to the detriment of the frequency of the *β-CN*A¹* allele (CHESSA et al. 2020). The latter is similar to the results presented in this study (Figure 1C, the red dotted line indicates the trend over time), possibly due to the response of a selection process not necessarily related to the farmers' knowledge of the nutritional quality of A2-type milk, but rather to the indirect positive effects of the **A²/*A²* genotype and to the increase in milk yield, protein and fat content or technological performance for the production of curd (MOHAN et al. 2021, VIGOLO et al. 2022). However, there is a strong interest among producers in the new market for specialty milks, which has led them to use male genotyping as a criterion for sire selection, which may partly explain the observed trend. On the other hand, it must be emphasized that A2-type milk can also be obtained from the **A³, *D, *E* and **H²* variants, as all of them do not release *β-casomorphin-7* after digestion compared to the other variants (**B, *C, *F, *G* and **H¹*).

Five variants of the *α_{S2}-CN* protein *(*A, *B, *C, *D* and **E*) have been reported (GALLINAT et al. 2013). Two SNPs present on the genotyping chip allow the differentiation of the **A* and **D* variants. The former is reported to be the most common in all breeds (AHMED et al. 2017, GALLINAT et al. 2013). Conversely, the latter is only reported in Angler (0.05, n=19), German Yellow (0.07, n=23), Hinterwälder (0.02, n=21), Hungarian Grey Steppe (0.03, n=20), Limpurger (0.03, n=20) and Vorderwälder (0.10, n=24), with low frequencies as in this study (GALLINAT et al. 2013). Meanwhile, in Turkish Holstein, allele **D* frequency is reported as 0.98 (n= 168) (ARDICLI et al. 2018). On the other hand, the *CSN1S2*E* variant was found only in the Sarabi breed (0.02, n= 22) (GALLINAT et al. 2013). In the synthetic breed Karan Fries (India), animals with genotype **A/*A* showed higher total milk yield, higher protein and calcium concentration (MOHAN et al. 2021), however, ARDICLI et al. (2018) show that it is the heterozygous genotype **A/*D* that shows the best performance for these variables.

The five *κ-CN* variants that can be reconstructed from the 12 SNPs present in the genotyping chips used in this study are **A, *A1, *B, *J* and **I* of the 14 *κ-CN* variants *(*A, *A¹, *B, *B², *C, *D, *E, *F¹, *F²,*G¹, *G², *H, *I* and **J*) reported (CAROLI et al. 2009). The frequency of the **A* variant reported here was higher than that reported in nine *B. taurus* (GALLINAT et al. 2013) and four *B. indicus* breeds (AHMED et al. 2017, GALLINAT et al. 2013), in Italian Holstein (CHESSA et al. 2020), Colombian Creole Hartón del Valle

(NARANJO et al. 2007) and Karan Fries (India) (MOHAN et al. 2021), but lower than reported in Colombian Holstein for the regions of Antioquia and Nariño (SOLARTE et al. 2009, PADILLA-DOVAL et al. 2021). In contrast, the **B* allele has a high frequency of 0.9 in Italian Jersey (n=622) (CHESSA et al. 2020), values far from those reported in the BON in the present work. On the other hand, in *B. indicus* breeds, the **H* variant is more frequent ($f_{(*H)} > 0.5$) compared to what was found here (**H*= 0.024).

Regarding the effect of the *CSN3* gene on traits related to milk production and milk quality, a meta-analysis (MAHMOUDI et al. 2020) including 48 articles with genotypes of 30471 animals for the **A* and **B* variants, in specialized and local breeds, taking into consideration four different inheritance models, reports a higher daily milk yield in animals with the **B/*B* and **A/*B* genotypes ($p< 0.05$) in the recessive and additive models. In the dominant and additive models, the **B/*B* genotype improves milk fat yield. In a dominant and recessive model, the **A/*A* genotype decreases milk protein percentage (MAHMOUDI et al. 2020). These results are confirmed by MOHAN et al. (2021) who report that the **B/*B* genotype is associated with higher milk yield and higher protein, calcium and phosphorus content. The same genotype is also reported to produce milk with a better fatty acid profile (HEWA NADUGALA et al. 2022). In addition, the **B/*B* genotype has been associated with shorter coagulation time, higher stability of micelle structure, higher gel strength and higher cheese yield (POULSEN et al. 2017, ZEPEDA-BATISTA et al. 2017, KRUCHININ et al. 2023).

