JAST (Journal of Animal Science and Technology) TITLE PAGE

Upload this completed form to website with submission

ARTICLE INFORMATION	Fill in information in each box below
Article Type	Research article
Article Title (within 20 words without abbreviations)	*CSN1S1* and *CSN3* genetic variants affect milk quality traits in the buffalo population in Serbia
Running Title (within 10 words)	*CSN1S1* and *CSN3* genetic variants affect buffalo milk quality traits
Author	Stefan Stepic1, Jovana Kuveljic2, Ivan Zivotic2, Predrag Perisic1, Vladan Bogdanovic1, Aleksandar Ignjatovic1, Maja Zivkovic2
Affiliation	1 Institute of Animal Science, Faculty of Agriculture, University of Belgrade, Belgrade, Serbia 2 Laboratory for Radiobiology and Molecular Genetics, "Vinca" Institute of Nuclear Sciences, National Institute of the Republic of Serbia, University of Belgrade, Belgrade, Serbia
ORCID (for more information, please visit https://orcid.org)	Stefan Stepic, ORCID ID: 0000-0003-2277-5213, Jovana Kuveljic, ORCID ID: 0000-0001-6109-9460, Ivan Zivotic, ORCID ID: 0000-0002-5073-5537, Predrag Perisic, ORCID ID: 0000-0003-3067-4658, Vladan Bogdanovic, ORCID ID: 0000-0002-8345-1258, Aleksandar Ignjatovic, ORCID ID: 0000-0001-6713-4415, Maja Zivkovic, ORCID ID: 0000-0002-0447-6626,
Competing interests	No potential conflict of interest relevant to this article was reported.
Funding sources State funding sources (grants, funding sources, equipment, and supplies). Include name and number of grant if available.	This study was funded by the Ministry of Science, Technological Development and Innovation of the Republic of Serbia (Agreement no. 451-03-66/2024-03/200017 and 451-03-65/2024-03/200116).
Acknowledgements	Not applicable.
Availability of data and material	Upon reasonable request, the datasets of this study can be available from the corresponding author.
Authors' contributions Please specify the authors' role using this form.	Conceptualization: Zivkovic M, Perisic P. Data curation: Stepic S, Ignjatovic A., Bogdanovic V, Perisic P. Formal analysis: Kuveljic J, Zivotic I, Zivkovic M. Methodology: Software: Validation: Investigation: Kuveljic J, Stepic S, Zivotic I, Ignjatovic A. Writing - Kuveljic J, Stepic S, Writing - review & editing: Zivkovic M, Perisic P, Bogdanovic V. Kuveljic J, Stepic S, Zivotic I, Ignjatovic A.
Ethics approval and consent to participate	The Ministry of Agriculture, Forestry and Water Management of the Republic of Serbia issued an ethical permit for the use of animals for research implementation, number 323-07-10974/2022-05.

CORRESPONDING AUTHOR CONTACT INFORMATION

For the corresponding author (responsible for correspondence, proofreading, and reprints)	Fill in information in each box below
First name, middle initial, last name	Jovana P. Kuveljic
Email address – this is where your proofs will be sent	jovana@vin.bg.ac.rs
Secondary Email address	
Address	Mike Petrovica Alasa 12-14, P.O. Box 522, 11001 Belgrade, Serbia
Cell phone number	+38163330470
Office phone number	+381113408566
Fax number	

Abstract

The Mediterranean water buffalo breed is the most common in Europe. Caseins are major milk proteins whose gene variants can affect milk yield, composition, and processing characteristics. The most prevalent type of milk protein, αS1 casein, has been associated with milk quality traits. κ casein has been associated with traits crucial for cheese manufacturing. Small-scale households in Serbia raise buffaloes primarily for their use to make milk and butter. This study aimed to assess the association of *CSN1S1* and *CSN3* genetic variants, solely and in haplotype, with milk quality traits in the water buffalo population in Serbia. The study involved 130 water buffaloes from Serbia. *CSN1S1* and *CSN3* genotypes were determined using sequencing analysis on an ABI PRISM 3130 Genetic Analyzer. A subgroup of 42 animals was analyzed for the composition of raw buffalo milk. Statistical analyses were performed using Statistica 8 software and Thesias software for haplotype analysis. We found that *CSN1S1* 472G>C was associated with higher protein and casein levels in milk. Haplotype analysis of variants *CSN1S1* 472G>C and *CSN3* 467C>T showed that haplotype 1C3C had significantly higher casein levels (*p*=0.00002) and protein levels (*p*=0.0004) in comparison with the reference haplotype, 1G3C. Our results showed that genetic variants *CSN1S1* 472G>C and *CSN3* 467C>T in haplotype significantly impact casein levels in buffalo milk and suggest that their haplotype analysis provides greater significance in association with milk casein level than individual analysis.

