Mediterranean River Buffalo CSN1S1 gene: search for polymorphisms and association studies.

Published version:

DOI:10.1071/AN13438

(Article begins on next page)

Mediterranean River Buffalo *CSN1S1* gene: search for polymorphisms and association studies

G. Cosenza [A,F], A. Pauciullo [B,F], N. P. P. Macciotta[C], E. Apicella [A], R. Steri [C], A. La Battaglia [D], L. Jemma [A], A. Coletta [E], D. Di Berardino [A], L. Ramunno [A]

[A]Dipartimento di Agraria, Università degli Studi di Napoli “Federico II”, via Università 100, Portici, NA 80055 Italy

[B]Istituto per il Sistema Produzione Animale in Ambiente Mediterraneo, Consiglio Nazionale delle Ricerche, via Argine 1085, Napoli, 80147 Italy

[C]Dipartimento di Agraria, Università degli Studi di Sassari, via De Nicola 9, Sassari, 07100 Italy

[D]Comunità Montana Alto Agri, Villa D’Agri di Marsicovetere (PZ), 85050 Italy

[E]Associazione Nazionale Allevatori Specie Bufalina, Via F. Petrarca 42-44 Località Centurano, Caserta, 81100 Italy

[F]Corresponding author: Email: giacosen@unina.it; alfredo.pauciullo@cnr.it

Abstract. The aim of the present work was to study the variability at *CSN1S1 locus* of the Italian Mediterranean river buffalo and to investigate possible allele effects on milk yield and its composition. Effects of parity, calving season and month of production were also evaluated. Three SNPs were detected. The first mutation, located at position 89 of 17th exon (c.628C>T), is responsible for the amino acid change p.Ser178 (B allele)/Leu178 (A allele). The other two polymorphisms, detected at the positions 144 (c.882G>A) and 239 (c.977A>G) of 19th exon respectively, are silent (3’ UTR).

Associations between the *CSN1S1* genotypes and milk production traits were investigated using 4,122 test day records of 503 lactations from 175 buffalo cows. Milk yield, fat and protein percentages were analyzed using a mixed linear model. A significant association between the c.628C>T SNP and the protein percentage was found. In particular, the CC genotype showed an average value of about 0.04% higher than the CT and TT genotypes. The allele substitution effect of the cytosine into the thymine was -0.014, with a quite low (0.3%) protein percentage (PP) contribution on total phenotypic variance. A large dominance effect was detected.

Furthermore, a characterization of the *CSN1S1* transcripts and a method based on *Mbo*I-ACRS-PCR for a rapid genotyping of c.628C>T were provided.

Additional keywords: Mediterranean river buffalo, *Bubalus bubalis***,** *CSN1S1***,** milk protein percentage, marker assisted selection

Introduction

Historically domestic water buffalo were divided into swamp and river subspecies that differ in morphology, behaviour, and chromosome number (2n=48 and 2n=50, respectively) (Ajmone-Masan P. *et al*, 2013). In particular, the Mediterranean river buffalo, is the only indigenous Italian breed of water buffalo. It is of the River subtype of water buffalo, and is genetically and ecological similar to the buffalo breeds of Hungary, Romania and the Balkan countries. Previously considered to belong to the Mediterranean buffalo group, it was officially recognised as a breed in 2000, based on its long isolation period and lack of interbreeding with other buffalo breeds.

The improvement of animal performances represents a priority for the Italian dairy Buffalo industry in order to fulfill the increasing market demand for mozzarella cheese. Italian buffalo stock consists of approximately 344,000 Mediterranean river buffaloes (http://faostat.fao.org/). In 2011, the average milk yield per lactation per buffalo cow (35,963 registered cows in the national herd book) was kg 2,223 with 8.49 and 4.65 % of fat and protein content, respectively (http://www.aia.it/aiawebsite/it/home). The milk is almost completely processed into cheese. A breeding program aimed at improving buffalo milk yield and composition is currently operating in Italy, however the low efficiency of the artificial insemination, the difficulties to detect the oestrus and the variability of its length are among the main causes of a very limited impact on the population (Barile, 2005).

Recent advances of molecular genetics offer the possibility to investigate genomic regions that affect traits of economic importance and to identify genetic polymorphisms useful for marker-assisted selection (MAS) programs. In the last decades, several association studies between milk production traits and markers located in milk protein genes have been carried out in cattle, sheep and goat (Ibeagha-Awemu *et al*, 2008; Martin *et al*, 2002; Mroczkowski *et al*, 2004). In particular, the goat αs1 encoding gene (*CSN1S1*) was found to be highly polymorphic with at least 17 alleles associated with qualitative and quantitative differences for the content of αs1casein (Ramunno *et al*, 2004; Ramunno *et al*, 2005), fat (Grosclaude *et al*, 1994; Chilliard *et al*, 2006), urea level (Schmidely *et al*, 2002; Bonanno *et al*, 2007; Avondo *et al*, 2009), fatty acid profile Chilliard *et al*, 2006) and milk yield (Yue *et al*, 2011). Associations between alleles at *CSN1S1 locus* and protein content were also observed for other species as for instance in bovine (Rando *et al*, 1998; Prinzenberg *et al*, 2003; Çardak, 2005) and ovine (Pirisi *et al*, 1999; Wessels *et al*, 2004).

