A novel duplex ACRS-PCR for composite CSN1S1–CSN3 genotype discrimination in domestic buffalo

Alfredo Pauciulloa, Sara Martorelloa, Kejsi Carkua, Carmine Versacea, Angelo Colettab and Gianfranco Cosenzaa

ABSTRACT
This short communication aimed to develop a fast and straightforward method for the simultaneous discrimination in buffaloes of the alleles CSN1S1 A/B and CSN3 X1/X2 using a single protocol. DNA was isolated from 219 individual blood samples. A duplex artificially created restriction site (ACRS-PCR) was accomplished using two pairs of primers generating 86 bp (CSN1S1) and 160 bp (CSN3). Amplicons were contemporary digested by MboI and HinfI for the identification of genotypes.

The double simultaneous amplification and digestion proved to be effective for allele identification. The method resulted particularly quick due to the small PCR amplicons and fast digest enzymes that allowed both a rapid amplification (about 1 h and 40 min) and digestion in 10 min. Population analysis indicated that minor allele frequencies were CSN1S1 A (0.425) and CSN3 X2 (0.306). Linkage disequilibrium (LD) showed r^2 = 0.46, and haplotype analysis revealed all four possible combinations with higher frequency for the CSN1S1 B-CSN3 X1 (0.553).

Considering the tremendous economic impact of the CSN1S1–CSN3 variants on the dairy production in buffalo, this method, applicable immediately after the birth from any DNA source, may speed up the selection of sires and dams’ lines with more favourable genotypes.

HIGHLIGHTS
- CSN1S1–CSN3 composite genotype influences dairy performances.
- Duplex ACRS-PCR identifies the CSN1S1–CSN3 composite genotype in a single protocol.
- The duplex ACRS-PCR method may speed up the selection of lines with more favourable genotypes.

Introduction
Caseins are the main proteins in ruminant milk and the influence of their genetic variants on milk composition and cheesemaking properties is well known to such an extent that sires selection takes into account their genotype at casein loci. In this respect, examples in cattle are represented by the variants of the κ-casein (CSN3 A and B) for dairy processing milk or the alleles of the β-casein (CSN2 A1 and A2) for drinking milk. The first has a quantitative effect on dairy yield, with the CSN3*B variant reacting more promptly with rennet, having a curd coagulation time significantly shorter than milk with CSN3*A so that sires with genotype CSN3 BB are favourite. The latter influences the composition in bio-peptides derived by enzymatic digestion of the β-casein. The CSN2*A1 variant releases the β-casomorphin-7 (BCM-7), an opioid affecting the gastrointestinal functions, including regulating motility and mucus production (Pal et al. 2015). Although the literature reports conflicting data about the BCM-7 effects, there is extensive evidence from animal trials and emerging evidence in humans that the BCM-7 is associated with slower gastrointestinal transit and hence increased gastrointestinal transit times (Brook-Taylor et al. 2017). In the last years, this result promoted the selection of sires with genotype CSN2 A2A2. The interactions among variants of the complete casein cluster have also effects on production traits. For instance, the β-casein (CSN2*B) together with the β-lactoglobulin (LGB*B) were found to be
more favourable for curd coagulation, firmness, and cheesemaking quality of milk (Di Stasio and Mariani 2000). In Italian Holsteins, Comin et al. (2008) found that CSN3 and CSN2 genotypes were strongly associated with milk coagulation traits and milk and protein yields, respectively. In particular, for coagulation time and curd firmness, the best CSN2–CSN3 composite genotypes were those with at least one B allele at both loci (A1/B–A/B, A2/B–B/B, and A2/B–A/B), whereas for milk and protein yield, the best genotype was the most frequent genotype (A2/A2–A/A) (Comin et al. 2008).

Moreover, the composite genotypes CSN1S1 BB-CSN2 A2A2-CSN3 BB showed the best coagulation characteristics and highest cheese yield in cattle (Perna et al. 2016; Albarella et al. 2020), so that composite genotype was proposed to be the most appropriate criterion for selection decisions.

