

MOLECULAR CHARACTERIZATION OF BUFFALO  $\alpha_{s1}$ -CASEIN GENE

**Biplab Kumar Pramanik<sup>1</sup>, Subhasis Batabyal<sup>1</sup>, Apratim Maity<sup>1,\*</sup>,  
Sachinandan De<sup>2</sup>, Saibal Chattopadhyay<sup>1</sup> and Abhijit Barui<sup>1</sup>**

Received: 28 March 2018

Accepted: 19 September 2022

**ABSTRACT**

The caseins (Alpha s1, Alpha s2, beta and kappa) comprise the major protein component of ruminant milk and are secreted in the form of stable calcium phosphate micelles. The present study was proposed to characterize  $\alpha_{s1}$ -casein gene in buffalo at the molecular level to determine complete  $\alpha_{s1}$ -casein cDNA sequence. Sequencing results were analyzed and aligned with available mRNA GenBank sequences of other domestic animals (Accession AAB34797, NM\_001009795, X72221, AY344966, NM\_181029, AY948385 and BC109618). The complete amino acid coding region of the buffalo  $\alpha_{s1}$ -casein was deduced from cloned cDNA generated from mammary tissue mRNA. Amino acid sequence comprises of 199 amino acid residues. When aligned with previously reported buffalo sequence (Accession-AY948385) it showed one amino acid (Glutamine) deletion at amino acid position 78.

**Keywords:** *Bubalus bubalis*, buffaloes, alpha s1-Casein, mRNA sequence, amino acid sequence

**INTRODUCTION**

The high-yielding Murrah is the Holstein-Friesian of the buffalo world. The protein content of milk is mainly divided in two components namely casein and whey protein. A very small fraction constitutes the milk fat globule proteins. But the main fraction is casein, constituting about 80% of the milk protein. The main fractions of casein are  $\alpha_{s1}$ ,  $\alpha_{s2}$ ,  $\beta$  and  $\kappa$ -casein. These proteins are expressed from the casein gene which is a 250 to 350 kb region, depending on species. This gene have been mapped in chromosome 6 of cattle and sheep whereas in human and buffalo it is situated in chromosome 4 and chromosome 7 respectively (Goldammer *et al.*, 2007). But in each case the casein genes are arranged in an order of  $\alpha_{s1}$ - $\alpha_{s2}$ - $\beta$ - $\kappa$  (Rijnkels, 2002).

Higher milk yield and therefore higher protein yield over the lactation has been correlated with the  $\alpha_{s1}$ -CN BB phenotype (Aleandri *et al.*, 1990; Sang *et al.*, 1994). However, the same phenotype has also been associated with lower protein concentration in milk (Ng-Kwai-Hang *et al.*, 1987).

<sup>1</sup>Faculty of Veterinary and Animal Sciences, West Bengal University of Animal and Fishery Sciences, Kolkata, India, \*E-mail: toapu77@gmail.com

<sup>2</sup>Animal Genomics Laboratory, ABTC, National Dairy Research Institute, Karnal, India

Casein appears in the milk as stable calcium phosphate micelles. The association of milk protein genotype with composition and properties of milk can be exploited as an additional criterion in selecting bulls for artificial insemination. For example, cheese yield can be enhanced by increasing the frequency of a milk protein genotype associated with increased milk yield and protein concentration. The present study was proposed to characterize  $\alpha_{s1}$ -casein gene in buffalo at the molecular level to determine complete  $\alpha_{s1}$ -casein mRNA sequence.

## MATERIALS AND METHODS

The work was carried out in the Department of Veterinary Biochemistry, Faculty of Veterinary and Animal Sciences, West Bengal University of Animal and Fishery Sciences and Animal Genomics Laboratory, Animal Biotechnology Center, National Dairy Research Institute.

### RNA isolation from mammary tissue

Five buffalo mammary tissues (1 to 2 g) were collected from New Delhi Municipal Corporation slaughterhouse immediately after the slaughter of the animal. Mammary tissues were collected from animals having milk in the mammary gland. The tissues were washed with DEPC treated water (0.1%) and stored in 2 ml RNA later solution (Ambion, USA) and brought to lab in an ice box to prevent any RNA degradation and were immediately processed for RNA isolation. From mammary tissue, the total RNA was isolated using *TRIzol* reagent (Invitrogen, Brazil) as per protocol. Finally, the RNA was precipitated by adding 0.5 ml isopropanol in aqueous phase and it was dissolved in RNase free distilled water.

