

The Study of Casein Alpha S1 (CSN 1S1) Gene Diversity in Ruminant Animals

Faith, E. A^{1*}, Yakubu, A², Sabuwa, A. B. M²

¹Department of Animal Science, College of Agriculture Lafia, Nasarawa State, P.M.B. 33, Nigeria

²Department of Animal Science, Faculty of Agriculture, Nasarawa State University, Keffi, Shabu-Lafia Campus, P.M.B. 135 Lafia, Nigeria

*Corresponding author:

Faith, E. A

Received: 01.03.2019

Accepted: 11.03.2019

Published: 31.03.2019

Abstract: This study was undertaken to explain the diversity of casein alpha S1 (CSN 1S1) gene in ruminant animals. The neighbour joining tree inferred the genetic distances among and between species of the CSN1S1 sequences. The tree showed a closer genetic distance between the sequences of *Capra hircus* and *Ovis aries*, *Bubalus bubalis* and *Bos taurus*, *Capra hircus* and *Bubalus bubalis* as well as *Bos taurus* and *Capra hircus*. UPGMA tree derived from the consensus sequences, indicates that *Bos taurus* sequences were more closer to *Bubalus bubalis* sequences, compared to the sequence of *Ovis aries* and *Capra hircus* respectively. The results of the substitution pattern were more homogenous between the sequence of *Capra hircus* and *Ovis aries* (0.00) and *Capra hircus* and *Bubalus bubalis* (0.00) followed by the sequence of *Bos taurus* and *Capra hircus* (0.01) whereas the lowest homogeneity substitution pattern was observed between the sequence of *Bos taurus* and *Bubalus bubalis* (0.31). The highest evolutionary divergence was observed between the sequence of *Ovis aries* and *Bubalus bubalis* and the evolutionary divergence appeared to be species wise. The results indicates that all the CSN1S1 sequences of the selected ruminant species were not significant and the test implies that the strict-neutrality for synonymous (dN) and non-synonymous (dS) substitution from the sequences used were neutral. This findings would aid in understanding of species genetic variation and similarities resulting from their sequences.

Keywords: Study, CSN 1S1 gene, sequences, divergence, neutrality, ruminant.

INTRODUCTION

The mammary gland is a complex organ both in structure and function. After pubertal mammogenesis it undergoes three physiological transitions during a lactation cycle: from involution to colostrogenesis, to lactation, then back to involution.

Proteins present in milk have different classes that are related to the type of polypeptide chains, with casein, beta-lactoglobulin, and alpha-lacto-albumin among them.

Caseins (α_{s1} , β , α_{s2} , and κ) in ruminants are coded by single-copy genes, physically linked in a region of approximately 300 kb, in the following order: *CSN1S1* (α_{s1} -Cn), *CSN2* (β -Cn), *CSN1S2* (α_{s2} -Cn), and *CSN3* (κ -Cn) [1].

Among caseins, the α_{s1} -Cn represents more than 40% in bovine milk, while in goat milk, it ranges from 0 to 25% due to the occurrence of polymorphism in the gene that codes for this protein [2]. As reviewed by Neveu *et al.*, [3], at least 18 alleles of the *CSN1S1*

gene have already been detected in goat breeds, and classified into four expression levels of 3.6, 1.6, 0.6, and 0 g/L/allele, and named as "high", "intermediate", "low" and "null" alleles. The protein synthesized through the expression of this gene consists of 199 amino acid residues. A, B, C, and E variants differ only in amino acid substitutions, while D and F variants result from the deletion of 11 and 37 amino acids, respectively [4]. Allele, characterized by the null production of α_{s1} -Cn in milk, is the result of the deletion of approximately 8 kilobases (kb) in the 3' region of this gene [5].

As reviewed by Moioli *et al.*, [6], the protein content in goat milk is affected by polymorphism in this gene, resulting in significant cheese yield differences. Consequently, due to economic implications, the effect of polymorphism in the *CSN1S1 locus* has been investigated, and allelic frequencies have been determined in several countries.