In ruminant species, the *CSN1S1* gene is characterized by an extremely split architecture with 19 exons, many of which (exons 5, 6, 7, 8, 10, 13 and 16) of small size (24 bp) (Ramunno *et al*, 2004; Koczan *et al*, 1991; Calvo *et al*, 2011). In buffalo, the αs1-casein gene codes for a precursor of 214 amino acids with a signal peptide of 15 amino acid residues (Ferranti *et al*, 1998; Sukla *et al*, 2007). Currently, several partial or complete bubaline *CSN1S1* cDNA sequences are available in EMBL (FJ392261; AJ005430; AY948385; EF025981; EF025982; EF025983; DQ111783). The similarity between buffalo and cattle, goat and sheep αs1-casein mRNA sequence is 97.2, 93 and 92.3%, respectively. A similar trend was observed comparing amino acid sequences of these species (Sukla *et al*, 2007).

Few polymorphisms have been reported for the *CSN1S1 locus* in buffalo. The occurrence of a αs1-casein B genetic variant characterized by a single amino acid substitution (p.Leu178Ser) as consequence of single nucleotide substitutions was found for the first time in Romanian Buffalo breed (Balteanu *et al*, 2008) and confirmed at amino acid level in Mediterranean water buffalo (Chianese *et al*, 2009). Furthermore in Indian water buffalo a novel *CNS1S1* allele has been characterized by c.620G>A substitution. It led to a p.Gly192Glu replacement in the peptide chain (Sukla *et al*, 2007).

Recently, additional SNPs were detected: a transition c.136G>A at exon 5, leading to a p.Val31Met substitution, a transition c.175A>G at exon 7, leading to a p.Ile44Val substitution in the peptide chain, one SNP (g.218T>C) in intron 5 and three SNPs (g.472G>C; g.547C>T; g.856T>C) in intron 6. In particular, it was reported that the g.472G>C substitution inactivates the intron 6 splice donor site promoting the skipping of exon 6 of the buffalo *CSN1S1* mRNA and, as consequence, triggering the synthesis of a defective protein lacking eight amino acids (Balteanu *et al*, 2013).

In recent years, many studies have been carried out for the identification of the genetic polymorphisms at the *loci* coding for the buffalo milk proteins (Cosenza *et al*, 2009a, b; Masina *et al*, 2007), and candidate genes responsible for the variation of the quali-quantitative characteristics of the Mediterranean water buffalo milk have been found (Cosenza *et al*, 2007; Pauciullo *et al*, 2010). Efforts in this direction also allowed to find significant associations with traits of economic interest, as milk yield (Pauciullo *et al*, 2012a, b) and milk coagulation properties (Bonfatti *et al*, 2012b).

The aim of our work was to study the variability at the Italian Mediterranean river buffalo *CSN1S1* cDNA and to investigate possible associations with milk yield and composition.

Materials and methods

Sampling

Individual blood samples were collected from 175 Italian river buffaloes randomly chosen and belonging to an experimental herd, located in Salerno province (Southern Italy).

In order to characterize the *CSN1S1* transcripts and to detected polymorphisms at this *locus*, individual milk samples were collected from 10 animals at comparable age, in third calving, at 120 days in milking and free of clinical mastitis and randomly chosen in different farms located in the province of Salerno and Caserta (Southern Italy). After collection, milk samples were immediately frozen and kept at -80 °C until analysis. Sampling was carried out in collaboration with the Italian National Association of Buffalo Breeds (ANASB).

RNA extraction

Total RNA was isolated from somatic cells (SCC range from 10,000 to 12,000/mL) present in individual milk samples by using NucleoSpin® Extract Kits (Macherey-Nagel). A digestion with 2U of DNase I (Ambion) in 1X DNase buffer was carried out according to the manufacture guidelines at 37°C for 30 min followed by the enzyme inactivation at 75°C for 5 min. The quantity, quality, purity, and integrity of RNA after DNase treatment were estimated by means of Nanodrop 2000c spectrophotometer (Thermo Scientific, Barrington, IL) and by electrophoresis on a denaturing agarose gel.

Reverse transcription, PCR and cloning

Total RNA was converted into cDNA by reverse transcription using Improm- II Reverse Transcriptase (Promega) with a final volume of 20µl. The reaction was performed using cDNA19R (5'-CAAAATCTGTTACTGCACA - 3'), a reverse primer complementary to nt 327–345 of 19th exon of buffalo *CSN1S1* cDNA sequence (accession number AY948385).