The purity and integrity of the RNA was verified (Sambrook and Russell, 2001).

### Synthesis of cDNA

Synthesis of cDNA from RNA was performed using Superscript III Reverse Transcriptase kit (Invitrogen, Brazil) as per protocol provided with the kit.

### Amplification of buffalo alpha $S_1$ -casein coding sequence

Cattle alpha  $S_1$ -casein mRNA sequence (GenBank Accession: BC109618) was used to design gene specific primers for amplification of buffalo alpha  $S_1$ -casein complete coding region. Forward primer (CSN1S1CF, 5'-CTTGCTGCTTCTTCCCAGTC-3') was located from 24<sup>th</sup> to 43<sup>rd</sup> position. Reverse primer (CSN1S1CR, 5'-CCTTTGACTCAGTGGCCTTT-3') was located from 990<sup>th</sup> to 971<sup>st</sup> position in the complimentary strand. Expected PCR product size was ~967.

PCR amplification of buffalo  $\alpha_{s1}$ -Casein complete coding fragment was done as per standard protocol (94°C for 35 seconds, 55°C for 30 seconds, 72°C for 30 seconds, 30 cycles). After PCR cycling, 5  $\mu$ l PCR product was checked in 2% agarose gel electrophoresis. The size of the products was compared with 100 bp DNA ladder (Fermentas UAB, Lithuania).

### Cloning of PCR product

PCR product was cloned in pGEM-T cloning vector (PCR product cloning kit, Promega). PCR products were purified before cloning. The ligation reaction was carried out as per the protocol described by suppliers.

### Transformation

A vial of XL1-Blue Super competent cells

(Agilent Technologies), 200  $\mu$ l aliquots, was thawed on ice. Ten  $\mu$ l of ligation mixture (plasmid DNA) was added and mixed by tapping with finger and incubated in ice for 30 minutes. Heat Shock was given at 42°C temperature for 90 seconds in a water bath. Immediately after heat shock, cold shock was given by transferring the tube in ice for 1 minute. Now 600  $\mu$ l of SOB media (Super Optimal Broth) and 100  $\mu$ l of 1 M glucose was added and incubated at 37°C in an incubator shaker (Lab therm-Kuhner -Switzerland) for 45 minutes with shaking. Finally, all the transformed cells were spread over LB agar plate containing Ampicillin (100  $\mu$ g/ml) and plates were incubated at 37°C overnight.

Recombinant clones were screened by Colony PCR. Positive clones were grown separately in LB broth containing ampicillin (20  $\mu$ g/ml) at 37°C overnight in an incubator shaker (Lab therm-Kuhner -Switzerland) with gentle shaking for purification of plasmid DNA.

### Recombinant plasmid isolation

QIAprep Spin Miniprep (Cat. No. 27106) plasmid purification kit (Qiagen GmbH, Hilden, Germany) were used to isolate recombinant plasmids from *E. coli* transformants. The eluted plasmid DNA was examined by Agarose (1%) gel electrophoresis for integrity and quantified spectrophotometrically. Isolated recombinant plasmids were custom sequenced using the T<sub>7</sub>F and SP<sub>6</sub>R primers in an automated ABI 377 sequencer (Perkin Elmer Applied Biosystem, Foster city, CA, USA) by Bangalore Genei Privet limited (Bangalore, India). The nucleotide sequence data was edited and analysed by BioEdit v7.0.9 (Hall, 1999).

## RESULTS AND DISCUSSIONS

### Amplification of buffalo alpha- s1 casein coding sequence

Total RNA was isolated from five mammary tissue samples and cDNA were synthesized. Primers were designed for amplification of buffalo alpha-s1 casein coding sequence from cDNA. The size of amplified PCR product was 967 bp.

### Cloning and sequencing of buffalo $\alpha_{s1}$ - casein coding region

Buffalo  $\alpha_{s1}$ - casein coding region amplicons were cloned in cloning vector. Recombinant clones were screened by colony PCR (Figure 1) and later confirmed by plasmid-DNA-PCR for presence of insert. Positive clones were grown separately in LB media containing ampicillin (100  $\mu$ g/ml) for purification of plasmid DNA.