β -casein has about 13 genetic variants, with the A1 and A2 variants being frequently found in dairy

Quick Response Code



Journal homepage:

<http://crosscurrentpublisher.com/ccijavs/>

Copyright @ 2019: This is an open-access article distributed under the terms of the Creative Commons Attribution license which permits unrestricted use, distribution, and reproduction in any medium for non commercial use (Non Commercial, or CC-BY-NC) provided the original author and source are credited.

cattle [2, 7]. The existing difference between the A1 and A2 variants of β -casein is caused by mutation of the A2 allele to the A1 allele at position 67 (histidine A1 and proline A2), being important in the variation of protein content and in the composition of milk protein [8].

Casein contains proteins that transform into opiate compounds called β -casomorphins (BCM). BCM binds to the A1 allele of the β -casein and it is believed that an ingestion of milk containing presence of this allele causes allergy and other diseases in the human body. This study was aimed to explain the diversity of casein alpha S1 (CSN 1S1) gene in ruminant animals.

MATERIALS AND METHODS

Sequences sources, alignment, translation and comparison

A total of sixteen (16) (CSN1S1) sequences, four each from four selected ruminant animals as thus: *Bos taurus* (cattle), *Ovis aries* (sheep), *Capra hircus* (goat) and *Bubalus bubalis* (buffalo) were extracted from the Genbank of the National Center for Biotechnology Information (NCBI) with a registered accession number as EU221571.1, EU221570.1, EU221569.1, EU221568.1 (*Bos taurus*), XM_027970719, XM_012179357.3, FJ695513.1, FJ440846.1 (*Ovis aries*), AJ439077.1, AJ439076.1, NM_001285695.1, KT253649.1 (*Capra hircus*) and XM_025289966.1, XM_025289965.1, XM_025289964.1, XM_025289962.1 (*Bubalus bubalis*).

The sequence alignments, translations and comparisons were performed using ClustalW as described by [9].

Phylogenetic tree and UPGMA tree construction

Neighbor-Joining trees was constructed each using P-distance model and pairwise deletion gap/missing data treatment. The construction was on the basis of genetic distances, depicting phylogenetic relationships among the CSN1S1 nucleotide sequences of the ruminant species investigated. The reliability of the trees was calculated by bootstrap confidence values [10], with 1000 bootstrap iterations using MEGA 6.0 software [11]. Unweighted pair group method using arithmetic average (UPGMA) trees for each gene was constructed with consensus sequences using same model as that of the phylogenetic tree.

Test of Homogeneity of Substitution Patterns between the consensus Sequence

A Monte Carlo test (1000 replicates) was used to estimate the P-values. P-values smaller than 0.05 are considered significant.

Estimates of Codon-based Evolutionary Divergence and Codon-based Test of Neutrality for analysis between sequences

The estimate of the codon-based evolutionary divergence between sequences were were conducted using the Nei-Gojobori model [12]. The analysis involved 4 nucleotide sequences. All positions containing gaps and missing data were eliminated.

RESULTS AND DISCUSSION

Table-1: Length of the CSN1 S1 sequences of the ruminant animals investigated (base pair)

Species	Sequence length variation			
<i>Bos taurus</i>	309,	338,	383,	392
<i>Ovis aries</i>	709,	733,	715,	739
<i>Capra hircus</i>	810,	824,	642,	692
<i>Bubalus bubalis</i>	710,	734,	746,	749

Source: Genbank of the NCBI

The sequence length varied between 309-749 base pair as indicated in Table-1 above.