However, the possible beneficial health effects of the bioactive peptides described for *κ-CN* include antithrombotic activity, reduction of inflammatory processes and oxidative stress, and improvement of the immune response (VARGAS-BELLO-PÉREZ et al. 2019). These effects are certainly related in some way to the polymorphism in the *CSN3* gene. For example, after *in vitro* digestion of *κ-CN*, two common IgE-binding epitopes are reported between the **A* and **E* alleles and only one in the **B* variant (LISSON et al. 2013), which can be partly explained by the low homology (58%) between human and bovine *κ-CN* (JAISWAL & WORKU 2022), which could imply a lower allergenicity of the *κ-CN*B* variant.

The above reports justify the targeting of breeding programs towards the selection of the *CSN3*B* allele. This is demonstrated by CHESSA et al. (2020) in Italian Holstein, where the frequency of this allele increased from 0.186 before 1990 to 0.355 in 2017. On the contrary, in the BON breed there is a slight decrease over time (Figure 2C, red dotted line) of this allele, and although the frequency of the homozygous **B/*B* genotype was low (0.077), the allele frequency has an important magnitude (Table 4), with high variation in its frequency between herds (Figure 2B), which would facilitate and justify a selection scheme in favor of this allele.

The **B* allele of the *β-LG* protein (Ensembl ID: ENSBTAT00000019538) is used as a reference to reconstruct the 11 reported variants (**A, *B, *C, *D, *E, *F, *G, *W, *I, *J* and **W*) in this protein (CAROLI et al. 2009). The 11 SNPs included in the genotyping chips allow the identification of the **A, *B, *C, *D, *E, *F* and **H* variants. In a previous study in BON cattle, the **A* variant was found to be the most frequent (0.51) (ROSERO-ALPALA et al. 2011), a value that contrasts with this study, but which can be explained by the limited sample size used (n= 30) and the origin of the samples (a single farm in one department vs. 15 herds in six departments). The **B* allele was the most frequent allele in several cattle populations (AHMED et al. 2017), as well as in the Holstein and Simmental populations of the Czech Republic (ČÍTEK et al. 2023), which is consistent with the results of this work. In contrast, the **A* variant is more common in the Gyr and Jersey breeds (BARBOSA et al. 2019, GAI et al. 2021).

Several authors have reported that the *β-LG*B/*B* genotype correlates with lower milk yield, but with higher protein, fat and casein content, additionally with shorter rennet formation time and lower somatic cell score (BARBOSA et al. 2019, GAI et al. 2021, ČÍTEK et al. 2023), this genotype was the most frequent in the BON population studied (Table 5). Different bioactive peptides with antihypertensive, antimicrobial, antilipemic, antioxidant and immunomodulatory activity have been reported from in vitro and in silico digestion of *β-LG* (DULLIUS et al. 2018, KIM et al. 2019, ČÍTEK et al. 2023) but reports on their bioactivity according to genetic variation are null.

There is extensive knowledge about the unique effect of casein variants on the properties and functionality of milk. However, some authors suggest that due to their proximity, it is possible that they are inherited as a linkage group rather than independently (VILLALOBOS-CORTÉS et al. 2023) and therefore the joint effect, i.e. the haplotypes, needs to be investigated (HEWA NADUGALA et al. 2022). The results of this investigation show that linkage between the *CSN1S1-CSN2-CSN1S2-CSN3* genes is low (Figure 3A), except for the **A1* variants (LOD= 0.66) of the *CSN2* (BCN_8491) and *CSN3* (CSN3_AY380228_13111) genes, which may partly explain the 28 haplotypes found and the low frequencies (<5%) of most of them (22 haplotypes). In contrast, linkage between variants of the same gene was logically higher, such as between

variants of the *CSN3* or *LGB* genes (Figure 3B).

AHMED et al. (2017) found 84 haplotypes for the *CSN1S1-CSN2-CSN1S2-CSN3* genes. However, unlike this study, AHMED et al. 2017 constructed the haplotypes from the protein variants rather than the genotypes, which would explain the higher number of haplotypes, nevertheless the haplotypes *C-A²-A-A-A* (0.156), *B-A²-A-A-A* (0.119) and *B-A²-A-A-A-A* (0.085) were the most common in their report. In another study using only *β-casein* and *κ-casein*, the most frequent haplotypes in the Italian Holstein breed were *A²-A* (0.474) and *A¹-A* (0.218), whereas in the Italian Jersey breed they were *A²-B* (0.646) and *B-B* (0.171) (CHESSA et al. 2020).