The PCR was performed using: cDNA19R and cDNA1F (5'- AACCCAGCTTGCTGCTT - 3'), a forward primer corresponding to nt 1–17 of partial 1st exon of buffalo *CSN1S1* cDNA sequence (accession number AY948385). The PCR reaction mix comprised 20 μl of RT reaction product, 50 mM KCl, 10 mM Tris–HCl, 0.1% Triton X-100, 2 mM $MgCl_2$, 10 pmol of each primer, dNTPs each at 0.2 mM, 5 U of Taq DNA Polymerase (Promega, Madison, WI), with a final volume of 100 μl.

The amplification protocol consisted of 39 cycles: the first cycle involved a denaturation step at 97 °C for 2 min, an annealing step at 57 °C for 30 s and an extension step at 72 °C for 1 min and 30 s. The next 37 cycles were performed under the following conditions: 94 °C for 30 s, 57 °C for 30 s. and 72 °C for 1 min and 30 s. In the 39th cycle, the final extension step was carried out at 72 °C for 10 min.

The amplified products were first analyzed by electrophoresis on 3% agarose gel in TBE 1X buffer (Bio-Rad) and then cloned in pCR2.1-TOPO plasmid by using the TOPO TA cloning kit (Invitrogen, Pro, Milan, Italy). White recombinant clones were randomly chosen and screened by PCR using standard vector primers M13. Recombinant clones underwent plasmid purification by PureYield™ Plasmid Midiprep System (Promega, USA) and then sequenced on both strands at CEINGE - Biotecnologie Avanzate (Naples, Italy).

DNA extraction

DNA was extracted from leukocyte, using the procedure described by Gossens and Kan (1981). Briefly, fresh buffy coat samples were washed twice with distilled water and NaCl 1.8% to remove the excess of red cells, protein digestion was carried out with 500 μl of proteinase K solution (2 mg/ml of proteinase K, 1% w/v SDS and 0.02 M EDTA). Proteins were extracted using phenol-chloroform method followed by DNA precipitation with cold isopropanol. The isolated DNA was then resuspended in 100 µl TE buffer pH 7.6 (10 mM Tris, 1mM EDTA). DNA concentration and $OD_{260/280}$ ratio of the samples were then measured by the Nanodrop ND-2000C Spectrophotometer (Thermo Scientific).

CSN1S1 locus genotyping

In order to genotype 175 individual samples of water buffalo for the c628C>T mutation, a method based on ACRS-PCR (Amplification Created Restriction Site PCR) was developed according to Lien *et al* (1992). The ACRS-PCR was performed using: ACRS17F (5'- CAATACCCTGATGCCC***G***AT - 3') as forward and ACRS17R (5'- CACCACAGTGGCATAGTAG - 3') as reverse, corresponding to nt 70–88 and complementary to nt 137–155 of the 17^{th} exon of buffalo *CSN1S1* cDNA sequence (EMBL HE573919), respectively. According to the method, the forward primer was modified by changing C→G in position 17 in order to provide a restriction site for the *Mbo*I (!***G***ATC) endonuclease.

The amplification protocol consisted of 39 cycles: the first cycle involved a denaturation step at 95 °C for 5 min. The next 38 cycles were performed under the following conditions: 95 °C for 45 s, 63.2 °C for 45 s. and 72 °C for 20 s. In the 39^{th} cycle, the extension step was carried out at 72 °C for 10 min. PCR reaction mix comprised: 100 ng of genomic DNA, 50mM KCl, 10mM Tris–HCl, 0.1% Triton X-100, 3mM $MgCl_2$, 5 pmol of each primer, dNTPs each at 400 µM, 2.5U of *Taq* DNA Polymerase (Promega, Madison, WI), and 0.04% BSA, with a final volume of 25 µl.

Digestion and electrophoresis condition

Each PCR amplification product was digested with 10U of *Mbo*I after incubation for 5 h at 37 °C; following the supplier's guidelines. The restricted fragments were analyzed directly by electrophoresis on 3% TBE agarose gel in 1 X TBE buffer and stained with ethidium bromide.

Association study

Associations between *CSN1S1* genotypes and milk production traits were carried out on 4,122 test day records of 503 lactations from 175 buffalo cows, supplied by the Italian Association of Buffalo Breeders

(ANASB). Milk yield (MY), fat (FP) and protein percentages (PP) were tested with the following mixed linear model:

$$y_{ijklmno} = Month_i + Par_j + Sea_K + DIM_l + \alpha_{s1m} + c_n(\alpha_{s1m}) + e_{ijklmno} \qquad [*]$$