Genetic variants associated with milk production (Pauciullo et al. 2012a, 2012b) or affecting the milk protein yield (Cosenza et al. 2015) are also known in river buffalo. In this respect, the occurrence of two DNA transversions at CSN1S1 and CSN3 might have a great economic impact on dairy yield and, consequently, on the selling price of buffalo milk mainly established by fat content and protein, and finally on sire selection for the genetic improvement of the breed. The first mutation, AJ005430:c.578C>T, realised at the 83rd nucleotide of the exon 17 of CSN1S1 is responsible for the amino acid replacement p.Ser178(B allele)/Leu178(A allele) of the mature αs1-casein (Chianese et al. 2009). The CSN1S1 B allele has been associated with higher protein content (Cosenza et al. 2015). The second transversion, HQ677596:c.536C>T, realised at the nucleotide 377 of the exon 4 of CSN3 leads to the amino acid change p.Ile135(X1 allele)/Thr135(X2 allele) of the mature α-casein (Mitra et al. 1998).

The effects of CSN1S1–CSN3 composite genotypes of these variants on milk production traits and milk coagulation properties have been well described in the Mediterranean buffalo (Bonfatti et al. 2012, 2013). For instance, the genotype AA-X1X2 exhibited the shortest rennet coagulation time (RCT) and a similar effect was revealed for the curd-firming time (K20). In addition to shortest RCT and K20, genotype AA-X1X2 also revealed the largest curd firmness (A30). Furthermore, recently Zicarelli et al. (2020) demonstrated that milk carrying CSN1S1*B and CSN3*X1 alleles resulted in a greater curd yield.

This short communication aimed to develop, for the first time, a simple and fast method for the simultaneous discrimination of the alleles CSN1S1 A/B and CSN3 X1/X2 in a single protocol. This method might be helpful for buffalo genotyping and animal selection independently from the milk production.

**Material and methods**

**Sampling and DNA isolation**

Sample collection was carried out on a total of 219 Italian Mediterranean river buffaloes belonging to one advanced dairy farm located in the Piedmont region (Northern Italy). Individual blood samples were collected during the routine prophylaxis of the farm by an official veterinarian of ASL (Local Sanitary Unit) of the Ministry of Health. For this reason, the Animal Care and Use Committee approval was not necessary.

Genomic DNA was isolated using the NucleoSpin Tissue kit (Macherey-Nagel, Düren, Germany). Concentrations and OD260/280 ratios were measured with the Nanodrop ND-1000 Spectrophotometer (Thermo Fisher Scientific Inc., Waltham, MA).

**Duplex ACRS-PCR and sequencing**

The investigated mutations (CSN1S1 AJ005430:c.578C>T and CSN3 HQ677596:c.536C>T) do not alter recognition sites of endonuclease. Therefore, we decided to set up a duplex Artificially Created Restriction Site (ACRS)-PCR for allele discrimination. For this reason, two pairs of primers were designed: 5’-CAATACCCCTGATGCCCGAT-3’ (forward) and 5’-CACCACAGTGCCATAGTAG-3’ (reverse) amplified a fragment of 86 bp at the CSN1S1 (Cosenza et al. 2015), whereas 5’-TGTTTGAGCCTACAAATGCAA-3’ (forward) and 5’-GTTGGTCTTCATGTGCC-3’ (reverse) gave an amplicon of 160 bp for the CSN3. In both methods, both forward primers were modified: the first by changing C→G to create the restriction site for Mbo I (I/GATC), and the second by replacing CT→GA to generate G1 ANTCC restriction site for the Hind I endonuclease.

Duplex PCR amplification was carried out in a final volume of 15 μL, including 50 ng of genomic DNA, 1× PCR Buffer (Promega, Madison, WI), 2.5 mM MgCl2, 5 pmol of each primer, dNTPs 200 μM each, 1 U of Taq DNA Polymerase (Promega). The thermal conditions were: 95°C for 4 min, 40 cycles at 95°C for 30 s, annealing at 55°C for 40 s, and extension at 72°C for 20 s. A final extension was carried out at 72°C for 5 min. The product specificity was confirmed by electrophoresis, loading 5 μL of PCR product on 3.0% agarose gel in 0.5× TBE buffer, stained with Sybr Green.
The remaining 10 µL were contemporary digested with 1 µL of both Fastdigest Mbo I and Hinf I endonucleases (Thermo Fisher Scientific Inc.) for 10 min at 37°C according to manufacturer’s guidelines. The digested products were resolved on 3.5% agarose gel in 0.5x TBE buffer stained with Sybr Green for genotype detection.

