Italian Mediterranean river buffalo CSN2 gene structure and promoter analysis

Gianfranco Cosenza, Maria Feligini, Andrea Mancusi, Annalisa D'Avino, Angelo Coletta, Dino Di Berardino & Luigi Ramunno

Published online: 07 Mar 2016.

Submit your article to this journal ⊘

Article views: 26

View related articles ⊘

Italian Mediterranean river buffalo *CSN2* gene structure and promoter analysis

Gianfranco Cosenza[1], Maria Feligini[2], Andrea Mancusi[1],
Annalisa D'Avino[1], Angelo Coletta[3], Dino Di Berardino[1],
Luigi Ramunno[1]

[1]Dipartimento di Scienze del Suolo, della Pianta, dell'Ambiente e delle Produzioni Animali,
Università di Napoli "Federico II", Italy

[2] Istituto Sperimentale Italiano Lazzaro Spallanzani, Lodi, Italy

[3]Associazione Nazionale Allevatori Specie Bufalina (ANASB), Caserta, Italy

Corresponding author: Luigi Ramunno. Dipartimento di Scienze del Suolo, della Pianta, dell'Ambiente e delle Produzioni Animali, Università di Napoli "Federico II". Via Università 133, 80055 Portici (NA), Italy - Tel. +39 081 2539004 - Fax: +39 081 7762886 - Email: ramunno@unina.it

ABSTRACT - The nucleotide sequence of the whole buffalo β-casein encoding gene (*CSN2*) plus 1,476 nt at the 5' flanking region and 51 nt at the 3' flanking region was determined. The gene is spread over 10.2 kb and consists of 9 exons varying in length from 24 (exon 5) to 498 bp (exon 7) and 8 introns from 92 bp (intron 5) to 2259 bp (intron 1). Furthermore, highly conserved sequences, mainly located in the 5' flanking region, were found between this gene and the β-casein encoding genes of other species. The comparison between the obtained promoter and exonic regions and buffalo sequences present in EMBL evidenced different polymorphic sites. Finally, 5 interspersed repeated elements (4 in the bovine *CSN2* gene) were also identified at 3 different locations of the sequenced region: 5' untranscribed region, intron 1, and intron 4.

Key words: **Mediterranean river buffalo,** β-casein, CSN2 gene, Nucleotide sequence.

Introduction - In ruminants the four caseins (αs1, β, αs2, and κ) represent about 80% of milk proteins. Of these, the β-casein **represents the most abundant casein fraction (about 40%) (**Grosclaude *et al*., 1987). The calcium-sensitive casein encoding genes (*CSN1S1*, *CSN2*, *CSN1S2)* show similar gene structures, with several small exons and a low exon/intron ratio. **At present, the complete sequences** of *CSN2* gene in the bovine (Bonsing *et al*., 1988, EMBL no. **M55158), ovine (**Provot *et al*., 1995, EMBL no. X79703), and caprine (Cosenza *et al*., 2005, EMBL no. **AJ011018) species are available in** EMBL. On the contrary, for the buffalo specie only cDNA sequences (EMBL no. AJ005432, DQ317447, DQ631829, DQ191170, DQ191171, DQ191172, AY599833, AJ005165) and promoter sequence (EMBL no. AJ005165) are available. **In this paper, we report the complete nucleotide sequence of the Mediter-**ranean river buffalo β-casein encoding gene (*CSN2*) and the analysis of its promoter.