Keywords (3 to 6): Water buffalo, *CSN1S1*, *CSN3*, genetic variant, haplotype, milk quality traits

Introduction

The water buffalo (*Bubalus bubalis*) population in Europe descends from Asian wild buffalo, and the breed is known as the Mediterranean water buffalo. The largest buffalo population in Europe is in Italy, followed by Romania and Bulgaria, and buffalo rearing is on the rise in Southern Europe (FAOSTAT, Food and Agriculture Organization of the United Nations, https://www.fao.org/faostat/en/#home). Buffalo milk contains more proteins, fat, and minerals than cow milk [1,2]. Increasing awareness of buffalo milk's health benefits consequently has led to the rise in buffalo breeding worldwide and the manufacturing of products from buffalo milk.

Caseins are major milk proteins, accounting for almost 80% of proteins in buffalo milk. There are four types of caseins: αS1, αS2, β, and κ casein, regulated by genes in the same cluster on chromosome 7: *CSN1S1*, *CSN1S2*, *CSN2*, and *CSN3*, respectively. Casein gene variants can affect milk yield, composition, and processing characteristics [3]. As αS1 casein is the most abundant of all milk proteins, it was the first to be investigated at the protein and gene levels, and associated with milk quality traits. *CSN1S1* variant 628C>T, located in exon 17, leads to amino acid change, Ser178 Leu, and this gives αS1 casein variants A and B, respectively. Variant B has shown a significant effect on milk protein percentage [4]. In the buffalo population of Romania, a variation of the αS1 casein B variant has been discovered. The 472G>C substitution in the B variant led to the skipping of exon 6 and the synthesis of a protein lacking eight amino acids [5]. However, the effect of this variant on milk quality traits was not investigated. κ casein has been associated with the stabilization and size of casein micelles [6], traits important for cheese production. It has been established that *CSN3* exon 4 codes for most of the mature κ casein. The most significant variant, 467C>T, which leads to an amino acid change, has been associated with a higher percentage of caseins, proteins and fat in cattle milk [7].

In Serbia, buffaloes are reared in small households mainly for local use of milk and butter. Even though the importance of buffalo conservation has been recognized and an *in situ* program was established nearly twenty years ago [8], there is still a need for selection and refinement. The genetic background of the buffalo population in Serbia has never been investigated, and this is necessary prior to estimating the potential for selection in the Serbian buffalo population. As milk and dairy products are primarily in use, this study aimed to assess the association of casein *CSN1S1* and *CSN3* gene variants, solely and in haplotype, with milk quality traits in the water buffalo population in Serbia.

Materials and Methods

Serbian buffalo population

The study involved 130 animals raised in the Raška region (Serbia) within the Pešter area, specifically in Novi Pazar and Tutin municipalities. In the Pešter area, domestic buffalo are extensively reared on pastures, representing an *in situ* form of conservation. They are primarily raised for milk and butter production, with less emphasis on meat production. During winter, the animals are kept in well-insulated barns as domestic buffalo are not tolerant to cold temperatures. They are taken to pasture in spring and remain there throughout the growing season. Animals were carefully chosen according to the Central Herd Register (Republic of Serbia), so there was no direct kinship between individuals.

The Ministry of Agriculture, Forestry and Water Management of the Republic of Serbia issued an ethical permit for the use of animals for research implementation, number 323-07-10974/2022-05. All rules related to animal welfare were followed.