where: y_{ijkmno} is the test-day record of MY, FP or PP; $Month_i$ is the fixed effect of the i-th month of production (12 levels); Par_j is the fixed effect of the j-th parity (6 levels: 1 to 5, >5); Sea_K is the fixed effect of the k-th calving season (4 levels: autumn, winter, spring, summer); DIM_l (Days In Milk) is the fixed effect of the l-th stage of lactation (30 levels of 10 days each); α_{s1m} is the fixed effect of the m-th genotype at the c.628C>T SNP of *CSN1S1* gene (3 levels: CC, CT, TT); c_n is the random effect of individual cow (175 levels), nested within α_{s1} genotype; and $e_{ijklmno}$ is the random residual. Pairwise comparisons among different levels of fixed effects included in model were performed using a Bonferroni adjusted test. (Co)variance matrices of random effects of cow and residual were assumed to be diagonal, $\mathbf{I}\sigma^2c$ and $\mathbf{I}\sigma^2e$, respectively. They allow for the REML estimation of variance components associated to individual cow (σ^2_c) and residual (σ^2_e). Variance component associated to the αs1 *locus* ($\sigma^2_{\alpha s1}$) was estimated running a mixed model having the same structure of [*] but with the αs1 genotype treated as random. Contributions of αs1 *locus* ($r^2_{\alpha s1}$) and cow (r^2_c) to the total phenotypic variance of the trait was calculated as the ratio between $\sigma^2_{\alpha s1}$ and σ^2_c , respectively and the sum of all variance components (i.e. $\sigma^2_{\alpha s1} + \sigma^2_c + \sigma^2_e$).

In order to estimate the average of gene substitution effect (α) and a possible dominance effects (d), gene effect was treated as a covariable, represented by the number of T alleles at the αs1 *locus* (0, 1, 2). Finally, an interaction between alleles at the SNP *locus* was considered (Banos *et al*, 2008; Barendse *et al*, 2008).

Results and discussion

a) Characterization of CSN1S1 transcripts

Analysis of the cloned RT-PCR fragments

The mRNAs extracted from individual milk samples obtained from 10 Mediterranean river buffaloes, randomly chosen in the province of Salerno and Caserta, was investigated through the clone analysis. 10 positive clones for each individual were screened. The electrophoretic analysis of the PCR products and the subsequent sequencing of clones, showed at least two populations of transcripts for each examined individual. The most represented population (about 90 %) was correctly assembled, followed by that one deleted of the first triplet of the 11th exon. The last event is a constitutive allele independent event which takes place during the maturation of the pre-mRNA. The first codon of exon 11 (CAG), coding for glutamine in position 78, is in fact competitively eliminated from the mRNA because it is recognized as cryptic site of splice instead of the canonic site (AG) located at the end of the 10th intron. This feature was already observed in buffalo (Ferranti *et al*, 1999) and it is common event for the *CSN1S1* gene of other ruminants, like sheep (Ferranti *et al*, 1998), goat (Ramunno *et al*, 2005) and cattle (Ferranti *et al*, 1999). Furthermore, as already observed in the afore mentioned species, also for the river buffalo it is reasonable to hypothesize the existence of other transcripts different from those reported in this study. Their undetectable amounts opens a new opportunity of investigation in the field of the transcript analysis for buffalo milk protein.

Polymorphism detection

In order to detect polymorphisms at the Mediterranean river buffalo *CSN1S1 locus*, the correctly assembled transcripts were sequenced. The analysis of the sequences showed a transcript of 1083 bp, spanning from the 16th nt of the 1st exon to the 345th nt of the 19th exon. The comparison of the obtained sequences (EMBL HE573919 and HE573920) showed 3 transitions. The first was located at the position 89 of the 17th exon (c.628C>T) and the other two at the position 144 and 239 of the 19th exon, respectively (c.882G>A and c.977A>G). The last two mutations are silent because they are located in the 3 'UTR, whereas the first is a miss-sense SNP (p.Ser178Leu). This amino acid change at the buffalo *CSN1S1 locus* was already observed and it characterizes the genetic variants of the αs1-casein, named B and A, respectively (Balteanu *et al*, 2008; Chianese *et al*, 2009).

Since the presence of the cytosine in position 628 characterizes even other *CSN1S1* sequences of buffalo (FJ392261, AY948385, AJ005430, DQ111783) and of other ruminants, such as goat (AJ504710.2), sheep (NM_001009795), bison (EU862388) and cattle (X59856), its presence might be indicative of an ancestral condition. According to the sequence analysis of the cDNA samples belonging to individuals with informative genotypes, the G in position 882 and the A in position 977 are in *cis* with the T at 89th nucleotide of exon 17. Although the complete genomic sequence of the river buffalo *CSN1S1* is not available, a close distance exists between the exons 17 and 19 in the homologous sequences of the other ruminants. For instance, this DNA region is only 1962 bp long in the bovine *CSN1S1* gene (EMBL acc. No. X59856), therefore, a condition of linkage disequilibrium can be assumed for these SNP in river buffalo.

b) Association of CSN1S1 polymorphism on milk yield and composition

Genotyping of Mediterranean river buffalo CSN1S1 alleles

The c.628C>T does not alter or create any restriction site, therefore, we established a screening method based on the *Mbo*I-ACRS-PCR to identify the carriers of this mutation in an easy and rapid way.