Material and methods - Genomic DNA of two Mediterranean river buffaloes was extracted from leukocytes obtained from blood samples. Primers for amplification and sequencing were designed by means of DNASIS-Pro software (Hitachi), using the sequence of the buffalo *CSN2* cDNA (Das *et al*., 2000) and the complete sequence of the bovine *CSN2* gene (Bonsing *et al*., 1988) as templates. Additional primers, designed on newly determined intron sequences, were also used for sequencing. A typical 50 µl reaction mix comprised: 100 ng of genomic DNA, 50 mM KCl, 10 mM Tris-HCl (pH 9.0), 0.1% Triton X-100, 3 mM MgCl$_2$, 200 nmol of each primer, dNTPs each at 400 µM, 2.5 U of *Taq* DNA Polymerase (Promega, Madison, WI), 0.04% BSA. The amplification programs consisted of 31 cycles. The first one characterised by a denaturation at 97°C for 2 min, a primers annealing at 46-62 °C for 45 sec, and an extension step at 72°C for 2 min. The next 30 cycles involved a denaturation step at 94°C for 45 sec, annealing at 46-62 °C for 45

sec, and extension at 72°C for 2 min. In the last cycle, the extension time was 10 min long. All the amplified fragments were sequenced in both directions. The transcription binding site prediction was carried out by AliBaba2 program: http://www.gene-regulation.com/pub/programs/alibaba2/index.html

Results and conclusions - *Structure of buffalo CSN2 gene* - We sequenced the whole *CSN2* gene plus 1,472 nt at the 5' flanking region and 51 nt at the 3' flanking region from two Mediterranean river buffaloes. The buffalo *CSN2* gene extends over 10.2 kb including 1.09 kb of exonic regions and 8.06 kb of intronic regions with a total similarity with the corresponding bovine sequence of about 96%. The main feature of the buffalo *CSN2* gene is its extremely simple architecture. It contains 9 exons ranging in size from 24 (exon 5) to 498 bp (exon 7) and 8 introns from 92 bp (intron 5) to 2259 bp (intron 1). Exon 1 contains the first 44 bp of the 5'-UTR. Exon 2 encodes the remaining 12 bp of the UTR, the entire signal peptide (45 bp), and the first two codon of the mature protein. Exon 7, that is the longest exon in the gene, encodes for about 82% of the mature protein, while exon 8 contains the last codon of mature protein, the stop codon (TAA, nt 4-6), and the first 36 bp of the 3'-UTR. Exon 9 comprises the remaining 322 bp of the 3'-UTR. All splice junctions follow the 5' GT/ 3'AG splice rule. The first 30 nt of the 3' flanking regions are well conserved among species, indicating that they are involved in the mechanism of the 3' end processing of the primary transcript. The comparison between the two obtained sequences of the buffalo *CSN2* gene showed only a SNP (transition C→T) at nt 274 of exon 9, which took place near the polyadenilation site (nt 306-311) and identifies a silent allele at the *CSN2 locus* that we named *CSN2* A1 (EMBL no. FM946182). By comparing our exons sequences with the published sequence of the buffalo *CSN2* cDNAs we found 13 mutations, 3 of which are non conservative: T→A (Ile→Asn), G→C (Lys→Asn), and T→C (Met→Thr) in position 20, 78, and 179 of exon 7, respectively (Table 1). *Buffalo CSN2 promoter analysis* - We determined the nucleotide sequence of the *CSN2* promoter (1472 bp) (EMBL no. FM986648). Comparative analysis of the proximal 5'-flanking region of *CSN2* was performed for bovine (X14711), caprine (AF409096), ovine (X79703), rabbit (X15735), human (AF027807), rat (M10936), murine (X13484), porcine (AY452035), and our newly sequenced buffalo *CSN2*. Within the analyzed region different congruent and putative binding sites for transcription factors were found to be highly conserved among species: the TATA Box (ATATAT) (-34/-29), 6 GATA1, 6 nuclear factor

Figure 1.   Promoter sequence of the buffalo CSN2 gene. Numbering is relative to the first nucleotide of exon 1 (+1). Congruent and putative factors are underlined, in shaded bold letters and in italics.

