B-3468
[1-4]

Indian J. Anim. Res.,
Print ISSN:0367-6722 / Online ISSN:0976-0555

AGRICULTURAL RESEARCH COMMUNICATION CENTRE
www.arccjournals.com/www.ijaronline.in

Molecular characterization and genetic variability of Alpha Casein gene, CSN1S1 in Bikaneri camel (*Camelus dromedarius*) milk

Karan Veer Singh*, S. Jayakumar, S.P. Dixit and Z.S. Malik[1]

National Bureau of Animal Genetic Resources,
GT Road Bye Pass, Karnal-132 001, Haryana, India
Received: 21-07-2017 Accepted: 16-08-2017 DOI: 10.18805/ijar.B-3468

ABSTRACT

Camel milk is an important protein source for the nomadic communities living in the arid lands of the world, In recent years there has been an increase in consumption of non-bovine milk as an alternative protein source for humans. Camel milk seems to be containing larger amount of total proteins, such as lactoferrin and immunoglobulins as compared to the cow milk, which may be responsible for the better antimicrobial properties. The casein fraction of milk proteins consists of four caseins, namely αs1-casein, αs2-casein, β-casein, and κ-casein. Casein genetic polymorphisms are important due to their effects on quantitative traits and technological properties of milk. This work was designed to study occurrence of polymorphism of α-casein in native Bikaneri camel (*Camelus dromedarius*) raw milk sample and to characterize these variants on molecular level.

Key words: Alpha- casein, Casein gene polymorphism, Camel milk production traits, *Camelus dromedarius.*

INTRODUCTION

Camels belong to the family of Camelidae, were domesticated around 3000–6000 years ago (Sushma *et al.*, 2014). It has two species, double-humped camel (*Camelus bactrianus*, 70 chromosomes) and single-humped camel (*Camelus dromedaries*, 74 chromosomes). The single-humped camel inhabits dry, desert habitat of Africa, Arab and west Central Asia while the double-humped inhabits eastern Central Asia and China (Cui *et al*. 2007). In recent years there has been an increasing interest in consumption of non-bovine milk as an alternative protein source for humans, and is being promoted as healthy food due to its therapeutic value such as antioxidant activity. Camel milk is an important protein source (Kappeler *et al.*, 1998), especially for the people living in the arid lands of the world, where other sources of proteins are scarce (Konuspayeva *et al.*, 2009). In composition camel milk is more similar to goat milk and contains less short-chain fatty acids than cow, sheep and buffalo milks, and it contains about 3 times greater vitamin-C than cow milk.

It is worth noting that, camel milk seems to be containing larger amount of total proteins, such as lactoferrin and immunoglobulins as compared to the cow milk, which may be responsible for the better antimicrobial properties (Farah, 1993). Moreover unlike cow milk, camel milk is reported to have antidiabetic (Pauciullo *et al.*, 2012) and anti-hypertensive effects. Camel milk fat lacks β-lactoglobulin and is rich in immunoglobulins which are compatible with human milk (Shamsia, 2009). Studies in India have revealed that populations who regularly consume camel milk have zero incidences of type 1 diabetics as compared to 5.5 percent in other communities that do not consume this milk (Rachagani *et al.*, 2008). One of the biggest hurdles in camel milk processing is its incompatibility with the Ultra High Temperature (UHT) exposure, which is present day dairy industries use to preserve milk.

The milk composition of camel has been studied previously by many authors. Total protein content ranges from 2.4 to 5.3 percent (Konuspayeva *et al.*, 2009; Nikkah, 2011a, 2011b) and is divided into caseins (CN ~80%) and whey proteins. (Giambra *et al.*, 2013; Ereifej *et al*., 2011; Ikonen *et al*., 2008; Medrano *et al*., 1990). There are four main types of milk caseins, i.e. αs1-casein, αs2-casein, β-casein, and κ -casein, encoded by four genes, CSN1S1, CSN1S2, CSN2 and CSN3, respectively. In camel milk α-casein (22percent) is the second main fraction after β-casein (65percent) (Kappeler *et al.*, 1998; El Agamy *et al*., 2006). Polymorphism of the blood groups, enzymes and milk proteins could be used as biological tools to improve the genetic merit in animal breeding studies. Casein genetic polymorphisms are important and well-known due to their effects on quantitative traits and technological properties of milk (Frajman *et al.*, 2004). The objective of this study was to investigate genetic and phylogenetic analysis of Bikaneri camel's milk casein gene.