Genetic analysis

Blood samples for genetic analyses were collected from the jugular vein (*v. jugularis*) into 3 ml EDTA-coated vacuum tubes. The samples were labeled and stored at +4°C until transportation to the laboratory, where they were then preserved at -20°C. DNA was extracted from blood samples of 130 animals using the phenol-chloroform extraction method. *CSN1S1* (Gene ID: 102396531) 472G>C and *CSN3* (Gene ID: 102395364) 445G>A, 467C>T, 471C>T and 516A>C genotypes were determined using sequencing analysis on an ABI PRISM 3130 Genetic Analyzer (Applied Biosystems, USA). BigDye™ Terminator v3.1 Cycle Sequencing Kit was applied to terminate the PCR reaction and for sequencing purposes following the manufacturer's protocol (Applied Biosystems, USA). Primers I5CZS1-F: 5' -ACT TAG CAA GGA GAT AAT GCA AGA A-3' and E7BCZS1-R: 5' - CTC AGT TGA TTC ACT CCC AAC ATC-3' were used for amplifying the genomic region from intron 5 to exon 7 of *CSN1S1* [5]. Primers 5' -CGC TGT GAG AAA GAG GAA AGA TTC-3' and 5' -AGA TTC AAG GAG TAT ACC AAT TGT TG-3' were used for amplifying exon 4 of *CSN3* [9]. Forward primers were utilized for the sequencing analysis. Detection of genotypes was conducted using Sequencing Analysis Software V4.0 (Applied Biosystems, USA). Genotyping succeeded in 128 samples for the *CSN1S1* variant and 125 for the *CSN3* variants. A quarter of the samples were genotyped for *CSN1S1* 472G>C by the PCR-RFLP method. The PCR product was digested with *TaaI* restriction enzyme, and the products of digestion were visualized on an 8 % polyacrylamide gel stained with silver nitrate. Genotypes were consistent with the results of sequencing.

Milk composition analysis

Milk samples were collected from a subgroup of 42 animals in the same lactation stages three times at equal intervals. Immediately after sampling, the milk was cooled, and no preservatives were added to the samples. Milk was sampled during winter when the animals were housed in barns. The composition of raw buffalo milk samples was determined by the following methods: titratable acidity according to the Soxhlet-Henkel method [10], total solids by the standard drying method at 102±2°C [11]; fat contents according to the Gerber method [12]; nitrogen and casein nitrogen contents by the Kjeldahl method [13], while the protein and casein contents were calculated as the nitrogen contents and casein nitrogen contents multiplied by 6.38.

Statistical analysis

Deviations from Hardy–Weinberg equilibrium were assessed by the chi-square (χ^2) test. Values of continuous variables are presented as mean ± standard deviation (SD). Casein correlations with other investigated milk quality traits were analyzed with the Spearman test and presented as Spearman R-value. T-test was used to compare the values of continuous variables with a normal distribution within two groups. The nonparametric Mann–Whitney U test was used to compare the values of continuous variables with a skewed distribution within two groups. The strength of the association between *CSN1S1* and *CSN3* genotypes and milk components was assessed using a multiple linear regression model and presented as β coefficient. For the regression analysis we have applied linear model as follows:

$$Y=\beta_0+\beta_1X_1+\beta_2X_2+\epsilon$$

where:

1. Y is the dependent variable (e.g., casein level in milk).
2. β_0 is the intercept of the regression line.
3. β_1 is the coefficient representing the effect of the first genetic variant (X_1).
4. β_2 is the coefficient representing the effect of the second genetic variant (X_2).
5. ϵ is the error term, which accounts for the variation in Y not explained by the linear relationship with X_1 and X_2.

The proportion of the variance in the casein and protein level that can be explained by the *CSN1S1* gene variant is presented as the R^2 from the regression model. Homogeneity of Variance was confirmed by the Levene's Test. The *p*-value < 0.05 was considered statistically significant. Statistical analyses were performed using the Statistica Version 8 software package [14].

Haplotype analysis was performed by the publicly available Thesias software (www.genecanvas.org). Thesias applies the stochastic-EM (Expectation–Maximization) algorithm to estimate haplotype frequencies and

their associated effects on the phenotype of interest [15,16]. For haplotype-phenotype association it uses the likelihood ratio test. Thesias assumes the additivity of the haplotype effects. The software then fits the following linear model to estimate the effects of haplotypes on the continuous trait. The Thesisas software was also used to estimate the LD parameters within the studied groups and to calculate the percentage of variance explained by the haplotypes. According to the Thesias software manual, *CSN1S1* and *CSN3* haplotypes' effect on milk quality trait levels are given as mean value for one dose of each haplotype along with 95% CI compared to the reference, most frequent haplotype. A *p*-value <0.008 was considered statistically significant because we applied Bonferroni correction for six milk quality traits being assessed. Statistical power of the study was calculated by the on-line Post-hoc Power Calculator (https://clincalc.com/stats/Power.aspx) [17].

Results

Variants, *CSN1S1* 472G>C (**Supplementary Figure 1**) and *CSN3* 467C>T were in Hardy–Weinberg equilibrium in the Serbian Buffalo population. The genotype distribution and allele frequencies are presented in **Table 1**. Three variants were additionally investigated, two of them, *CSN3* 445G>A and *CSN3* 516A>C, were not polymorphic in the investigated Serbian Buffalo population, meaning that only wild-type alleles G for *CSN3* 445G>A and allele A for *CSN3* 516A>C, were present. And, the third variant *CSN3* 471C>T was in complete linkage disequilibrium (LD) with variant *CSN3* 467C>T (**Supplementary Figure 2**).