For the validation and confirmation of the duplex ACRS-PCR genotype results, twelve informative samples (2 AA-X1X1, 3 AA-X2X2, 3 AB-X1X2, 3 BB-X1X1, 1 BB-X2X2) that resulted both double homozygous and heterozygous by ACRS-PCR were amplified in single PCR reactions and sequenced by Sanger Technology at Microsynth (Vienna, Austria).

**Statistical analysis**

Allele frequencies were calculated by Popgene software (University of Alberta, Edmonton, Alberta, Canada), whereas Phase 2.1 (Stephens and Scheet 2005) was used to determine the haplotype frequencies, and Haploview 4.2 (Barrett et al. 2005) was used to define linkage disequilibrium.

**Results and discussion**

Domestic buffalo is an expanding breeding reality worldwide. This is mainly due to the rheological characteristics of its milk, known to have higher protein and fat content compared to other ruminant’s milk. These properties make it particularly suitable for the cheesemaking with a higher yield production (Zicarelli 2004; Zicarelli et al. 2020). Despite the potential economic impact on the dairy industry, increasing buffalo milk production with optimal percentages of fats and proteins through an adequate genetic selection process still reveal delays. This is also due to the lack of fast methods for the discrimination of favourable alleles for more profitable breeding strategies, such as the A/B and X1/X2 alleles at αs1- and κ-casein genes, respectively (Bonfatti et al. 2012; Bonfatti et al. 2013; Cosenza et al. 2015; Zicarelli et al. 2020).

To date, RP-HPLC of milk protein and/or DNA sequencing were the most used methods for genotype determination at CSN1S1 and CSN3 genes (Chianese et al. 2009; Bonfatti et al. 2012; Bonfatti et al. 2013; Zicarelli et al. 2020). However, both methods require several working steps and have limitations. For instance, RP-HPLC cannot be used for sires genotyping, whereas sequencing needs the development of separated PCR protocols, purifications, formamide treatments before sequencing, etc. To overcome these restrictions, in this study, we developed for the first time a duplex ACRS-PCR method followed by a digestion for the contemporary discrimination of CSN1S1 A/B and CSN3 X1/X2 alleles.

The double simultaneous amplification and digestion proved to be effective for genotype identification (Figure 1). For CSN1S1 AJ005430:c.578C>T, the digestion of the PCR product (86 bp) with the Mbo I endonuclease produced two fragments (70 and 16 bp) for the presence of the cytosine (allele B), and undigested fragment (86 bp) for the presence of the thymine (allele A). Simultaneously, for the SNP CSN3 HQ677596:c.536C>T, using the endonuclease Hinf I, in presence of the cytosine (allele X1), the fragment remained undigested (160 bp), whereas in the occurrence of the thymine (allele X2), two fragments (137 and 23 bp) were obtained (Figure 1).

The method resulted particularly fast due to the small PCR amplicons and the use of fast digest enzymes that allowed both a rapid amplification
(about 1 h and 40 min) and the digestion after 10 min of incubation time only. Sanger sequencing was performed for the validation of the method. Electropherogram analysis confirmed the result of the genotyping by ACRS-PCR (Figure 2). However, the whole procedure (set up of two amplification protocols, PCR purification, sequencing reaction, sequencing run) resulted in time-consuming compared to the method proposed herein, and more expensive in terms of costs. We assessed a saving of 75% for the costs sustained with our method versus the sequencing. However, we do not exclude a certain range of saving, depending on the market evolution, country, buyers–sellers agreements for reagents’ orders, etc.

LD analysis performed with our data showed $r^2=0.46$, and haplotype analysis revealed all four possible combinations with higher frequency for the CSN1S1 B-CSN3 X1 (Table 1). Conversely, a high linkage disequilibrium ($p<.001$) between CSN1S1 and CSN3 genes were found by Bonfatti et al. (2012). However, this is a relatively rare event. In fact, it is well-known that many recombination events characterise the casein cluster (Caroli et al. 2009; Pauciullo et al. 2019).