MATERIALS AND METHODS

Sample collection and DNA isolation: Milk and blood samples from camel (*C. dromedarius*) were collected from Bikaner region, Rajasthan. Milk samples 10mL was collected

*Corresponding author's e-mail & address: karan_veer@yahoo.com
[1]Lala Lajpat Rai University of Veterinary and Animal Sciences, Hissar, Haryana.

with Sodium azide and stored at -4^{0}C. Blood samples (10 mL) were obtained and stored at 20^{0}C temperature until DNA extraction. The genomic DNA was isolated from the samples using the standard protocol (Shashikanth, 1999). For mRNA extraction, milk somatic cells, from fresh camel milk samples were gained by centrifugation (10 min, 2200 X g), and washed twice with phosphate buffered saline/0.5 M EDTA according to Boutinaud *et al.,* (2002). Invisorb Spin RNA Mini Kit (Invitek GmbH, Germany) was used for extraction of total RNA.

PCR Amplification and purification: To amplify targeted DNA sequences by polymerase chain reaction (PCR), Primer3 v.0.4.0 (Rozen and Skaletsky, 2000) was used to design five set of primer pairs using camel α-casein mRNA-sequence (GenBank ID: AJ012628-AJ012630) as a preliminary template. The PCR amplification was carried out in a volume of 25 µl, containing 100 ng genomic DNA, 0.5 µl m of each primer, 0.3 µl of *Taq* DNA polymerase (Fermentas), 0.5 µl MgCl2, and 0.5 µl of dNTPs. Amplification was carried out in a GeneAmp9600 thermal cycler (PerkinElmer) employing the following conditions: 94C for 2 min; 30 cycles at 94C for 1 min, X C for 1 min, and 72C for 1 min; and final extension at 72C for 5 min. The amplified PCR product was assessed by Agarose gel electrophoresis using 1%EE Agarose using 1X TAE buffer and visualized under UV light.

Sequencing reaction and sequence analyses: PCR products were sequenced directly using 50 ng (2.0 ul) of PCR product and 4 pM (1.0 ul) of primer, 4 ul of BigDye Terminator ready reaction kit (Perkin Elmer), and 3.0 ul of double distilled water to adjust the volume to 10.0 ul. Cycle sequencing was carried out in a GeneAmp9600 thermal cycler (Perkin Elmer) employing 30 cycles at 96 ^{o}C for 10 s, 50 ^{o}C for 5 s, and 60 ^{o}C for 4 min. Products were purified by alcohol precipitation followed by 70% alcohol washing. Purified samples were dissolved in 10 ul of 50% Hi-Di formamide and analyzed in an ABI3700 automated DNA Analyzer (Perkin Elmer). Resulting chromatograms were compared by alignment using Chromas software ver2.6.4 (Technelysium Pty. Ltd., Tewantin, Queensland, Australia).

Statistical analyses: Sequences used for analysis were downloaded from NCBI, aligned in the ClustalX package (Thompson *et al.,* 1994), Nucleotide BLAST program at NCBI was used for sequence homology searches in public databases. The Sequence data from amplified gene product were edited manually using Chromas Ver.2.6.4. Multiple sequence alignments were performed using MegAlign tool of LASERGENE software (DNA STAR, Inc., Madison, WI, USA). The GENEDOC package was used for formatting the sequences to make them compatible with the desired software. Allele and genotype frequencies were calculated using PopGene program v. 1.31 (Yeh *et al.,* 1999). To test the distribution of genotypes on the base of Hardy-Weinberg, chi-square test was performed. For determination of phylogenetic relationship of casein among different mammalian species, nucleotide sequences from the GenBank and Ensemble databases were used.