Table 2 presents the composition of milk for the studied subgroup of 42 buffaloes and correlations between milk casein percentage and the other five milk quality traits. Caseins were significantly positively correlated with protein, acidity SH°, total solids, and solids-not-fat, and were not significantly correlated with milk fat percentage.

Figure 1 shows individual associations of variants *CSN1S1* 472G>C and *CSN3* 467C>T with protein and casein levels in milk. We observed that minor allele C (*CSN1S1* 472G>C), by the dominant model of inheritance, GG vs. GC+CC, was associated with higher protein levels in milk, (p=0.0008, t-test) and higher casein levels, (p=0.00007, t-test). The study power for this association was 90% for α=0.05. Proportion of variance in caseins explained by *CSN1S* 472G>C was 33%, and for proteins it was 25%, by the regression model. Variant *CSN3* 467C>T was not associated with protein and casein levels in milk by the dominant model of inheritance, CC vs. CT+TT. However, by the multiple linear regression model, we found significant effects of *CSN1S1* 472G>C (β=0.62, p=0.000013) and *CSN3* 467C>T (β=0.32, p=0.012) on milk casein levels. Substitution effect for variant *CSN1S1* 472G>C, individually, is β=0.58 for caseins and β=0.50 for proteins.

As *CSN1S1* and *CSN3* are located on chromosome 7 but are not in a strong LD (r^2=0.17, D'=-1), there

was a significant rationale to analyze them as haplotype. The Thesias software set the 1G3C haplotype as a reference because it was the most frequent haplotype in the investigated buffalo population. Haplotype frequencies of *CSN1S1* 472G>C and *CSN3* 467C>T and their effects on milk quality traits are presented in **Table 3**. Compared with the reference haplotype 1G3C (mean casein level [95% CI] = 3.34 [3.18 - 3.50], haplotypes 1G3T and 1C3C had significantly higher casein levels in milk (1G3T - mean casein level [95% CI] = 3.68 [3.48 - 3.88], p=0.02 and 1C3C - mean casein level [95% CI] = 3.90 [3.72 - 4.06], p=0.00002). In addition, haplotype 1C3C had significantly higher protein levels (mean protein level [95% CI] = 5.14 [4.88 - 5.40], p=0.0004) in comparison with the reference haplotype 1G3C (mean casein level [95% CI] = 4.52 [4.32 – 4.72]). To provide reliable results, we set stringent significance cut-off values by applying the Bonferroni correction ($p < 0.008$), and consequently, we did not consider the association of 1G3T haplotype with casein level significant. The study power for haplotype effect on casein levels was 85% for α=0.008, and for haplotype effect on protein levels was 65% at the same stringent significance level of 0.008. The percentage of variance explained by these haplotypes was 40% for caseins and 31% for proteins.

Discussion

The main buffalo species in Europe is the water buffalo, with the Mediterranean water buffalo as the predominant breed. There are genetic variants among geographic regions, and their detection is essential for the authentication of Protected Denomination of Origin (PDO) products. These tests are often based on analysis of the major milk protein genes. Casein genes have a great impact on milk quality, as expected because caseins are the most abundant proteins in milk. Also, casein genes are of interest in evolutionary studies, as well as in selection and refinement [18]. According to the Food and Agriculture Organization of the United Nations (FAO/DAD-IS, https://www.fao.org/dad-is/browse-by-country-and-species/en/), in Serbia, the buffalo population is increasing, but it is still marked as endangered. It is crucial to genetically characterize the local buffalo population for future selection and breeding. Knowledge of the genetic background and careful selection are essential to avoid inbreeding and to maintain genetic diversity in the population [19]. Inbreeding has various adverse effects on milk production and quality traits due to the small effective size of the population, decreased response to selection, and reduced animal performance. Significant inbreeding has already been noticed in buffalo populations [20], and the genetic potential of never-selected buffalo populations might be useful for reintroduction purposes [21]. Milk protein and fat percentages in the investigated Serbian buffalo population were comparable with other buffalo populations [3,22-23], though different from the milk quality reported in a previous study on buffalos from the Pešter region [24], which reported much lower fat percentages in milk.

