Journal of Biotechnology Research
ISSN(e): 2413-3256, ISSN(p): 2413-8878
Vol. 4, Issue. 2, pp: 4-8, 2018
URL: http://arpgweb.com/?ic=journal&journal=16&info=aims

Academic Research Publishing Group

Original Research Open Access

Evaluation of Functional Coding of Non-Synonymous (Nssnp) and Genetic Relationship of Alpha Casein S1 Gene in Some Selected Ruminants

Dauda A.*
Department of Animal Science University of Calabar P.M.B 1115 Calabar, Nigeria

Duwa H.
Department of Animal Science University of Maiduguri P.M.B 1069 Maiduguri, Nigeria

Abstract

Bioinformatics analysis of functional coding of non-synonymous (nsSNP) and genetic relationship of Alpha casein s1 gene in sheep, goat and cattle was performed using A total of twenty (20) alpha casein s1 nucleotide sequences comprising sheep (6), goat (7) and cattle (7) were retrieved from the GenBank (NCBI) (www.ncbi.nlm.nih.gov). The Genbank accession numbers of the sequences were ACJ46473, ACJ46472, AAB34798, AAB34797, XP_012034747 and NP_001009795 (sheep) NP_001272624, ALJ30148, CAA51022, XP_017904622, XP_017904621, XP_017904620 and XP_017904618 (goat) ACG63494, ABW98949, ABW98945, ABW98943, XP_015327132, XP_005208086 and XP_015327137 (cattle). The results of Functional analyses of coding nsSNP of the alpha casein S1 gene of sheep, goat and cattle showed both deleterious/harmful and beneficial/neutral for the three species. The Tajima's neutrality test showed positive values for all the species. The phylogenetic relationship revealed that similarity and dissimilarity may be due evolution and selection pressure. The study concluded that information revealed will help in genotype-sphenotype selection for milk quality.

Keywords: Evaluation; Functional; Non-synonymous; Genetic; Evolutionary.

1. Introduction

The caseins are the major milks proteins of mammals. Their dual function for the suckling infant is to serve as a major source of amino acid, as well as to transport phosphate and calcium in sufficient amounts to support growth of bones [1]. Effects of milk protein polymorphism on milk production traits (milk, fat, and protein yield and also fat and protein percentage) have been investigated during the past decades and, in some cases, results are still conflicting [2-4]. For centuries, animal breeders have greatly and effectively manipulated the genomes of livestock species and enhanced production traits in their herds by selecting superior individuals as predecessors for the next generations. This indirect use of hereditary information without involving molecular knowledge may give useful results in terms of production but may ignore some reproductive traits of animal. Hence, there is need to use selection methods that are based on genomic studies [5]. However, there is a long history of research on the use of genetic markers to identify quantitative trait loci and their use in marker-assisted selection but with limited implementation in practical breeding programs [6]. The objective of this study, therefore, was to investigate using bioinformatic to assess non-synonymous polymorphism and genetic relationship of alpha casein s1 gene of cattle, sheep and goat with a view to providing relevant genetic information for marker assisted selection in the studied species.

2. Materials and Method

A total of twenty (20) alpha casein s1 nucleotide sequences comprising sheep (6), goat (7) and cattle (7) were retrieved from the GenBank (NCBI) (www.ncbi.nlm.nih.gov). The Genbank accession numbers of the sequences were ACJ46473, ACJ46472, AAB34798, AAB34797, XP_012034747 and NP_001009795 (sheep) NP_001272624, ALJ30148, CAA51022, XP_017904622, XP_017904621, XP_017904620 and XP_017904618 (goat) ACG63494, ABW98949, ABW98945, ABW98943, XP_015327132, XP_005208086 and XP_015327137 (cattle). Sequences alignment, translation and comparison of the alpha casein s1 gene of the various species was done with ClustalW as described by Larkin, *et al.* [7] using IUB substitution matrix, gap open penalty of 15 and gap extension penalty of 6.66**.** *In silico* functional analysis, missense mutations was obtained using PROVEAN (Protein Variant Effect Analyzer) with threshold value of -2.5. PROVEAN collects a set of homologous and distantly related sequences from the NCBI NR protein database using BLASTP (ver.2.2.25) with an E-value threshold of 0.1. The sequences were clustered based on a sequence identity of 80% to remove redundancy using the CD-HIT program (ver.4.5.5) [8]. If the PROVEAN score is smaller than or equal to a given threshold, the variation is predicted as deleterious [9]. The evolutionary history was inferred using the Neighbor-Joining method [10]. The bootstrap consensus tree inferred from 1000 replicates [11] is taken to represent the evolutionary history of the taxa analyzed [11]. Branches corresponding to partitions reproduced in less than 50% bootstrap replicates are collapsed. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) are shown next to the branches [11]. The evolutionary distances were computed using the Poisson correction method [12] and are in