Recombinant plasmid DNA was purified and checked in agarose (1%) gel eletrophoresis. Purified plasmids were sent for sequencing in commercial laboratory. The nucleotide sequence data was edited and analysed by BioEdit v7.0.9 (Hall, 1999). The sequence data acquired was blasted ([www.ncbi.nlm.nih.gov/BLAST](http://www.ncbi.nlm.nih.gov/BLAST)) to determine that the sequence was of  $\alpha_{s1}$ -casein gene. Obtained sequence and deduced amino acid sequence along with various positions were shown in Figure 2.

The entire sequence data of bovine and caprine  $\alpha_{s1}$ -casein gene has been generated (Koczan *et al.*, 1991). The bovine sequence extends over 17.508 kb with 1.138 kb exonic and 16.370 kb intronic portions. It splits into 19 exons ranging in length from 24 to 385 bp and 18 introns stretching in size from 90 to 1967 bp. The caprine  $\alpha_{s1}$ -casein sequence was almost completely homologous with its bovine equivalent. The  $\alpha_{s1}$ -casein coding region

length is of 645 bp for both cattle and buffalo. The size of the  $\alpha_{s1}$ -casein cDNA ORF is 621 bp in sheep and goat but, the length of the same in camel and human are 669 bp and 558 bp respectively. The 645 bp coding region of buffalo  $\alpha_{s1}$ -casein cDNA could be decoded into 214 amino acids. The signal peptide is shaped by first 15 amino acids whereas the remaining amino acids constitute the mature peptide (Sukla *et al.*, 2007).

The obtained amino acid sequence of buffalo  $\alpha_{s1}$ -casein was of 199 amino acid residues which correlate with Farrell *et al.*, 2004. Alignment of multiple  $\alpha_{s1}$ -casein amino acid sequences [Clustax programme version 2.0.11 (Larkin *et al.*, 2007) and GeneDoc programme version 2.7.000 (Nicholas and Nicholas, 1997)] of other domestic animals with that of obtained sequence and one reported amino acid sequence of buffalo (Accession no. AY-948385) clearly shows one mutation at position 78 (from glutamine to lysine) (Figure 3) which correlates with the findings of Ferranti *et al.*, 1999.

This amino acid deletion may be due to a codon skipping which is likely due to an erroneous 3' cryptic splice site usage when exons 10 and 11 are spliced. Same phenomenon has also been found at the protein level in the other ruminant species like caprine (Ferranti *et al.*, 1997) and ovine (Ferranti *et al.*, 1995). The sequence variation and its consequent effect on the physiological systems could be determined by further study.

Construction of sequence dependent phylogenetic tree of mature amino acid sequences of cattle, buffalo, sheep, goat, horse, mouse, and human using FigTree v1.2.2 programme (Rambaut, 2008) indicates buffalo sequences are different. Separate clads were noticeable for buffalo, cattle, goat, sheep, horse, and mouse in the phylogenetic tree constructed from the data of the amino acid sequences (Figure 4). Human formed distinct cluster elucidating a separate location in the cladogram.

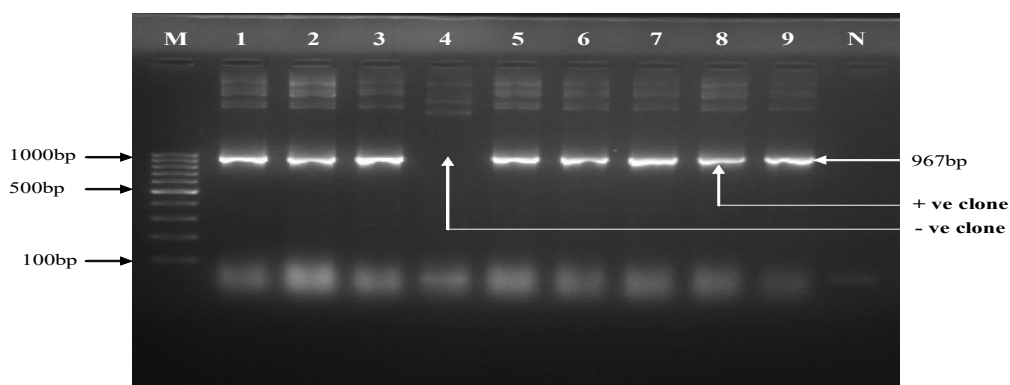


Figure 1. Agarose (2%) gel electrophoresis of colony PCR amplicon of  $\alpha_{s1}$ -casein mRNA sequence of 967 bp.