The amplified fragment includes the last 86 bp of 17th exon. Therefore, the digestion with such endonuclease produces an undigested fragment of 86 bp in individuals homozygous for the thymine and two fragments of 70 and 16 bp (not visible electrophoretically) in buffaloes homozygous for the cytosine. The heterozygous individuals produce a pattern characterized by all 3 restriction fragments: 86, 70 and 16 bp (Figure 1). The investigated population was in Hardy–Weinberg equilibrium. Genotype distribution and allele frequencies are reported in Table 1. The frequency of the thymine was 0.33. This result is in agreement with data reported by Chianese *et al* (2009) and Bonfatti *et al* (2012a) and it confirms that the most common genetic variant in Mediterranean breed reared in Campania (Italy) is the αs1-CN B variant.

Association study

All environmental factors included in model [*] affected significantly all the traits considered (Table 2) except the calving season for FP. The genotype at the *CSN1S1 locus* was significantly associated with protein percentage ($p<0.04$). In particular (Table 3), the CC genotype showed an average value of about 0.04% higher than the CT and TT genotypes. The allele substitution effect of a cytosine in a thymine was -0.014±0.014, with a quite low contribution of the *CSNSI locus* to the total phenotypic variance of PP ($r^2_{\alpha s1}$=0.003) PP. A large effect of dominance (-0.028±0.019) was also observed. Often such effect is not detected or considered non relevant because numerically much lower than the additive effect. Although dominance effects are not important in the estimation of breeding values, being not transmitted in the offspring, they might have an impact on allele substitution effect in the population as recently reported (Pauciullo *et al*, 2012a, b).

Recently, Bonfanti *et al* (2012a) estimated the effects of the *CSN1S1* (B and A alleles, c.628C>T transition) -CSN3 (k-casein, X1 and X2 alleles) genotypes on milk production traits and milk coagulation properties in Mediterranean water buffalo. In particular, these authors report that genotypes did not affect milk protein content, but the composite genotype AB-X1X1, compared to genotype AA-X1X1, was associated with an increased fat content and they indicate a role for the casein genes in the variation of the coagulation properties of the buffalo milk. However, the same authors (Bonfatti *et al*, 2012b) reported that the increased proportion of the αs1-CN on the total casein (TCN) content is associated with genotypes carrying the allele *CSN1S1* A. On the contrary, genotypes associated with a marked decrease of the αs1-CN on the TCN (composite genotypes AB-X1X1 and BB-X1X2) are associated with marked increases in the proportion of αs2-CN.

Although results of the present work need to be confirmed with large-scale studies, they might be of great economic interest for the buffalo dairy industry. In fact, increases in average protein content would lead consequently to an expected increase of mozzarella production.

Conclusions

The present study reports a characterization of the Mediterranean water buffalo *CSN1S1* transcripts. At least two populations of transcripts were detected. The most represented population (about 90 %) was correctly assembled, followed by that one deleted of the first triplet of the 11th exon.

Also the study confirms the existence of genetic polymorphisms at this *locus* and it offers a method based on the *Mbo*I-ACRS-PCR for a rapid genotyping at DNA level for the *CSN1S1* A and B alleles. Furthermore, a significant association between the c.628C>T SNP and the protein percentage was found. In particular, the CC genotype showed an average value of about 0.04% higher than the CT and TT genotypes.

Therefore, further studies are necessary to better determine the real effects of the transition c.628C>T on milk composition. Besides, an investigation on larger population are needed in order to validate its application.

In addition, it is necessary to investigate the remaining polymorphisms detected in the 3' un-translated regions (UTR). In fact, it is well known that the sequences in the 3' UTR can affect the mechanism of mRNA regulation, such as de-adenylation and degradation. Therefore, it is reasonable to hypothesize that these observed mutations might influence, directly or indirectly, the gene expression and, consequently, the milk protein composition.

Acknowledgements

This research was financially supported by the Italian Ministry for Agriculture and Forestry Policy-MiPAF (INNOVAGEN project).

References

Ajmone-Masan P, Marco Milanesi M, Bombal, Colli L, Bacciu N, Ramunno L, Cosenza G, Coletta A, Zicarelli L, Fritz E, Koltes J, Reecy JM, Williams J, Iamartino D, Sonstegard TS, Van Tassell C P, The International Buffalo Consortium (2013) Biodiversity among Buffalo genomes. *International Plant and Animal Genome Conference Future Dates, XXI Conference*, January 12-16, - San Diego

Avondo M, Pagano RI, Guastella AM, Criscione A, Di Gloria M, Valenti B, Piccione G, Pennisi P (2009) Diet selection and milk production and composition in Girgentana goats with different αs1-casein genotype. *Journal of Dairy Research* **76**, 202–209. doi: 10.1017/S0022029909003914

Barile VL (2005) Improving reproductive efficiency in female buffaloes. *Livestock Production Science* **92**,183-194. doi:10.1016/j.livprodsci.2004.06.014

Balteanu VA, Carsai TC, Vlaic A (2013) Identification of an intronic regulatory mutation at the buffalo αs1-casein gene that triggers the skipping of exon 6. *Molecular Biology Reports* **40**, 4311–4316. doi: 10.1007/s11033-013-2518-2

Balteanu VA, Pop FD, Vlaic A, Rusu AR (2008) Multiple mutations events are characterizing alphas1 casein Brv and beta casein Crv alleles discovered in Romanian buffalo breed. *Bulletin of University Of Agricultural Sciences And Veterinary Medicine Cluj-Napoca. Animal Science and Biotechnologies* **65**, 478.