RESULTS AND DISCUSSION

Sequences were edited and initially aligned using SEQUENCHER and then optimally aligned visually. Multiple sequences downloaded from NCBI were aligned by ClustalX package. In the present study direct sequencing of the complete coding region of camel samples revealed full sequence similarity to α-casein A of Kappeler *et al.* (1998) who reported the occurrence of two variants (A & B) in camel αS1- casein, it did not revealed any variability as reported earlier. The sequencing revealed that the size of alpha (s1)-casein cDNA was of 1087 bp with GC content of 42.5 percent. These results were comparable with other submitted sequences in NCBI for α-casein gene (1094 bp). The gene coded 222 amino acids precursor with a signal peptide of 15 amino acid residues, Coding sequences were translated to amino acids using the EBI online translation tool. On the whole, the camel sequence shares a similar organization with the bovine counterpart, with some differences (AY380228). Genetic polymorphism at α-casein locus the estimated minimum values of x^2 (0.2082) evidenced for a good relatedness between empirically obtained and theoretically expected genotypes population. The observed H indicates the representativeness of the samples, despite of their limited size as well as it emphasizes that current selection has not eliminated (suspended) any of the alternative forms (alleles) of the α-casein gene.

Milk protein composition traits are associated with protein genetic variants. Camel α-casein may exist as several variants. Erhardt (2016) and workers have worked on casein gene polymorphism in milk samples of camels from different regions of Sudan (Africa) by isoelectric focusing, where three protein patterns named α-casein A, C, and D were identified. The major allele A revealed frequencies of 0.79 (Lahaoi), 0.75 (Shanbali), 0.90 (Arabi Khali), and 0.88 (Arabi Gharbawi) in the different ecotypes. α-casein C variant shows a single G > T nucleotide substitution in the exon 5, leading to a non-synonymous amino acid exchange (p.Glu30 > Asp30, GenBank ID: JF429138) in comparison to α-casein A and D. At cDNA level, no further single nucleotide polymorphisms could be identified in α-casein A, C, and D. The variants A and C are characterized by missing of exon 18 as compared to CSN1S1* B, due to DNA insertion of 11 bp at intron 17 which alter the pre-mRNA spliceosome machinery, multispecies alignment of Martin *et al*. (2003) already confirms. Alpha α-casein C (Shuiep *et al.,* 2013) and α-casein B (Kappeler *et al.,* 1998) are both characterized by p.Glu30 > Asp30 in the deduced mature protein sequence, and are only differing by missing or presence of the eight amino acids encoded by exon 16. It is often skipped during the processing of mRNA, as it is known in sheep (Ferranti *et al.,* 1995; 1998).

Table 1: Estimated A/T and G/C nucleotide content of the α-casein gene in different species.

Species	Accession no	T/A	G/C
Cattle	X00564	59.13	40.87
Water buffalo	AJ005430	59.02	40.98
Sheep	X03237	59.01	40.99
Goat	X59836	58.18	41.82
Pig	X54973	59.09	40.91
Camel	AJ012628	57.5	42.50
Human	NM-001890	60.45	39.55
Yak	AF194983	67.99	32.01

Beside exon skipping as a reason for casein allele, simultaneous occurrence of skipped and non-skipped forms of the same CSN1S1 allele are usual in sheep, goat, cattle, pig, and human. However, the characteristics of camel α-casein have not been well-documented. Further research and detailed DNA sequencing studies are needed to clarify complete sequence differences between α-casein *A, B, and C in detail, probably the occurrence of this variant is characteristic of other *C. dromedarius* ecotypes.

It is well known that the variation in mRNA and protein is primarily due to alternative splicing, duplication, and insertion/deletion events in addition to nucleotide mutations. Alternative splicing as a reason for the presence of minor and major fractions of camel milk protein is already described for camel lactophorin, a major whey protein in camel milk by Kappeler *et al.,* 1999. About 75 percent of the protein is expressed as a long variant and the minor fraction (25 percent) of camel lactophorin is a shortened variant, characterized by the lack of exon 2. Long and short variants of α S1-casein also occur in ovine milk as a result of differential splicing of the heterogeneous nuclear RNA (Ferranti P. *et al.,* 1995), as well as it was also showed in goat (Ramunno L. *et al.,* 2005; Leroux C. *et al.,* 1992) and cattle where, for instance, the skipping of the exon 4 results in the A variant (Mohr U. *et al.,* 1994).