Finally, among some effects of haplotypic variants on milk production, quality and technological performance, the haplotype *BB-A²A²-BB* of the *CSN1S1-CSN2-CSN3* genes (with a frequency of 3.5% in BON) has been associated with higher milk protein content, better coagulation, better ability to retain fat and protein in milk (PERNA et al. 2016a,b). The *BB-A²A²-AA* combination of *CSN1S1-CSN2-CSN3* genes correlated with low or no rennet-induced coagulation, this haplotype was found with a frequency of 10.1% in BON, however, this same haplotype was associated with better but acid-induced coagulation characteristics (KETTO et al. 2017), which may be of interest for cheese producers with milk of BON origin. Additionally, the haplotypes *CC-A²A²-BB* (0.3% in BON), *BC- A2A2-BB* (0.29% in BON) and *BB-A¹A²-AA* (19.6% in BON) with a cumulative frequency of 20.19% in BON, are associated with good rennet formation, coagulation time and cheese yield (KETTO et al. 2017).

The *BB-BB* haplotype of the *CSN2-CSN3* genes (only found as heterozygous in BON with a cumulative frequency of 0.6%) with higher cheese yield (CHESSA et al. 2020). The *B-A¹-B* variant of *α_{S1}-CN*, *β-CN* and *κ-CN* proteins was associated with higher milk yield and protein content, found in heterozygous state in BON with cumulative frequency close to 40% (HEWA NADUGALA et al. 2022). Haplotype *B-A¹-A* proteins *α_{S1}-CN, β-CN and κ-CN* were associated with higher milk fat content (HEWA NADUGALA et al. 2022). In the Italian Reggiana breed the *C-A2-B* haplotype was correlated with higher fat and protein content (CAROLI et al. 2009, HEWA NADUGALA et al. 2022), with a frequency of 0.3% in BON. Finally, the possible effects on human health in terms of allergenicity and intolerance to milk and milk products of haplotypic variants or composite genotypes have not been studied, probably due to the existence of synergistic combinations between the variants and the complexity of analysing the results.

CONCLUSION

The BON cattle population showed a high variation in the SNPs evaluated, as 62.79% of the SNPs were polymorphic. The SNPs present in the genotyping chips allow the reconstruction of the most important *casein* and *β-lactoglobulin* variants from the point of view of their relationship with milk production, quality and technological performance. The high frequency of the protein variants of interest (*CSN1S1*B, β-CN*A2, κ-CN*B* and *β-LG*B*) in BON cattle suggests a potential in terms of production, quality, yield and functionality of the milk produced by this breed, without ignoring the non-genetic factors that can modify it. The high frequency in both sexes and the variation between farms of the *β-CN*A²* and *κ-CN*B* variants could be used as a basis for a future genetic improvement program in the breed, with the aim of producing quality milk with a function in human health. These results indicate that there is no linkage in the casein genes, but there is linkage between some intragenic variants useful for indirect genotyping. It will be necessary to extend this study to the complete sequencing (including UTRs) of the *CSN1S1, CSN2, CSN1S2, CSN1S2, CSN3, LAA* and *LGB* genes in order to detect new SNPs that characterize the breed or have a potential effect on protein quality, as well as to replicate it in the other Criollo breeds present in Colombia. Also, to study not only the individual effect of the variants, but also the effect of the *caseins*/*α-LA/β-LG* as a whole, to understand the possible synergistic or opposite effect on the production, quality, yield and functionality of cow's milk.

ACKNOWLEDGMENTS

The authors would like to thank the livestock professionals who facilitated the collection of information.

REFERENCES

AHMED AS et al. 2017. Milk protein polymorphisms and casein haplotypes in butana cattle. Journal of Applied Genetics 58 :261-271.

AL KALALDEH M et al. 2023. Detection of genomic regions that differentiate Bos indicus from Bos taurus ancestral breeds for milk yield in Indian crossbred cows. Frontiers in Genetics 13: 1082802.

AMAYA A et al. 2022. Selection indexes to optimise genetic and economic progress in colombian Blanco Orejinegro Cattle. Livestock Science 263: 105015.

ARDICLI S et al. 2018. Effect of STAT1, OLR1, CSN1S1, CSN1S2, and DGAT1 genes on milk yield and composition traits of Holstein breed. Revista Brasilera de Zootecnia 47: e20170247.

ARRIETA L et al. 2021. Enfermedad podal en bovinos: prevalencia y asociación con algunas variables. Revista de la Facultad de Medicina Veterinaria y Zootecnia 68: 66-74.

BARBOSA S et al. 2019. Genetic association of variations in the Kappa-Casein and β-Lactoglobulin genes with milk traits in Girolando cattle. Revista Brasileira de Saúde e Produção Animal 20: e0312019.