Banos G, Woolliams JA, Woodward BW, Forbes AB, Coffey MP (2008) Impact of single nucleotide polymorphisms in leptin, leptin receptor, growth hormone receptor, and diacylglycerol acyltransferase (DGAT1) gene loci on milk production, feed, and body energy traits of UK dairy cows. *Journal of Dairy Science* **91**, 3190-3200. doi: 10.3168/jds.2007-0930

Barendse W, Harrison BE, Bunch RJ, Thomas MB (2008) Variation at the Calpain 3 gene is associated with meat tenderness in zebu and composite breeds of cattle. *BMC Genetics* **9**, 41. doi: 10.1186/1471-2156-9-41

Bonanno A, Finocchiaro R, Di Grigoli A, Sardina MT, Tornambè D, Gigli I (2007) Energy intake effects at pasture on milk production and coagulation properties in Girgentana goats with different αs1-casein genotypes. *Proceedings Symposium of the International Goat Association (IGA) The quality of goats products*, CRA-Uze, Bella (PZ), Italy, p. 70.

Bonfatti V, Giantin M, Gervaso M, Coletta A, Dacasto M, Carnier P (2012a) Effect of CSN1S1-CSN3 (α(S1)-κ-casein) composite genotype on milk production traits and milk coagulation properties in Mediterranean water buffalo. *Journal of Dairy Science* **95(6)**, 3435-3443. doi: 10.3168/jds.2011-4901

Bonfatti V, Giantin M, Gervaso M, Rostellato R, Coletta A, Dacasto M, Carnier P (2012b) CSN1S1-CSN3 (αs1-k-casein) composite genotypes affect detailed milk protein composition of Mediterranean water buffalo. *Journal of Dairy Science* **95**, 6801–6805. doi: 10.3168/jds.2012-5601

Calvo JH, Dervishi E, Sarto P, Berzal B, Serrano M, Joy M (2011) Direct Submission EMBL JN560175.

Çardak AD (2005) Effects of genetic variants in milk protein on yield and composition from Holstein-Friesian and Simmentaler cows. *The South African Journal of Animal Science* **35 (1)**, 41-47.

Chianese L, Quarto M, Pizzolongo F, Calabrese MG, Caira S, Mauriello R, De Pascale S, Addeo F (2009) Occurrence of genetic polymorphism at the alpha(s1)-casein locus in Mediterranean water buffalo milk. *International Dairy Journal* **19 (4)**, 181–189. doi:10.1016/j.idairyj.2008.09.008

Chilliard Y, Rouel J, Leroux C (2006) Goats alpha-s1-casein genotype influences its milk fatty acid composition and delta-9 desaturation ratios. *Animal Feed Science and Technology* **131,** 474–487. doi:10.1016/j.anifeedsci.2006.05.025

Cosenza G, Pauciullo A, Mancusi A, Nicodemo D, Di Palo R, Zicarelli L, Di Berardino D, Ramunno L (2007) Mediterranean river buffalo oxytocin-neurophysin I (OXT) gene: structure, promoter analysis and allele detection. *Italian Journal of Animal Science* **6 (Suppl.2)**, 303-306. doi: 10.4081/ijas.2007.s2.303

Cosenza G, Feligini M, Mancusi A, D'Avino A, Coletta A, Di Berardino D, Ramunno L (2009a) Italian Mediterranean river buffalo CSN2 gene structure and promoter analysis. *Italian Journal of Animal Science* **8 (Suppl.2)**, 57-59. doi: 10.4081/ijas.2009.s2.57

Cosenza G, Pauciullo A, Feligini M, Coletta A, Colimoro L, Di Berardino D, Ramunno L (2009b) A point mutation in the splice donor site of intron 7 in the αs2-casein encoding gene of the Mediterranean River buffalo results in an allele-specific exon skipping. *Animal Genetics* **40 (5)**, 791. doi: 10.1111/j.1365-2052.2009.01897.x.

Ferranti P, Chianese L, Malorni A, Migliaccio F, Stingo V, Addeo F (1998) Copresence of deleted protein species generates structural heterogeneity of ovine alpha(s1)-casein. *Journal of Agricultural and Food Chemistry* **46 (2)**, 411-416. doi: 10.1021/jf970609s

Ferranti P, Scaloni A, Caira S, Chianese L, Malorni A, Addeo F (1998) The primary structure of water buffalo alpha(s1)- and beta-casein identification of phosphorylation sites and characterization of a novel beta-casein variant. *Journal of Protein Chemistry* **17 (8)**, 835-844.