The database search of sequences for a possible match to the DNA sequence of casein gene was conducted using the BLAST (NCBI). In the phylogenetic trees, constructed from the sequences of the alpha (S1)-casein mRNA as well as protein sequences, it has been observed that camel, buffalo, cattle, goat and sheep formed a cluster with a closer relationship between cattle and buffalo followed by goat and sheep based on genetic distance (F_{st} values) using MEGA5 (Tamura *et al.,* 2011). Comparing gene sequence with other reported sequence in this region showed estimated homology as 94, 94, 94, 82, 94, 97 and 96 percent, respectively with that of cattle, goat, sheep, pig, buffalo, equine and human. A similar trend was observed when compared with amino acid sequences of these species.

However, this could be a starting point for further characterization of genetic diversity in camel milk proteins. Milk protein variability would help in studying association concerning milk performance traits in camel. Moreover, at phenotypic level, milk protein could be used to increase economic value of milk, for screening of breeds and population for phenotyping animals for breed characterization.

CONCLUSION

Camel milk has developed a high reputation as a healthy nutrition with most of its therapeutic value ascribed to its biological properties such as antioxidant activity. Casein genes have been deeply investigated in ruminants, whereas little information is available in camel, the present study can be a starting point for further characterization of genetic diversity in camel milk proteins. The genetic polymorphism of milk proteins can be studied and used further for better diversity, genetic structure preservation assessment of populations and the relationship with their dairy production traits.

ACKNOWLEDGEMENT

The authors thank the camel owners in Bikaner region for providing milk and blood samples. We also thank Mr Ramesh Rana, technical officer, ICAR-NBAGR who assisted in collection of samples and analysis. The facilities and support provided by the ICAR-NDRI for milk sample analysis is greatly acknowledged.

REFERENCES

Boutinaud, M., Rulquin, H., Keisler, D. H., Djiane, J., and Jammes, H. (2002). Use of somatic cells from goat milk for dynamic studies of gene expression in the mammary gland. *Journal of Animal Science*, **80:** 1258-1269.

Cui P., Ji R., Ding F., Qi D., Gao H. and Meng H. (2007). A complete mitochondrial genome sequence of the wild two-humped camel (*Camelus bactrianus ferus*): an evolutionary history of camelidae. *BMC Genomics*. **8:** 241-245.

El Agamy, E. I. (2006). Camel milk. In Y.W. Park, and F.W. Haenlein (*Eds.*), Handbook of Non-bovine Mammals Blackwell Publisher. Iowa, NJ, USA: (pp. 297-344).

Ereifej, K. I., Alu'datt, M. H., AlKhalidy, H. A., Alli, I., and Rababah, T. (2011). Comparison and characterisation of fat and protein composition for camel milk from eight Jordanian locations. *Food Chemistry*, **127**: 282-289.

Erhardt, Georg., Shuiep, El., Lisson, Maria., Weimann, Christina., Wang, Zhaoxin., El Zubeir, Ibtisam., Pauciullo, Alfredo. (2016). Alpha S1-casein polymorphisms in camel (*Camelus dromedarius*) and descriptions of biological active peptides and allergenic epitopes. *Tropical Animal Health and Production*, **48**: 879-887. 9p.

Farah, Z. (1993). Composition and characteristics of camel milk. *J. Dairy. Res.* **60**: 603-626.

Ferranti, P., Chianese, L., Malorni, A., Migliaccio, F., Stingo, V., & Addeo, F. (1998). Copresence of deleted protein species generates structural heterogeneity of ovine as1-casein. Journal of Agricultural and Food Chemistry, 46, 411-416.

Ferranti, P., Malorni, A., Nitti, G., Laezza, P., Pizzano, R., Chianese, L. (1995). Primary structure of ovine ás1-caseins: localization of phosphorylation sites and characterization of genetic variants A, C and D. *J Dairy Res*.;**62**: 281–96.

Frajman, P., and Dovc, P. (2004) Milk production in the post-genomic era. *Acta agriculturae slovenica*, **84**: 109-119.

Giambra, I.J., Zubeir, E.l., and Erhardt, G. (2013). Biochemical and molecular characterization of polymorphisms of ás1-casein in Sudanese camel (*Camelus dromedarius*) milk. *Int. Dairy J*. **28**: 88-93.

Ikonen, T., Ojala M., and Syväoja, E.L. (2008). Effects of composite casein and beta-lactoglobulin genotypes on renneting properties and composition of bovine milk by assuming an animal model. *J. Dairy Sci*. **71**: 188-195.