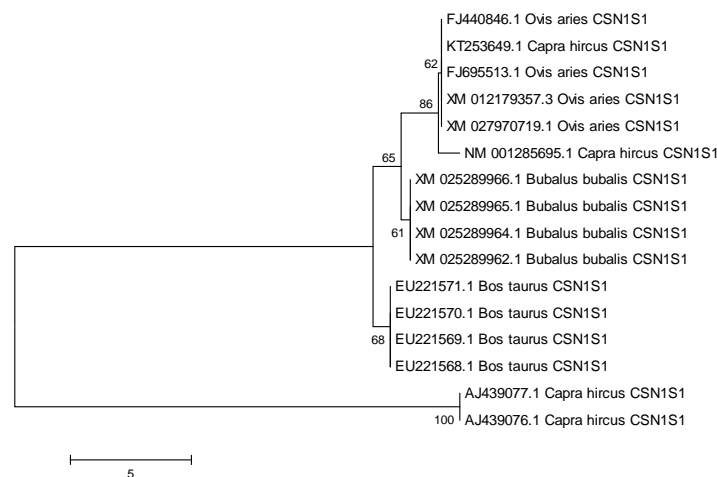


Fig-1: Neighbour joining tree

The phylogenetic tree of the neighbour joining (Figure-1) inferred the genetic distances among and between species of the CSN1S1 sequences. The tree indicate a closer genetic distance between the sequences of *Capra hircus* and *Ovis aries*, *Bubalus bubalis* and *Bos taurus*, *Capra hircus* and *Bubalus bubalis* as well as *Bos taurus* and *Capra hircus*. The neighbour-joining

tree clearly revealed that clustering was largely species-wise and this findings is in line with the report of Yakubu *et al.*, [13]. The species wise clustering might be due to species specific residues [14] and such patterns of the sequences may be explained by gene conversion and balancing selection.

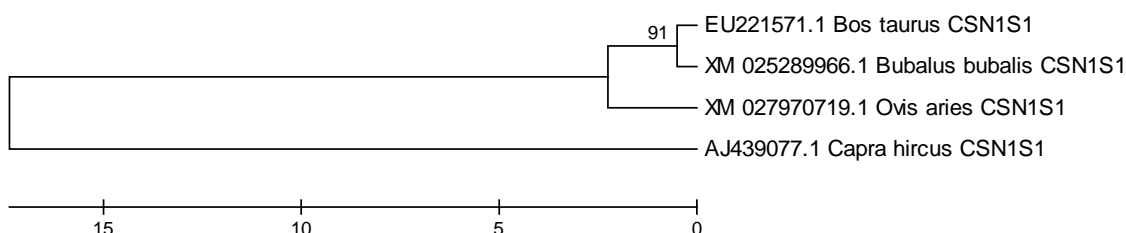


Fig-2: UPGMA tree construction of the CSN1S1 sequences

UPGMA tree derived from the consensus sequences (Figure-2), indicates that *Bos taurus* sequences were more closer to *Bubalus bubalis* sequences, compared to the sequence of *Ovis aries* and

Capra hircus respectively. Faith *et al.*, [15] also reported a similar result and in their findings on DGAT 2 gene, they observed that the sequences of sheep and goat were more closer than those of Cattle.

Table-1: Test of the Homogeneity of Substitution Patterns between the consensus Sequences of CSN1S1 gene in ruminant

S/N	Sequence	Species	<i>Bos taurus</i>	<i>Capra hircus</i>	<i>Ovis aries</i>	<i>Bubalus bubalis</i>
1	EU221571.1	<i>Bos taurus</i>	-	1.63	0.09	0.01
2	AJ439077.1	<i>Capra hircus</i>	0.01	-	5.47	5.30
3	XM_027970719.1	<i>Ovis aries</i>	0.05	0.00	-	0.00
4	XM_025289966.1	<i>Bubalus bubalis</i>	0.31	0.00	1.00	

The estimates of the disparity index per site are shown in table 1 for each sequence pair below the diagonal. The analysis involved 4 nucleotide sequences.