Ferranti P, Lilla S, Chianese L, Addeo F **(**1999). Alternative nonallelic deletion is constitutive of ruminant alpha(s1)-casein. *Journal of Protein Chemistry* **18 (5)**, 595-602.

Gossens M, Kan YW (1981). DNA analysis in the diagnosis of hemoglobin disorders. *Methods in Enzymology* **76**, 805-817.

Grosclaude F, Ricordeau G, Martin P, Remeuf F, Vassal L, Bouillon J **(**1994) From gene to cheese: caprine alpha s1 casein polymorphism, its effects and its evolution. *INRA Productions Animales* **7 (1)**, 3–19.

Ibeagha-Awemu EM, Kgwatalala P, Zhao X (2008) A critical analysis of production-associated DNA polymorphisms in the genes of cattle, goat, sheep and pig. *Mammalian Genome* **19 (9)**, 591-617. doi: 10.1007/s00335-008-9141-x

Koczan D, Hobom G, Seyfert HM **(**1991) Genomic organization of the bovine alpha-S1 casein gene. *Nucleic Acids Research* **19 (20)**, 5591-5596.

Lien S, Alestrom, P., Klungland H., Rogne S. (1992) Detection of multiple beta-casein (CASB) alleles by amplification created restriction sites (ACRS). *Animal Genetics* **23 (4)**, 333-338. doi:10.1111/j.1365-2052.1992.tb00155.x

Martin P, Szymanowska M, Zwierzchowski L, Leroux C (2002). The impact of genetic polymorphisms on the protein composition of ruminant milks. *Reproduction Nutrition Development* **42**, 433-459. doi: 10.1051/rnd:2002036

Masina P, Rando A, Di Gregorio P, Cosenza G, Mancusi A (2007). Water buffalo kappa-casein gene sequence. *Italian Journal of Animal Science* **6 (Suppl.2),** 353-355. doi: 10.4081/ijas.2007.s2.353

Mroczkowski S, Korman K, Erhardt G, Piwczyński D, Borys B (2004) Sheep milk protein polymorphism and its effect on milk performance of Polish Merino. *Archiv Tierzucht*, Dummerstorf **47**, 114-121.

Pauciullo A, Cosenza G, D’Avino A, Colimoro L, Nicodemo D, Coletta A, Feligini M, Marchitelli C, Di Berardino D, Ramunno L (2010) Sequence analysis and genetic variability of stearoyl CoA desaturase (SCD) gene in the Italian Mediterranean river buffalo. *Molecular and Cellular Probes* **24 (6)**, 407-410. doi: 10.1016/j.mcp.2010.07.009

Pauciullo A, Cosenza G, Steri R, Coletta A, Jemma L, Feligini M, Di Berardino D, Macciotta NPP, Ramunno L (2012a) An association between Single Nucleotide Polymorphisms at the Oxytocin locus and milk yield in Italian Mediterranean river buffalo. *Journal of Dairy Research* **79 (2)**, 150-156. doi: 10.1017/S0022029911000914

Pauciullo A, Cosenza G, Steri R, Coletta A, La Battaglia A, Di Berardino D, Macciotta NPP, Ramunno L (2012b) A single nucleotide polymorphism (SNP) in the promoter region of River buffalo Stearoyl CoA Desaturase gene (SCD) is associated with milk yield. *Journal of Dairy Research* **79 (4)**, 429-435. doi: 10.1017/S0022029912000507

Pirisi A, Piredda G, Papoff CM, Di Salvo R, Pintus S, Garro G, Ferranti P, Chianese L (1999). Effects of sheep alpha s1-casein CC, CD and DD genotypes on milk composition and cheesemaking properties. *Journal of Dairy Research* **66 (3)**, 409-419.

Prinzenberg EM, Weimann C, Brandt H, Bennewitz J, Kalm E, Schwerin M, Erhardt G (2003) Polymorphism of the Bovine *CSN1S1* Promoter: Linkage Mapping, Intragenic Haplotypes, and Effects on Milk Production Traits. *Journal of Dairy Science* **86**, 2696–2705. doi:10.3168/jds.S0022-0302(03)73865-X

Rando A, Di Gregorio P, Ramunno L, Mariani P, Fiorella A, Senese C, Marletta D, Masina P (1998) Characterization of the CSN1AG allele of the bovine αs1-casein locus by the insertion of a relict of a long interspersed element. *Journal of Dairy Science* **81 (6)**, 1735–1742. doi:10.3168/jds.S0022-0302(98)75741-8

Ramunno L, Cosenza G, Rando A, Illario R, Gallo D, Di Berardino D, Masina P (2004) The goat αs1-casein gene: gene structure and promoter analysis. *Gene* **334**, 105-111 doi:10.1016/j.gene.2004.03.006

Ramunno L, Cosenza G, Rando A, Pauciullo A, Illario R, Gallo D, Di Berardino D, Masina P (2005) Comparative analysis of gene sequence of goat *CSN1S1* F and N alleles and characterization of *CSN1S1* transcript variants in mammary gland. *Gene* **345 (2)**, 289-299. doi:10.1016/j.gene.2004.12.003

Schmidely Ph, Meschy F, Tessier J, Sauvant D (2002) Lactation response and nitrogen, calcium, and phosphorus utilization of dairy goats differing by the genotype for alpha S1-casein in milk, and fed diets varying in crude protein concentration. *Journal of Dairy Science* **85 (9)**, 2299–2307. doi:10.3168/jds.S0022-0302(02)74310-5

Sukla S, Bhattacharya TK, Venkatachalapathy RT, Kumar P, Sharma A (2007). Molecular cloning and characterization of buffalo alpha(s1)-casein gene. DNA Seq. **18 (5)**, 334-340.