*Corresponding Author

the units of the number of amino acid substitutions per site. The analysis involved 20 amino acid sequences. All positions containing gaps and missing data were eliminated. There were a total of 172 positions in the final dataset. Evolutionary analyses were conducted in MEGA7 [13]. The Tajima test statistic [14] was estimated using MEGA7. All positions containing gaps and missing data were eliminated from the dataset (Complete deletion option). The abbreviations used are as follows: m = number of sites, S = Number of segregating sites, p_s = S/m, $\Theta = p_s/ a_1$, and π = nucleotide diversity. D is the Tajima test statistic [15].

3. Results

The Functional analyses of coding non-synonymous single nucleotide polymorphism (nsSNP) of the alpha casein S1 gene of sheep, goat and cattle are presented in Table 1, 2 and 3 respectively. Amino acid substitutions of the wild type alleles located in the putative peptide coding region of sheep, goats and cattle were obtained from the alignment of deduced amino acid sequences of sheep, goats and cattle. The result of functional analysis of coding nsSNP of the alpha casein S1 gene of sheep revealed that variants (A11L, L14E, P17N, P44I, A69M, E76K and P80S) returned deleterious or harmful while the rest returned beneficial or neutral. The result of functional analysis of the alpha casein S1 gene of goat showed that variants (A13L, R16P, E29V, E54L, D58I, E62S, D71A, Q74M and S82Y) returned deleterious while the remaining variants are neutral or beneficial. The result of functional analysis of the alpha casein S1 gene of cattle showed that nine variants are deleterious and nine are beneficial. The Results from Tajima's Neutrality test are presented in Table 4. The result showed that cattle have the highest value for nucleotide diversity and all the species revealed positive values for (D) Tajima's test. The phylogenetic relationship of nucleotide sequence of the three species (sheep, goat and cattle) is shown in Figure 1. The result revealed that all the three species intermingle with each other showing close relationships.

Table-1. Functional analysis of coding nsSNP of the alpha casein S1 gene of sheep using PROVEAN

Variant	PROVEAN Score	Prediction
A11L	-4.199	Deleterious
L14E	-5.172	Deleterious
P17N	-5.824	Deleterious
H19N	-1.555	Neutral
K22R	-1.369	Neutral
G25V	-0.692	Neutral
S27P	-1.284	Neutral
V30E	-0.281	Neutral
F38N	-2.805	Neutral
P44I	2.854	Deleterious
S56D	-2.148	Neutral
D58K	-1.177	Neutral
Q66D	0.739	Neutral
A69M	-2.917	Deleterious
S71D	-2.176	Neutral
S74Q	-2.176	Neutral
E76K	-3.100	Deleterious
P80S	-4.528	Deleterious

Default threshold is −2.5, that is; Variants with a PROVEAN score equal to or below −2.5 are considered "deleterious" while Variants with PROVEAN score above −2.5 are considered "neutral". G = glycine, A = Alanine, L = leucine, M = methionine, F = phenylalanine, W = tryptophan,Q = glutamine, E = glutamic acid, S = serine, P = proline, V = valine, Y = tyrosine, R = arginine, N = asparagine, T = threonine, C = cysteine.