The results indicated that the substitution pattern were more homogenous between the sequence of *Capra hircus* and *Ovis aries* (0.00) and *Capra hircus*

and *Bubalus bubalis* (0.00) followed by the sequence of *Bos taurus* and *Capra hircus* (0.01) whereas the lowest homogeneity substitution pattern was observed between the sequence of *Bos taurus* and *Bubalus bubalis* (0.31). Ogah *et al.*, [16] and Faith and Owoeye [15] reported the minimum Dxy value of (0.00) between Cattle, Sheep and Goat.

Table-2: Estimates of Codon-based Evolutionary Divergence between CSN1S1 Sequences

Sequence	Species	<i>Bos taurus</i>	<i>Capra hircus</i>	<i>Ovis aries</i>	<i>Bubalus bubalis</i>
EU221571.1	<i>Bos taurus</i>	-			
AJ439077.1	<i>Capra hircus</i>	4.00	-		
XM_027970719.1	<i>Ovis aries</i>	33.33	33.83	-	
XM_025289966.1	<i>Bubalus bubalis</i>	1.00	5.00	37.08	-

The number of synonymous differences per sequence from between sequences are shown in Table-

2. Analyses were conducted using the Nei-Gojobori model [12]. The analysis involved 4 nucleotide

Quick Response Code



Journal homepage:

<http://crosscurrentpublisher.com/ccijavs/>

Copyright @ 2019: This is an open-access article distributed under the terms of the Creative Commons Attribution license which permits unrestricted use, distribution, and reproduction in any medium for non commercial use (Non Commercial, or CC-BY-NC) provided the original author and source are credited.

sequences. All positions containing gaps and missing data were eliminated. The results indicated that the highest evolutionary divergence was observed between

the sequence of *Ovis aries* and *Bubalus bubalis* and the evolutionary divergence appeared to be species wise.

Table-3: Codon-based Test of Neutrality for analysis between CSN1S1 sequences

Sequence	Species	<i>Bos taurus</i>	<i>Capra hircus</i>	<i>Ovis aries</i>	<i>Bubalus bubalis</i>
EU221571.1	<i>Bos taurus</i>		-0.55	-1.02	0.96
AJ439077.1	<i>Capra hircus</i>	0.59	-	-0.84	-0.20
XM_027970719.1	<i>Ovis aries</i>	0.31	0.40	-	-1.73
XM_025289966.1	<i>Bubalus bubalis</i>	0.34	0.84	0.09	-

The probability of rejecting the null hypothesis of strict-neutrality ($dN = dS$) (below diagonal) is shown in Table-3. Values of P less than 0.05 are considered significant at the 5% level. The test statistic ($dN - dS$) is shown above the diagonal. dS and dN are the numbers of synonymous and nonsynonymous substitutions per site, respectively. The results indicates that all the CSN1S1 sequences of the selected ruminant species were not significant and the test implies that the strict-neutrality for synonymous (dN) and non-synonymous (dS) substitution from the sequences used were neutral. As shown in Table-3.

CONCLUSION

The sequence length of the CSN1S1 gene varied between 309-749 base pair and the neighbour joining inferred the genetic distances among and between species of the CSN1S1 sequences. The closest genetic distance were observed between the sequences of *Capra hircus* and *Ovis aries*. Results of the substitution pattern were more homogenous between the sequence of *Capra hircus* and *Ovis aries*, *Capra hircus* and *Bubalus bubalis*. The highest evolutionary divergence was observed between the sequence of *Ovis aries* and *Bubalus bubalis*. The non- significant strict-neutrality test for synonymous (dN) and non-synonymous (dS) substitution from the sequences used were neutral. This findings would aid in understanding of species genetic variation and similarities resulting from their sequences.