Kappeler, S., Farah, Z., and Puhan, Z. (1998). Sequence analysis of *Camelus dromedaries* milk caseins. *Journal of Dairy Research*, **65:** 209-222.

Kappeler, S., Farah, Z., and Puhan, Z. (1999). Alternative splicing of lactophorin mRNA from lactating mammary gland of the camel (*Camelus dromedarius*). *Journal of Dairy Science*, **82:** 2084-2093.

Konuspayeva, G., Faye, B., and Loiseau, G. (2009). The composition of camel milk: a meta-analysis of the literature data. *Journal of Food Composition and Analysis*, **22:** 95-101.

Leroux, C., Mazure, N., Martin, P., (1992). Mutations away from splice site recognition sequences might cis-modulate alternative splicing of goat as1-casein transcripts. Structural organization of the relevant gene. *J Biol Chem.*; **267:** 6147–57.

Martin, P., Ferranti, P., Leroux, C., and Addeo, F. (2003). Non-bovine caseins: quantitative variability and molecular diversity. In [P. F. Fox, & O. L. H. McSweeney (Eds.], Advanced dairy chemistry. Proteins, Vol. I (pp. 277e317). New York, NY, USA: Kluwer Academic/Plenum Publisher.

Medrano, J.F., and Cordova, E.A. (1990). Genotyping of bovine kappa-casein loci following DNA sequence amplification. *Biotechnology*. **8:** 144-146.

Mohr, U., Koczan, D., Linder, D., Hobom, G., and Erhardt, G. (1994). A single point mutation results in A allele-specific exon skipping in the bovine alpha s1-casein mRNA. *Gene*.;143:187–92

Nikkah, A. (2011a). Equidae, camel, and yak milks as functional foods: a review. *Journal of Nutrition and Food Sciences*, **1:** 100-116.

Nikkah, A. (2011b). Science of camel and yak milks: human nutrition and health perspectives. *Food and Nutrition Sciences*, **2:** 667-673.

Pauciullo, A., Shuiep, E., Cosenza, G., Ramunno, L. and Erhardt, G. (2012). Molecular characterization and genetic variability at ê-casein gene (CSN3) in camels. *Gene*. **15:** 22-30.

Rachagani, S., and Gupta, I.D., (2008). Bovine kappa-casein gene polymorphism and its association with milk production traits. *Genet. Mol. Biol*. **13:** 893-897.

Ramunno, L., Cosenza, G., Rando, A., Pauciullo, A., Illario, R., Gallo, D., (2005). Comparative analysis of gene sequence of goat CSN1S1 F and N alleles and characterization of CSN1S1 transcript variants in mammary gland. *Gene*.; **345:**289–299.

Rozen, S., Skaletsky, H. (2000). Primer3 on the WWW for general users and for biologist programmers. *Methods Mol Biol*. **132:**365-86.

Shamsia (2009). Nutritional and therapeutic properties of camel and human milks. International Journal of Genetics and Molecular Biology,**1:**52:58.

Shashikanth, P.B., (1999). Study on DNA polymorphism in cattle and buffalo. Ph.D. Thesis, NDRI Deemed University, Karnal, India.

Shuiep E. S., Giambra I. J., El Zubeir I.E. M., Erhardt G. (2013). Biochemical and molecular characterization of polymorphisms of ás1-casein in Sudanese camel (Camelus dromedarius) milk. International Dairy Journal **28** : 88-93.

Sushma Prasad, Sharique A. Ali1, P. Banerjee, Jyoti Joshi, Upasna Sharma and R. K. Vijh (2014).Identification of SNPs and their validation in camel (Camelus bactrianus and Camelus dromedarius). *Journal of Agriculture and Veterinary Science* **7**: 2 (2)65-70.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., & Kumar, S. (2011). MEGA5: Molecular Evolutionary Genetics Analysis Using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods. Molecular Biology and Evolution, **28**(10), 2731–2739.

Thompson, J. D., Higgins, D. G., & Gibson, T. J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Research, 22(22), 4673–4680.

Yeh, F.C., Yang, R.C., Boyle, T., (1999). POPGENE 32. Version 1.31. Population genetics software. Available from: http://www.ualberta .ca/~fyeh/fyeh/.