In this study, we have analyzed the *CSN1S1* gene variant, which was detected in the neighboring Romanian buffalo population but has not been found in any other buffalo population [5]. *CSN1S1* 472G>C is an intronic variant that leads to skipping exon six and produces a protein lacking eight amino acids. We detected this variant in the Serbian buffalo population of 128 animals with a frequency of 0.23 for the minor C allele, while in the Romanian buffalo population of 160 animals, it was 0.18. We have found an association between the variant *CSN1S1* 472G>C and the milk casein and protein levels. Even though Balteanu et al. implied that the 472G>C variant causes a defective protein, functional studies are needed to support that premise. Our results suggest that milk from animals with the minor C allele has more caseins and proteins. The association of variant *CSN1S1* 472G>C with milk quality traits has not been previously analyzed, so our results are the first regarding this variant's effect on milk composition. It would be of interest to explore the presence of this variant in other local neighboring populations so we could define our buffalo population in terms of milk protein-coding genes compared with those of neighboring populations. A limitation of our study was that we measured the total casein percentage in milk samples and not the percentage of specific casein types, which would have given us information on αS1 casein presence in milk samples. As shown in cattle [7], variants in casein genes affect the concentration of the specific caseins in milk. T allele of gene variant *CSN3* 467C>T was associated with a lower concentration of αS1 casein and a higher concentration of αS2 and κ casein in cow milk.

Variant *CSN3* 467C>T has been explored by others in buffalo breeds, but in smaller sample sizes and mainly not in association with milk quality traits [3,9,25–28]. Fan et al. performed an extensive analysis of *CSN3* gene variants. It has shown that in Chinese water buffalo population frequencies of minor alleles were found to be less than 0.1 for investigated variants (*CSN3* 445G>A, 467C>T 471C>T and 516A>C) [9]. This indicates that the Chinese water buffalo population is very different from other water buffalo populations around the world due to migration and crossbreeding events. Since, variant *CSN3* 467C>T was among these investigated variants, it gave us the reason to assume that other variants investigated in that study could be polymorphic in Serbian buffalo population. Nevertheless, *CSN3* variants 445G>A and 516A>C were not polymorphic in our buffalo population and *CSN3* 471C>T was in complete LD with variant 467C>T. Variant 471C>T is silent and it does not lead to an amino acid change, while *CSN3* 467C>T changes the amino acid threonine to isoleucine and has a functional effect on κ casein protein. This variant is the only *CSN3* variant shared between buffalo and cattle, while other variants found in buffalo *CSN3* gene do not exist in cattle, and *vice versa* [9]. Allele T of the variant *CSN3* 467C>T has been associated with higher casein, protein, and fat content in cow milk [7]. Our study analyzed six milk quality traits, and the minor allele T showed a significant effect on a casein level in milk. Of previously mentioned studies, just one addressed milk quality traits and associated allele T with higher protein and casein levels in buffalo milk [26].

We have analyzed *CSN1S1* 472G>C and *CSN3* 467C>T haplotype effects on milk quality traits. Haplotype analysis can provide insight into the true effect of variants inherited together in a population. The most abundant haplotype in the investigated Serbian buffalo population, with a frequency of 0.40, was haplotype 1G3C, inferred from wild-type alleles of both variants. By comparison, haplotype 1C3C consisted of *CSN1S1* 472G>C minor allele C and *CSN3* 467C>T wild type allele C, had higher levels of caseins and proteins in buffalo milk. So far, only two variants in casein genes have been analyzed in haplotype *CSN1S1* 628C>T and *CSN3* 467C>T, and mainly with regard to milk yield and coagulation properties [3,26]. Our individual analysis showed that allele T of gene variant *CSN3* 467C>T was associated with higher levels of caseins in milk, and we confirmed this in haplotype analysis, showing that haplotype 1G3T had a higher level of caseins than haplotype 1G3C. Despite this, the haplotype 1C3C had the largest increase in milk casein levels. Based on individual variant analysis, it would be expected that the 1C3T haplotype has the greatest effect on casein, but this haplotype is so rare that it was not even detected in our buffalo population.

In conclusion, we have shown that genetic variants *CSN1S1* 472G>C and *CSN3* 467C>T have a significant impact on casein levels in buffalo milk. *CSN1S1* minor allele C was associated with higher levels of caseins and proteins in milk. In addition, haplotype analysis of these genetic variants provides greater validity and reliability in association with casein levels compared with individual analysis. Our results contribute to knowledge of the buffalo genome by characterizing the buffalo population in Serbia, which has not been performed to date, and provide potentially valuable perspectives for applications in improving milk quality.