REFERENCES

1. Ferretti, L., Leone, P., & Sgaramella, V. (1990). Long range restriction analysis of the bovine casein genes. *Nucleic acids research*, 18(23), 6829-6833.
2. Farrell Jr, H. M., Jimenez-Flores, R., Bleck, G. T., Brown, E. M., Butler, J. E., Creamer, L. K., ... & Swaisgood, H. E. (2004). Nomenclature of the proteins of cows' milk—Sixth revision. *Journal of dairy science*, 87(6), 1641-1674.
3. Neveu, C., Riaublanc, A., Miranda, G., Chich, J. F., & Martin, P. (2002). Is the apocrine milk secretion process observed in the goat species rooted in the perturbation of the intracellular transport mechanism induced by defective alleles at the α_{1S} Cn locus?. *Reproduction Nutrition Development*, 42(2), 163-172.
4. Brignon, G., Mahe, M. F., Ribadeau-Dumas, B., Mercier, J. C., & Grosclaude, F. (1990). Two of the three genetic variants of goat α_{1S} -casein which are synthesized at a reduced level have an internal deletion possibly due to altered RNA splicing. *European Journal of Biochemistry*, 193(1), 237-241.
5. Martin, P., Ollivier-Bousquet, M., & Grosclaude, F. (1999). Genetic polymorphism of caseins: a tool to investigate casein micelle organization. *International Dairy Journal*, 9(3-6), 163-171.
6. Moioli, B., D'andrea, M., & Pilla, F. (2007). Candidate genes affecting sheep and goat milk quality. *Small Ruminant Research*, 68(1-2), 179-192.
7. Vercesi Filho, A. E., Camargo, G. M. F., Cardoso, D. F., & Zadra, E. F. (2012). Identificação de alelos A1 e A2 para o gene da betacaseína na raça Gir Leiteiro. *IX Simpósio Brasileiro de Melhoramento Animal. Anais... João Pessoa, Paraíba*.
8. Bonfatti, V., Di Martino, G., Cecchinato, A., Vicario, D., & Carnier, P. (2010). Effects of β - κ -casein (CSN2-CSN3) haplotypes and β -lactoglobulin (BLG) genotypes on milk production traits and detailed protein composition of individual milk of Simmental cows. *Journal of dairy science*, 93(8), 3797-3808.
9. Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., ... & Thompson, J. D. (2007). Clustal W and Clustal X version 2.0. *bioinformatics*, 23(21), 2947-2948.
10. Felsenstein, J. (1985). Confidence limits on phylogenies: an approach using the bootstrap. *Evolution*, 39(4), 783-791.
11. Tamura, K., Stecher, G., Peterson, D., Filipski, A., & Kumar, S. (2013). MEGA6: molecular evolutionary genetics analysis version 6.0. *Molecular biology and evolution*, 30(12), 2725-2729.
12. Nei, M., & Gojobori, T. (1986). Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. *Molecular biology and evolution*, 3(5), 418-426.
13. Yakubu, A., Faith, E.A., Peters, S. O., Takeet, M. I., de Donato, M., & Imumorin, S. O. (2014). In silico molecular analysis of the evolution and differentiation of lactoferrin gene in some ruminant and non-ruminant animals. Proceedings of the 37th Annual Conference of Genetics Society of

- Nigeria (GSN), Federal University, Lafia, 21st-24th October, 2013, 539-548.
14. Takahashi, K., & Nei, M. (2000). Efficiencies of fast algorithms of phylogenetic inference under the criteria of maximum parsimony, minimum evolution, and maximum likelihood when a large number of sequences are used. *Molecular Biology and Evolution*, 17(8), 1251-1258.
 15. Elijah, A. F., & Ayoadele, O. O. (2017). Genetic diversity of lactoferrin gene in-silico on selected Mammalian species. *Biotechnology in Animal Husbandry*, 33(2), 171-180.
 16. Ogah, D. M., Faith, E. A., Idris, U., & Abdulmaliki, R. O. (2017). Genetic diversity of DGAT1 gene in-silico on some selected mammalian species. Proceedings of the 41ST Annual Conference of Genetics Society of Nigeria (GSN), Federal University, of Agriculture Makurdi, Benue State, Nigeria. 15th -19th, 370-374.