CAIVIO-NASNER S et al. 2021a. Diversity analysis, runs of homozygosity and genomic inbreeding reveal recent selection in Blanco Orejinegro cattle. Journal of Animal Breeding and Genetics 138: 613-627.

CAIVIO-NASNER S et al. 2021b. Frequency of genotypic markers for genetic disorders, colour, polledness, and major genes in Blanco Orejinegro cattle. Tropical Animal Health and Production 53: 546.

CAIVIO-NASNER S et al. 2021c. Genetic parameters and trends for reproductive traits in Blanco Orejinegro cattle from Colombia. Semina Ciências Agrárrias 42: 2523-2538.

CAROLI AM et al. 2009. Invited Review: Milk protein polymorphisms in cattle: Effect on animal breeding and human nutrition. Journal of Dairy Science 92: 5335-5352.

CHESSA S et al. 2020. The effect of selection on casein genetic polymorphisms and haplotypes in Italian Holstein cattle. Italian Journal of Animal Science 19:833-839.

ČÍTEK J et al. 2023. CSN1S1 and LALBA polymorphisms and other factors influencing yield, composition, somatic cell score, and technological properties of cow's milk. Animals 13: 2079.

DANDINE-ROULLAND C & PERDRY H. 2018. Genome-Wide data manipulation, association analysis and heritability estimates in R with gaston 1.5. Human Heredity 83: 6.

DULLIUS A et al. 2018. Whey Protein hydrolysates as a source of bioactive peptides for functional foods – biotechnological facilitation of industrial scale-up. Journal of Functional Foods 42: 58-74.

FARRELL HM et al. 2004. Nomenclature of the proteins of cows' milk--sixth revision. Journal of Dairy Science 87: 1641-1674.

FITZGERALD R et al. 2020. Application of in silico approaches for the generation of milk protein-derived bioactive peptides. Journal of Functional Foods 64: 103636.

GAI N et al. 2021. Effect of Protein genotypes on physicochemical properties and protein functionality of bovine milk: A review. Foods 10: 2409.

GALLINAT JL et al. 2013. DNA-based identification of novel bovine casein gene variants. Journal of Dairy Science 96: 699-709.

LÓPEZ-HERRERA et al. 2001. Mecanismos moleculares de resistencia a las enfermedades vesiculares virales del ganado criollo colombiano blanco orejinegro (BON). Iatreia 14: 280.

HEWA NADUGALA B et al. 2022. The effect of casein genetic variants, glycosylation and phosphorylation on bovine milk protein structure, technological properties, nutrition and product manufacture. International Dairy Journal 133: 105440.

JAISWAL L & WORKU M. 2022. Recent perspective on cow's milk allergy and dairy nutrition. Critical Reviews in Food Science and Nutrition 62: 7503-7517.

KETTO I et al. 2017. Effects of milk protein polymorphism and composition, casein micelle size and salt distribution on the milk coagulation properties in norwegian red cattle. International Dairy Journal 70: 55-64.

KIM Y et al. 2019. Alpha-Casein and Beta-Lactoglobulin from cow milk exhibit antioxidant activity: A plausible link to antiaging effects. Journal of Food Science 84: 3083-3090.

KOLENDA M & SITKOWSKA B. 2021. The Polymorphism in various milk protein genes in Polish Holstein-Friesian dairy cattle. Animals 11: 389.

KRUCHININ A et al. 2023. Effect of CSN3 gene polymorphism on the formation of milk gels induced by physical, chemical, and biotechnological factors. Foods 12: 1767.

KYSELOVÁ J et al. 2019. Physiochemical characteristics and fermentation ability of milk from Czech Fleckvieh cows are related to genetic polymorphisms of β-casein, κ-casein, and β-lactoglobulin. Asian-Australas Journal of Animal Sciences 32: 14-22.

LEÓN C et al. 2019. Genomic association study for adaptability traits in four Colombian cattle breeds. Genetics and Molecular Research 18: GMR18373.

LISSON M et al. 2013. Genetic variants of bovine β- and κ-Casein result in different immunoglobulin E-binding epitopes after in vitro gastrointestinal digestion. Journal of Dairy Science 96: 5532-5543.

LONDOÑO-GIL M et al. 2021. Genome-wide association study for growth traits in Blanco Orejinero (BON) cattle from Colombia. Livestock Science 243: 104366.

MACEDO L et al. 2020. Genomic Analysis of milk protein fractions in Brown Swiss cattle. Animals 10: 336.

MAHMOUDI P et al. 2020. A Meta-Analysis on association between CSN3 gene variants and milk yield and composition in cattle. Animal Genetics 51: 369-381.