Yue XP, Zhang XM, Wang W, Ma RN, Deng CJ, Lan XY, Chen H, Li F, Xu XR, Ma Y, Lei CZ (2011) The *CSN1S1* N and F alleles identified by PCR-SSCP and their associations with milk yield and composition in Chinese dairy goats. *Molecular Biology Reports* **38 (4)**, 2821–2825. doi: 10.1007/s11033-010-0428-0

Wessels G, Hamann H, Erhardt G, Distl O (2004) Genotype effects of milk protein polymorphisms on milk production in East Friesian dairy sheep. Berliner Munchener Tierarztl Wochenschr **117 (9-10)**, 414–419 [Article in German]

Table 1. Genotyping data and allele frequency of the c.628C>T SNP at the *CSN1S1* gene in the Italian Mediterranean river buffalo population.

	Genotypes				**Allelic frequency**	
	CC	**CT**	**TT**	**Total**	**C**	**T**
Observed	75	84	16	**175**	0.67	0.33
Expected	78.22	77.55	19.22			

$\chi 2=1.20$- $p\leq 0.05$ - d.o.f.=1

Table 2. Statistical significance of factors included in model

	P-value		
Effect	Milk yield	Fat	Protein
Genotype	0.63	0.93	0.04
Parity	<.001	<.001	<.001
Month of production	<.001	<.001	<.001
DIM	<.001	0.02	<.001
Season	<.001	0.48	0.02

Table 3. Least squares means of milk yield (kg/d), fat and protein percentage (%) for the genotypes at the *locus* c.628C>T of river buffalo *CSN1S1* gene estimated with model [*].

Genotype	Animals n.	Milk yield (kg/d)	Fat (%)	Protein (%)
CC	75	7.81±0.17	9.22±0.12	4.72±0.016[a]
CT	84	7.92±0.17	9.27±0.12	4.68±0.016[b]
TT	16	7.56±0.36	9.20±0.23	4.69±0.029[ab]

[a,b] Means within columns with different superscripts differ (P=0.038)

Table 4. Substitution effect of a cytosine with a thymine at the c.628C>T SNP in the *CSN1S1* gene (mean ± SE) and contribution of the αs1 polymorphism to the phenotypic variance for protein percentage.

Statistic	Value
α	-0.014±0.014
d	-0.028±0.019
$\sigma^2_{\alpha s1}$	0.0004
σ^2_c	0.0034
σ^2_e	0.1596
$r^2_{\alpha s1}$	0.003
r^2_c	0.021

α: Substitution effect;
d: dominance effect;
σ^2: variance components associated to the genotype (αs1); to the individual buffalo cow (c), to residuals (e);
r^2: contributions of genotype (αs1) and of individual buffalo cow (c) to the total phenotypic variance.

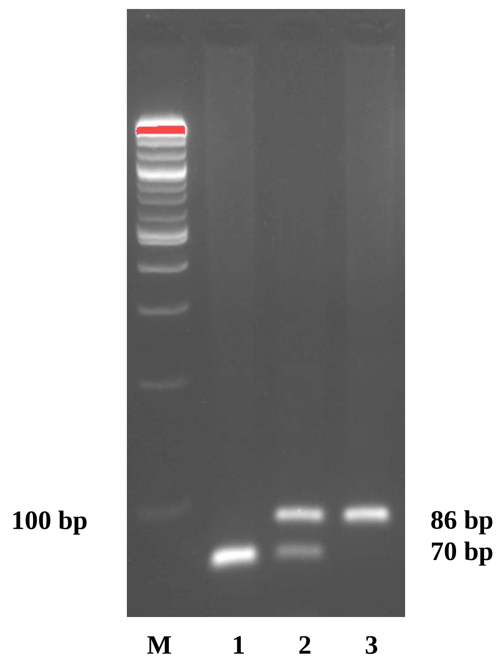

Figure 1. Observed genotypes after *Mbo*I digestion of fragments obtained by ACRS-PCR of a DNA region corresponding to the last 86 bp of the 17th exon into Mediterranean river buffalo *CSN1S1* gene. M=2-Log DNA ladder (0.1-10.0kb) (New England Biolabs); lane 1: *CSN1S1* C/C; lane 2: *CSN1S1* C/T; lane 3 *CSN1S1* T/T