Lane 1-3 and 5-9: amplification of 967 bp of  $\alpha_{s1}$ -casein mRNA sequence.

Lane 4: Negative colony without PCR amplification.

Lane M: DNA molecular weight marker (GeneRular 100 bp DNA ladder).

Lane N: PCR without template (negative control).

1	CTTGCTGCTT	CTTCCCAGTC	TTGGGTCAA	GATCTTGACA	ACCATGAAAC		
					M K L	Frame 2	
51	TTCTCATCCT	TACCTGTCTT	GTGGCTGTTG	CTCTTGCCAG	GCCTAACAG		
	L I L	T C L	V A V A	L A R	P K Q	Frame 2	
101	CCTATCAAGC	ACCAAGGACT	CCCTCAAGGA	GTCCTCAATG	AAAATTTACT		
	P I K H	Q G L	P Q G	V L N E	N L L	Frame 2	
151	CAGGTTTTTT	GTGGCGCCTT	TTCCAGAAGT	GTTTGAAAG	GAGAAGGTCA		
	R F F	V A P F	P E V	F G K	E K V N	Frame 2	
201	ATGAAGTGG	CACGGATATT	GGGAGTGAAT	CAACTGAGGA	TCAAGCCATG		
	E L S	T D I	G S E S	T E D	Q A M	Frame 2	
251	GAAGATATTA	AGCAAATGGA	AGCTGAAAGC	ATTTGCTCAA	GTGAGGAAAT		
	E D I K	Q M E	A E S	I S S S	E E I	Frame 2	
301	TGTTCCCAT	AGTGTGAGA	AGCACATTCA	AAAGGAAGAT	GTGCCCTCTG		
	V P I	S V E K	H I Q	K E D	V P S E	Frame 2	
351	AGCGTTACCT	GGGTTATCTG	GAACAGCTTC	TTAGACTGAA	AAAATACAAC		
	R Y L	G Y L	E Q L L	R L K	K Y N	Frame 2	
401	GTACCCAGC	TGAAATGTG	TCCCAATTTG	GCTGAGGAAC	AACTTCACAG		
	V P Q L	E I V	P N L	A E E Q	L H S	Frame 2	
451	TATGAAAGAG	GGAATCCATG	CCCAACAGAA	AGAACCTATG	ATAGGAGTGA		
	M K E	G I H A	Q Q K	E P M	I G V N	Frame 2	
501	ATCAGGAACT	GGCCTACTTC	TACCCTCAGC	TTTTCAGACA	ATTCTACCAG		
	Q E L	A Y F	Y P Q L	F R Q	F Y Q	Frame 2	
551	CTGGACGCCT	ATCCATCTGG	TGCCTGGTAT	TACGTTCCAC	TAGGCACGCA		
	L D A Y	P S G	A W Y	Y V P L	G T Q	Frame 2	
601	ATACCCTGAT	GCCCCATCAT	TCTCTGACAT	CCCTAATCCC	ATCGGCTCTG		
	Y P D	A P S F	S D I	P N P	I G S E	Frame 2	
651	AGAACAGTGG	AAAGACTACT	ATGCCACTGT	GGTGAAGAGT	CAAGTGAATT		
	N S G	K T T	M P L W	*		Frame 2	
701	CTGAGGGACT	CCACAATTAT	GGTCTTTGAT	GGTCTGAAA	ATTCCATGCT		
751	CTACACGTCT	TTTCATCTAT	CATGTCAAAC	CATTCTATCC	AAAGGCTTCA		
801	ACTGCTGTTT	TAGAATAGGG	CAATCTCAA	TTGAAGGCAC	TCTTTCTTCT		
851	TGAGTTCTCT	ACTGTATTTT	AGATGGTGTA	ACATCCTTAA	GTGAAATTGT		
901	CCTAACAGCT	TGTTACCTAA	ATTCCAGTAG	TATCATGCTG	GTATAAAGGC		
951	CACTGAGTCA	AAGGA					

Figure 2. Nucleotide sequence of buffalo alpha S1 casein gene and deduced amino acid sequence.