Acknowledgments

This study was funded by the Ministry of Science, Technological Development and Innovation of the Republic of Serbia (Agreement no. 451-03-66/2024-03/200017 and 451-03-65/2024-03/200116).

References

1. Li Z, Lu S, Cui K, Shafique L, Rehman S ur, Luo C, et al. Fatty acid biosynthesis and transcriptional regulation of Stearoyl-CoA Desaturase 1 (SCD1) in buffalo milk. BMC Genet. 2020 Dec 2;21(1):23.

2. Zicarelli L. Buffalo Milk: Its Properties, Dairy Yield and Mozzarella Production. Vet Res Commun. 2004; 28:127–35.

3. Bonfatti V, Giantin M, Gervaso M, Coletta A, Dacasto M, Carnier P. Effect of CSN1S1-CSN3 (αS1-κ-casein) composite genotype on milk production traits and milk coagulation properties in Mediterranean water buffalo. J Dairy Sci. 2012 Jun;95(6):3435–43.

4. Cosenza G, Pauciullo A, Macciotta NPP, Apicella E, Steri R, La Battaglia A, et al. Mediterranean river buffalo CSN1S1 gene: search for polymorphisms and association studies. Anim Prod Sci. 2015;55(5):654.

5. Balteanu VA, Carsai TC, Vlaic A. Identification of an intronic regulatory mutation at the buffalo αS1-casein gene that triggers the skipping of exon 6. Mol Biol Rep. 2013 Jul 3;40(7):4311–6.

6. Alexander M, Rojas-Ochoa LF, Leser M, Schurtenberger P. Structure, Dynamics, and Optical Properties of Concentrated Milk Suspensions: An Analogy to Hard-Sphere Liquids. J Colloid Interface Sci. 2002 Sep;253(1):35–46.

7. Heck JML, Schennink A, van Valenberg HJF, Bovenhuis H, Visker MHPW, van Arendonk JAM, et al. Effects of milk protein variants on the protein composition of bovine milk. J Dairy Sci. 2009 Mar;92(3):1192–202.

8. Perisic P, Bogdanovic V, Mekic C, Ruzic-Muslic D, Stanojevic D, Popovac M, et al. The importance of buffalo in milk production and buffalo population in Serbia. Biotechnology in Animal Husbandry. 2015;31(2):255–63.

9. Fan X, Zhang Z, Qiu L, Zhang Y, Miao Y. Polymorphisms of the kappa casein (CSN3) gene and inference of its variants in water buffalo (Bubalus bubalis). Arch Anim Breed. 2019 Dec 5;62(2):585–96.

10. Fabro MA, Milanesio HV, Robert LM, Speranza JL, Murphy M, Rodríguez G, et al. Technical Note: Determination of Acidity in Whole Raw Milk: Comparison of Results Obtained by Two Different Analytical Methods. J Dairy Sci. 2006 Mar;89(3):859–61.

11. IDF: Milk, cream and evaporated milk. In Determination of Total Solids Content (Vol. O21B). Brussels, Belgium: International Dairy Federation. 1987.

12. IDF: Determination of fat content. Gerber Butyrometers (Vol. 105) Brussels, Belgium: International Dairy Federation. 1981.

13. AOAC: Official Methods of Analysis, 15th edition. Washington, DC, USA: Association of Official Analytical Chemists, 1990. 807–809, 841-842.

14. Weiß CH. StatSoft, Inc., Tulsa, OK.: STATISTICA, Version 8. AStA Advances in Statistical Analysis.

2007 Sep 27;91(3):339–41.

15. Tregouet DA, Garelle V. A new JAVA interface implementation of THESIAS: testing haplotype effects in association studies. Bioinformatics. 2007 Apr 15;23(8):1038–9.

16. Tregouet DA, Escolano S, Tiret L, Mallet A, Golmard JL. A new algorithm for haplotype-based association analysis: the Stochastic-EM algorithm. Ann Hum Genet. 2004 Mar 9;68(2):165–77.

17. Rosner B. Fundamentals of Biostatistics. 7th ed. Boston, MA: Brooks/Cole; 2011.

18. Rehman S ur, Hassan F ul, Luo X, Li Z, Liu Q. Whole-Genome Sequencing and Characterization of Buffalo Genetic Resources: Recent Advances and Future Challenges. Animals. 2021 Mar 22;11(3):904.