1 (NF1), 2 signal transducer and activator of transcription 5 (STAT5) highly conserved in all *CSN2* promoters, 2 GR elements/progesterone receptor elements (GREs/PREs), 3 CCAAT/enhancer-binding protein-β (C/EBPβs), 2 Yin and Yang factor 1 (YY1), 1 activating protein-1 (AP-1), 6 simian virus 40 promoter factor 1 (Sp1), 1 CAATBOX, and 13 octamer factor 1 (OCT1) (Figure 1). The comparison between the obtained promoter sequence and that of the same gene of the buffalo reared in India (EMBL no. AY352050) evidenced 25 mutations. Of these, 4 (C→T, T→A, A→G, and G→C in position -1406, -1402, -956, -700, respectively) might be of particular interest as they would create 3 putative extra sites of regulation of *CSN2* gene expression in the Italian Mediterranean Buffalo: 1 OCT1 and 1 GATA1 (shaded in Figure 1). In addition, a polymorphic microsatellite sequence has been detected between nt 256 and 265 (AT^5, EMBL no. FM986648, *vs.* AT^6 EMBL no. AY352050). *Artiodactyla Retroposons* - The analysis of the buffalo *CSN2* gene evidenced a higher exon/intron size ratio than that observed in cattle as a consequence of a **non–long terminal repeat** retrotransposons (NLRs) (**truncated L1_BT line) located in intron 1 (nt 839-1097) of buffalo** *CSN2* gene that, at present, has been observed exclusively in the bovine species. Like the homologous gene in cattle, the buffalo *CSN2* gene is characterized by four DNA elements showing similarity to artiodactyla retroposon. **In** particular, the first is located in the distal region of the promoter (from -546 to -675) and appears to be an half Bov-B SINE**,** whereas the other three are all located in the intron 4, showing strong similarities with the full-length Bov-A2 and **Bov-t SINE, respectively (Lenstra** *et al*., 1993).

Table 1.   Mutations detected by comparing the sequence of the exons of the gene we sequenced* with the published sequence of the full and partial buffalo *CSN2* cDNAs**. Numbering is relative to the first nt of the corresponding exon.

| Exon | EMBL FM946182* | | EMBL** | aa change | Exon | EMBL FM946182* | | EMBL** | aa change |
	Location	Mutation				Location	Mutation		
1	29 nt	A→C	AY352050	-	7	327 nt	G→T	AY599833; DQ191171	-
5	6 nt	T→C	DQ317447	-	7	348 nt	C→T	AY599833	-
7	20 nt	T→A	AY599833	Ile→Asn	7	434 nt	T→C	AY599833	-
7	66 nt	G→A	AY599833	-	7	503 nt	C→T	DQ191171	-
7	78 nt	G→C	DQ191172	Lys→Asn	7	518 nt	T→C	AY599833; DQ191170-71-72	-
7	99 nt	G→T	AY599833	-	9	274 nt	G→T	AJ005165; DQ317447	-
7	179 nt	T→C	AJ005432	Met→Thr					

REFERENCES - **Bonsing,** J., Ring, J.M., Stewart, A.F., Mackinlay, A.G., 1988. Complete nucleotide sequence of the bovine β-casein gene. **Aust. J. Biol. Sci. 41:527-537. Cosenza,** G., Pauciullo, A., Gallo, D., Di Berardino, D., Ramunno, L., 2005. A *Ssp*I PCR-RFLP detecting a silent allele at the goat *CSN2 locus*. **J. Dairy Res**. 72:456-459. **Das,** P., Tiwari, G., Jain, S., Garg, L.C., 2000. Nucleotide sequence of the river buffalo beta-casein cDNA. J. Anim. Sci. 78:1390. **Grosclaude,** F., Mahé, M.F., Brignon, G., Di Stasio, L., Jeunet, R., 1987. A Mendelian polymorphism underlying quantitative variations of goat αs1-casein. **Genet. Sel. Evol. 19:399-412. Lenstra,** J.A., Van Boxtel, A.F., Zwacgstra, K.A., Schwerin, M., 1993. Short interspersed nuclear element (SINE) sequences of the Bovidae. Anim. Genet. 24:33-39. **Provot,** C, Persuy, M.A, Mercier, J.C., 1995. Complete sequence of the ovine beta-casein-encoding gene and interspecies comparison. Gene. 154:259-263.