The most favourable haplotype for dairy purposes observed in our study (CSN1S1 A-CSN3 X2 according to the original nomenclature) showed a frequency of 0.279. According to Zicarelli et al. (2020), buffaloes carrying these casein variants produced a greater curd yield due to higher proteins rather than milk production, and greater real curd yield/estimated curd yield ratio (RCY/ECY).

Considering the relatively low frequency of this haplotype (CSN1S1 A-CSN3 X2), the choice of animals with better dairy performances and the consequent indirect selective pressure applied in this advanced buffalo farm, whose exclusive attitude is dairy production, might leave very wide opportunities of genetic progress.

Conversely, the most frequent haplotype (0.553) observed in the present study was CSN1S1 B-CSN3 X1. According to the original nomenclature, this haplotype corresponds to a lower curd yield (Zicarelli et al. 2020). In general, the negative effect on the dairy characteristics was mainly attributed to the CSN3 X1 allele, indicated as the worst for the RCY/ECY ratio, independently from its association with any CSN1S1 genotype (Zicarelli et al. 2020).

Although the results refer to a single herd and should be tested population-wide, it is of great importance as the method appeared effective for the recognition of composite CSN1S1 and CSN3 genotypes in domestic buffalo, one of the most relevant dairy traits for its influence on milk production and cheese-making properties.

In conclusion, a new very fast protocol based on a single ACRS-PCR was set up for the simultaneous genotyping of CSN1S1 A/B-CSN3 X1/X2 alleles, known to have significant effects on dairy performances of river buffaloes. The protocol minimises the cost of analyses on a large scale (up to 75% cheaper than the sequencing) and gives most laboratories with basic equipment and average economic resources to perform this assay. Furthermore, considering the great

![Figure 2. Sanger sequencing of the informative DNA samples. Double homozygous and double heterozygous at CSN1S1 and CSN3 loci as confirmation of the developed genotyping method.](image-url)
economic impact on dairy production, this method, applicable immediately after the buffalo birth from any source of DNA, may speed up the selection of sires and dams’ lines with more favourable genotypes.

**Ethical approval**

All procedures observed the Directive 98/58/EC concerning the protection of animals. Sampling has been carried out during the routine prophylaxis of the farm by an official veterinarian of ASL (Local Sanitary Unit) of the Ministry of Health. Therefore, according to the Committee on the Ethics of Animal Experiments of the University of Torino (D.R. n. 2128 released on 06/11/2015) further ethics approval was not required.

**Acknowledgements**

The authors thank Morisiasco Ivan and Moris farm in the Piedmont region for providing river buffalo samples.

**Disclosure statement**

The authors declare no conflict of interest with any financial organisation regarding the material discussed in the manuscript.

**Funding**

This work was partially supported by the Italian Ministry of Research [Project no. PON01_00486 GENOBU] and by Ricerca Innovazione e Selezione per la Bufala [Project no. PAUA_RIC_N_COMP_20_02].

**ORCID**

Alfredo Pauciullo http://orcid.org/0000-0002-3140-9373

Gianfranco Cosenza http://orcid.org/0000-0001-6006-4987

**Data availability statement**

The authors confirm that the data supporting the findings of this study are available within the article.

**Table 1.** Genotypes, allele, and haplotype frequencies detected at the CSN1S1 and CSN3 genes in the Italian Mediterranean river buffalo breed.

| Genotypes | Allele frequencies | Allele frequencies |
|-----------|--------------------|--------------------|
| CSN3      | CSN1S1             | CSN3               |
|           | X1X1               | X1X2               | X2X2               | Total | A      | B      | X1     | X2     | BX1    | Frequency | SE     |
| CSN1S1    |                     | 0.425              | 0.575              | 0.694              | 0.306  | 0.425  | 0.575  | 0.694  | 0.306  | 0.553  | 0.0024  |
| AA        |                    |                    |                    |                    | 0.553  | 0.0024 |
| AB        |                    |                    |                    |                    | 0.279  | 0.0024 |
| BB        |                    |                    |                    |                    | 0.023  | 0.0024 |

**References**

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

Barrett JC, Fry B, Maller J, Daly MJ. 2005. Haploviz: analysis and visualization of LD and haplotype maps. Bioinformatics. 21(2):263–265.