19. Doekes HP, Bijma P, Windig JJ. How depressing is inbreeding? A meta-analysis of 30 years of research on the effects of inbreeding in livestock. Genes (Basel). 2021 Jun 1;12(6).

20. Santana ML, Aspilcueta-Borquis RR, Bignardi AB, Albuquerque LG, Tonhati H. Population structure and effects of inbreeding on milk yield and quality of Murrah buffaloes. J Dairy Sci. 2011 Oct;94(10):5204–11.

21. Strandén I, Kantanen J, Russo IRM, Orozco-terWengel P, Bruford MW. Genomic selection strategies for breeding adaptation and production in dairy cattle under climate change. Heredity (Edinb). 2019 Sep 18;123(3):307–17.

22. Li J, Liu S, Li Z, Zhang S, Hua G, Salzano A, et al. DGAT1 polymorphism in Riverine buffalo, Swamp buffalo and crossbred buffalo. Journal of Dairy Research. 2018 Nov 2;85(4):412–5.

23. Braun, P., & Preuss, S. Nutritional composition and chemico-physical parameters of water buffalo milk and milk products in Germany. Milchwissenschaft-milk Science International, 2008. 63, 70-72.

24. Becskei Z, Savić M, Ćirković D, Rašeta M, Puvača N, Pajić M, et al. Assessment of Water Buffalo Milk and Traditional Milk Products in a Sustainable Production System. Sustainability. 2020 Aug 15;12(16):6616.

25. Asim M, Saif-ur Rehman M, Hassan F ul, Awan FS. Genetic variants of CSN1S1, CSN2, CSN3 and BLG genes and their association with dairy production traits in Sahiwal cattle and Nili-Ravi buffaloes. Anim Biotechnol. 2023 Dec 27;34(7):2951–62.

26. Zicarelli L, Di Palo R, Napolano R, Tonhati H, De Carlo E, Gagliardi R, et al. Influence of αS1-casein and κ-casein polymorphism on the curd yield of Italian Mediterranean buffalo (Bubalus bubalis L.) milk. Int Dairy J. 2020 Jan 1;100.

27. El Nahas SM, Bibars MA, Taha DA. Genetic characterization of Egyptian buffalo CSN3 gene. Journal of Genetic Engineering and Biotechnology. 2013 Dec;11(2):123–7.

28. Medeiros NBC, GUERREIRO SLM, PEREIRA RSR, SÁ AL, RODRIGUES M, MEZZOMO R, et al. CSN1S1 and CSN3 gene variants in female Murrah buffaloes in the Brazilian Amazon. An Acad Bras Cienc. 2022;94(suppl 3).

Tables and Figures

Table 1. Genotypes distribution and allele frequencies of *CSN1S1* 472G>C and *CSN3* 467C>T variants in the investigated Serbian buffalo population

	Serbian Buffalo Population
***CSN1S1* 472 G>C**	N=128
Genotypes	Genotype frequencies, n (%)
GG	74 (57.81)
GC	48 (37.50)
CC	6 (4.69)
MAF (Allele C)	0.23
***CSN3* 467 C>T**	N=125
Genotypes	Genotype frequencies, n (%)
CC	43 (34.40)
CT	62 (49.60)
TT	20 (16.00)
MAF (Allele T)	0.41

MAF – minor allele frequency

311
312 **Table 2.** Composition of milk for the studied subgroup of 42 buffalos and casein correlations
313 with other milk components

Milk components (%)	Mean ± SD	Spearman R	*p* value
Casein	3.58 ± 0.29		
Protein	4.73 ± 0.41	0.93	<0.00001
Fat	9.05 ± 1.02	0.226	0.153
Acidity SH°	8.83 ± 1.08	0.316	0.04
Total solid	19.65 ± 1.25	0.47	0.0017
solids-not-fat	10.58 ± 0.49	0.61	0.00002

314
315

316 **Table 3.** Haplotype frequencies of *CSN1S1* 472G>C and *CSN3* 467C>T and their effects on
317 milk quality traits