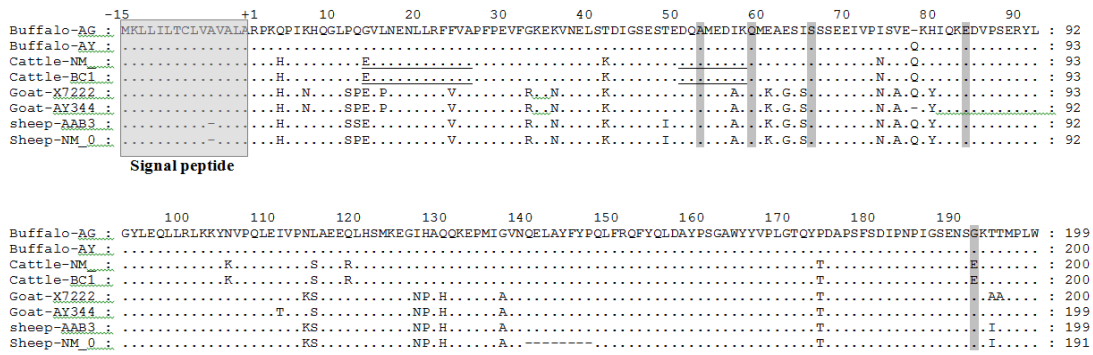


Figure 3. Multiple amino acid sequence alignment of  $\alpha_{s1}$ -casein in domestic animals. Underlined amino acids from 14 to 26 are deleted in Bovine alpha S1 casein ‘A’ variants and 51 to 58 are deleted in Bovine alpha S1 casein ‘H’ variants.

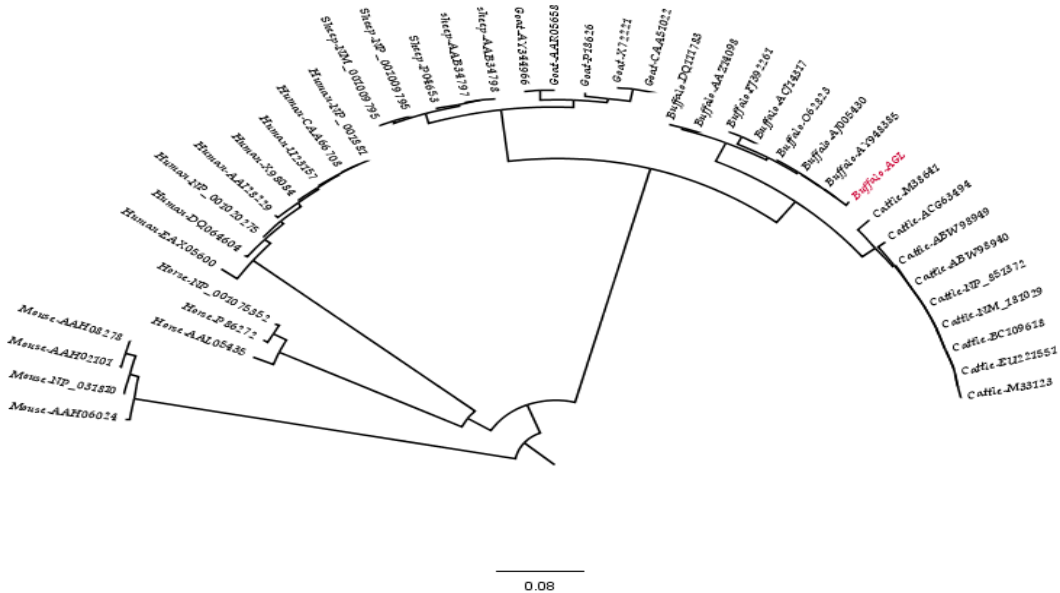


Figure 4. Phylogenetic tree constructed from amino acid sequence data of  $\alpha_{s1}$ -casein gene. The evolutionary relationship between different species of animals as well as human has been presented in this tree.