Bonfatti V, Gervaso M, Rostellato R, Coletta A, Carrier P. 2013. Protein composition affects variation in coagulation properties of buffalo milk. J Dairy Sci. 96(7):4182–4190.

Bonfatti V, Giantin M, Gervaso M, Coletta A, Dacasto M, Carrier P. 2012. Effect of CSN1S1-CSN3 (αS1-κ-casein) composite genotype on milk production traits and milk coagulation properties in Mediterranean water buffalo. J Dairy Sci. 95(6):3435–3443.

Brook-Taylor S, Dwyer K, Woodford K, Kost N. 2017. Systematic review of the gastrointestinal effects of A1 compared with A2 β-casein. Adv Nutr. 8(5):739–738.

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

Chianese L, Quarto M, Pizzolongo F, Calabrese MG, Caira S, Mauriello R, De Pascale S, Addeo F. 2009. Occurrence of genetic polymorphism at the αs1-casein locus in Mediterranean water buffalo milk. Int Dairy J. 19(4):181–189.

Comin A, Cassandro M, Chessa S, Ojala M, Dal Zotto R, De Marchi M, Carrier P, Gallo L, Pagnacco G, Bittante G. 2008. Effects of composite beta- and kappa-casein genotypes on milk coagulation, quality, and yield traits in Italian Holstein cows. J Dairy Sci. 91(10):4022–4027.

Cosenza G, Pauciullo A, Macciotta N, Apicella E, Steri R, La Battaglia A, Jemma L, Coletta A, Di Berardino D, Ramunno L. 2015. Mediterranean river buffalo CSN1S1 gene: search for polymorphisms and association studies. Anim Prod Sci. 55(5):654–660.

Di Stasio L, Mariani P. 2000. The role of protein polymorphism in the genetic improvement of milk production. Zootec Nutri Anim. 26:69–90.

Mitra A, Schlee P, Krause I, Blusch J, Werner T, Balakrishnan C, Pirchner F. 1998. Kappa-casein polymorphisms in Indian dairy cattle and buffalo: a new genetic variant in buffalo. Anim Biotechnol. 9(2):81–87.

Pal S, Woodford K, Kukuljan S, Ho S. 2015. Milk intolerance, beta-casein and lactose. Nutrients. 7(9):7285–7297.
Pauciullo A, Cosenza G, Steri R, Coletta A, Jemma L, Feligini M, Di Berardino D, Macciotta NP, Ramunno L. 2012a. An association analysis between OXT genotype and milk yield and flow in Italian Mediterranean river buffalo. J Dairy Res. 79(2):150–156.

Pauciullo A, Cosenza G, Steri R, Coletta AL, Battaglia A, Di Berardino D, Macciotta NP, Ramunno L. 2012b. A single nucleotide polymorphism in the promoter region of river buffalo stearoyl CoA desaturase gene (SCD) is associated with milk yield. J Dairy Res. 79(4):429–435.

Pauciullo A, Shuiep ET, Ogah MD, Cosenza G, Di Stasio L, Erhardt G. 2019. Casein gene cluster in camelids: comparative genome analysis and new findings on haplotype variability and physical mapping. Front Genet. 10:748.

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.

Stephens M, Scheet P. 2005. Accounting for decay of linkage disequilibrium in haplotype inference and missing-data imputation. Am J Hum Genet. 76(3):449–462.

Zicarelli L. 2004. Buffalo milk: its properties, dairy yield and mozzarella production. Vet Res Commun. 28:127–135.

Zicarelli L, Di Palo R, Napolano R, Tonhati HD, Carlo E, Gagliardi R, Di Luccia A, la Gatta B. 2020. Influence of αS1-casein and κ-casein polymorphism on the curd yield of Italian Mediterranean buffalo (Bubalus bubalis L.) milk. Int Dairy J. 100:104559.