CSN1S1* 472 G>C and *CSN3* 467 C>T Haplotypes	**1G3C**	**1G3T**	**1C3C**
Haplotype frequencies	0.40	0.38	0.22
Mean casein % [95% CI]	3.34 [3.18 - 3.50]	3.68 [3.48 - 3.88]	3.90 [3.72 - 4.06]
p value	reference haplotype	0.02	0.00002
Mean protein % [95% CI]	4.52 [4.32 - 4.72]	4.84 [4.58 - 5.10]	5.14 [4.88 - 5.40]
p value	reference haplotype	0.08	0.0004
Mean fat % [95% CI]	8.82 [8.20 - 9.44]	9.14 [8.38 - 9.92]	9.42[8.44 - 10.42]
p value	reference haplotype	0.56	0.31
Mean acidity SH° [95% CI]	8.56 [8.08 - 9.02]	9.24 [8.54 - 9.96]	8.56 [7.00 - 10.12]
p value	reference haplotype	0.13	0.99
Mean total solid % [95% CI]	19.22 [18.38 - 20.06]	19.94 [19.02 - 20.84]	20.34 [19.20 - 21.46]
p value	reference haplotype	0.33	0.09
Mean solids-not-fat % [95% CI]	10.38 [10.14 - 10.62]	10.78 [10.38 - 11.18]	10.88 [10.48 - 11.28]
p value	reference haplotype	0.08	0.06

318 * - The alleles in haplotypes are in the following order: *CSN1S1* 472G>C and *CSN3* 467C>T.
319 According to the Thesias software 1G3C haplotype was set as the reference haplotype, as the
320 most frequent haplotype in the investigated buffalo population. Haplotype 1C3T was not
321 detected in these buffaloes. *p* values <0.008 were considered statistically significant.

322
323
324

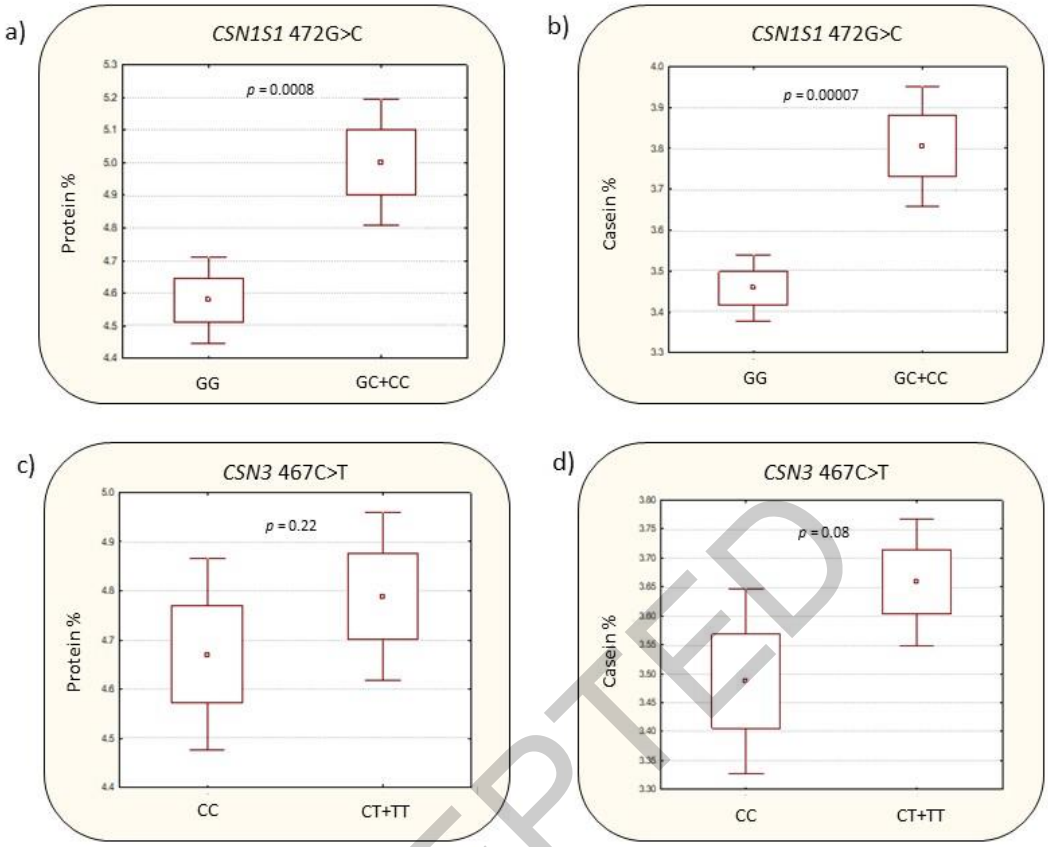

Figure 1. Association of *CSN1S1* 472G>C genotypes by the dominant model of inheritance, GG vs. GC+CC, with a) protein and b) casein percentage in buffalo milk and association of *CSN3* 467C>T genotypes by the dominant model of inheritance, TT vs. CT+TT, with c) protein and d) casein percentage in buffalo milk