## REFERENCES

- Aleandri, R., L.G. Buttazzoni, J.C. Schneider, A. Caroli and R. Davoli. 1990. The effects of milk protein polymorphisms on milk components and cheese-producing ability. *J. Dairy Sci.*, **73**(2): 241-255. DOI: 10.3168/jds.S0022-0302(90)78667-5
- Farrell, H.M. Jr., R. Jimenez-Flores, G.T. Bleck, E.M. Brown, J.E. Butler, L.K. Creamer, C.L. Hicks, C.M. Hollar, K.F. Ng-Kwai-Hang and H.E. Swaisgood. 2004. Nomenclature of the proteins of cows' milk-sixth revision. *J. Dairy Sci.*, **87**(6): 1641-1674. DOI: 10.3168/jds.S0022-0302(04)73319-6
- Ferranti, P., F. Addeo, A. Malorni, L. Chianese, C. Leroux and P. Martin. 1997. Differential splicing of pre-messenger RNA produces multiple forms of mature caprine alpha(s1)-casein. *Eur. J. Biochem.*, **249**(1): 1-7. DOI: 10.1111/j.1432-1033.1997.t01-5-00001.x
- Ferranti, P., S. Lilla, L. Chianese and F. Addeo. 1999. Alternative nonallelic deletion is constitutive of ruminant  $\alpha_{s1}$ -casein. *J. Protein Chem.*, **18**: 595-602. DOI: 10.1023/a:1020659518748
- Ferranti, P., A. Malorni, G. Nitti, P. Laezza, R. Pizzano, L. Chianese and F. Addeo. 1995. Primary structure of ovine  $\alpha_{s1}$ -casein: Localization of phosphorylation sites and characterization of genetic variants variants A, C and D. *J. Dairy Res.*, **62**: 281-296. DOI: 10.1017/S0022029900030983
- Goldamer, T., R. Weikard, M.N. Miziara, R.M. Brunner, R. Agrawala, A.A. Schaffer, J.E. Womack and M.E.J. Amaral. 2007. A radiation hybrid map of river buffalo (*Bubalus bubalis*) chromosome 7 and comparative mapping to the cattle and human genomes. *Cytogenet Genome Res.*, **119**(3-4): 235-241. DOI: 10.1159/000112067
- Hall, T.A. 1999. BioEdit: A user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucl. Acids. S.*, **41**: 95-98.
- Koczan, D., G. Hobom and H.M. Seyfert. 1991. Genomic organization of the bovine alpha-S1 casein gene: *Nucl. Acids. S.*, **19**(20): 5591-5596. DOI: 10.1093/nar/19.20.5591
- Larkin, M.A., G. Blackshields, N.P. Brown, R. Chenna, P.A. McGettigan, H. McWilliam, F. Valentin, I.M. Wallace, A. Wilm, R. Lopez, J.D. Thompson, T.J. Gibson and D.G. Higgins. 2007. Clustal W and Clustal X version 2.0. *Bioinformatics*, **23**(21): 2947-2948. DOI: 10.1093/bioinformatics/btm404a
- Ng-kwai-Hang, K.F., J.F. Hayes, J.E. Moxley and H.G. Monardes. 1987. Variation in milk protein concentrations associated with genetic polymorphism and environmental factors. *J. Dairy Sci.*, **70**(3): 563-570. DOI: 10.3168/jds.S0022-0302(87)80042-5
- Nicholas, K.B. and H.B. Nicholas Jr. 1997. *GeneDoc: A Tool for Editing and Annotating Multiple Sequence Alignments*. Available on: <https://www.psc.edu/biomed/genedoc/>
- Rambaut, A. 2008. *FigTree*. Institute of Evolutionary Biology, University of Edinburgh, Edinburgh. Available on: <http://tree.bio.ed.ac.uk/software/figtree>
- Rijnkels, M. 2002. Multispecies comparison of the casein gene loci and evolution of casein gene family. *J. Mammary. Gland Biol.*, **7**(3): 327-345. DOI: 10.1023/a:1022808918013
- Sambrook, J. and D. Russell. 2001. *Molecular Cloning. A Laboratory Manual*, 3<sup>rd</sup> ed. Cold Spring Harbor Press, New York, USA.



- Sang, B.C., B.S. Ahn, B.D. Sang, Y.Y. Cho and A. Djajanegara. 1994. Association of genetic variants of milk proteins with lactation traits in Holstein cows, p. 217-211. *In 7<sup>th</sup> Asian Australasian Animal Production Animal Science Congress Proceedings*, Bali, Indonesia.
- Sukla, S., T.K. Bhattacharya, R.T. Venkatachalapathy, P. Kumar and A. Sharma. 2007. Molecular cloning and characterization of buffalo  $\alpha_{s1}$ -casein gene. *DNA Seq.*, **18**(5): 334-340. DOI: 10.1080/10425170701388974