MARTÍNEZ A et al. 2012a. Genetic footprints of Iberian cattle in America 500 years after the arrival of Columbus PLOS ONE 7: e49066.

MARTÍNEZ R et al. 2012b. Desempeño de toretes de las razas criollas Blanco Orejinegro y Romosinuano en prueba de crecimiento en pastoreo. Revista Colombiana de Ciencias Pecuarias 25: 36-45.
MARTÍNEZ R et al. 2012c. Eficiencia productiva de la raza BON en el trópico colombiano. 1.ed. Corporación colombiana de investigación agropecuaria - AGROSAVIA. Bogotá. Colombia
MOHAN G et al. 2021. Casein (CSN) gene variants and parity affect the milk protein traits in crossbred (*Bos taurus* x *Bos indicus*) cows in sub-tropical climate. Tropical Animal Health and Production 53: 289.
MUKESH M et al. 2022. Demographic pattern of A1/A2 beta casein variants indicates conservation of A2 type haplotype across native cattle breeds (*Bos indicus*) of India. 3 Biotech 12: 167.
MUNTEAN I et al. 2022. Predictive factors for oral immune modulation in cow milk allergy. Nutrients 14: 494.
NARANJO J et al. 2007. Detección de variantes alélicas de la kappa-caseína en bovinos Hartón del Valle. Acta Agronómica 56: 43-47.
PADILLA-DOVAL J et al. 2021. Análisis genético de cinco polimorfismos de nucleótido simple de caseínas lácteas obtenidos con chips genómicos en ganado Holstein de Antioquia, Colombia. Revista de la Facultad de Medicina Veterinaria y Zootecnia 68: 137-149.
PARADIS E. 2010. pegas: an R package for population genetics with an integrated–modular approach. Bioinformatics 26: 419-420.
PERNA A et al. 2016a. The Influence of casein haplotype on quality, coagulation, and yield traits of milk from Italian Holstein cows. Journal of Dairy Science 99: 3288-3294.
PERNA A et al. 2016b. The Influence of casein haplotype on morphometric characteristics of fat globules and fatty acid composition of milk in Italian Holstein cows. Journal of Dairy Science 99: 2512-2519.
RIVEROS D et al. 2022. Relación entre la hormona antimülleriana y la reserva ovárica en vacas donadoras Brahman y Blanco Orejinegro. Revista de Investigaciones Veterinarias del Perú 33: e21000.
POULSEN N et al. 2017. Novel genetic variation associated to CSN3 strongly affects rennet-induced milk coagulation. International Dairy Journal 71: 122-130.
PURCELL S et al. 2007. PLINK: A tool set for whole-genome association and population-based linkage analyses. The American Journal of Human Genetics 81: 559-575.
ROCHA J et al. 2019. Tick burden in *Bos taurus* cattle and its relationship with heat stress in three agroecological zones in the tropics of Colombia. Parasites and Vectors 12: 73.
ROSERO-ALPALA J et al. 2011. Polimorfismo genético de beta-lactoglobulina y alphalactoalbúmina en el ganado criollo colombiano, mediante PCR-SSCP. Acta Agronómica 60: 339-346.
SANCHEZ M et al. 2020. Frequencies of milk protein variants and haplotypes estimated from genotypes of more than 1 million bulls and cows of 12 French cattle breeds. Journal of Dairy Science 103: 9124-9141.
SOLARTE E et al. 2009. Allelic frequencies of Kappa casein gene in Holstein breed in Nariño - Colombia. Livestock Research for Rural Development 21: 3.
THIRUVENGADAM M et al. 2021. β-Casomorphin: A complete health perspective. Food Chemistry 337: 127765.
VANVANHOSSOU S et al. 2021. First DNA sequencing in Beninese indigenous cattle breeds captures new milk protein variants. Genes 12: 1702.
VARGAS-BELLO-PÉREZ E et al. 2019. Bioactive peptides from milk: animal determinants and their implications in human health. Journal of Dairy Research 86: 136-144.
VIGOLO V et al. 2022. β-Casein Variants differently affect bulk milk mineral content, protein composition, and technological traits. International Dairy Journal 124: 105221.
VILLALOBOS-CORTÉS A et al. 2023. Characterization of casein variants in the Guaymi and Guabala breeds through a low-density chip arrangement. Journal of Applied Animal Research 51: 69-73.
ZEPEDA-BATISTA J et al. 2017. Potential influence of κ-casein and β-lactoglobulin genes in genetic association studies of milk quality traits. Asian-Australasian Journal of Animal Sciences 30: 1684-1688.