Table-2. Functional analysis of coding nsSNP of the alpha casein S1 gene of goat using PROVEAN

Variant	PROVEAN Score	Prediction
V10A	-2.410	Neutral
A13L	-3.876	Deleterious
R16P	-4.475	Deleterious
H19P	-1.612	Neutral
N22H	-0.860	Neutral
G25L	-1.504	Neutral
E29V	-2.689	Deleterious
E33N	-1.854	Neutral
L36R	-0.656	Neutral
P42F	-2.361	Neutral
V46F	-1.193	Neutral
E54L	-5.341	Deleterious
D58I	-3.632	Deleterious
E62S	-3.248	Deleterious
D66Q	-0.761	Neutral

D71A	-3.693	Deleterious
Q74M	-3.073	Deleterious
G78S	2.027	Neutral
S82Y	-4.760	Deleterious

Default threshold is −2.5, that is; Variants with a PROVEAN score equal to or below −2.5 are considered "deleterious" while Variants with PROVEAN score above −2.5 are considered "neutral". G = glycine, A = Alanine, L = leucine, M = methionine, F = phenylalanine, W = tryptophan,Q = glutamine, E = glutamic acid, S = serine, P = proline, V = valine, Y = tyrosine, R = arginine, N = asparagine, T = threonine, C = cysteine.

Table-3. Functional analysis of coding nsSNP of the alpha casein S1 gene of cattle using PROVEAN

Variant	PROVEAN Score	Prediction
V10E	-4.147	Deleterious
A13N	-4.850	Deleterious
R16L	-3.578	Deleterious
H19F	0.132	Neutral
H23P	-4.732	Deleterious
L26E	2.238	Neutral
V30K	-1.054	Neutral
N34E	-0.211	Neutral
R37S	-1.048	Neutral
V40I	-0.596	Neutral
F43E	-0.272	Neutral
V46E	1.802	Neutral
N53I	-3.983	Deleterious
S56M	-0.841	Neutral
E62S	-3.289	Deleterious
Q67I	-3.286	Deleterious
Q74L	-3.706	Deleterious
S82V	-5.094	Deleterious

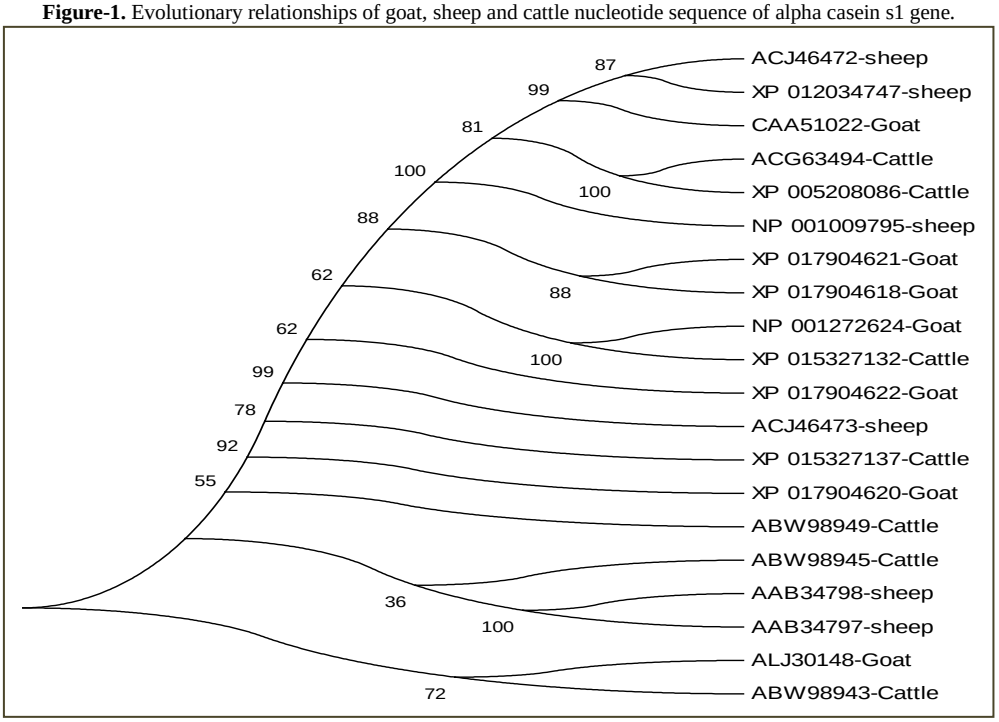
Default threshold is −2.5, that is; Variants with a PROVEAN score equal to or below −2.5 are considered "deleterious" while Variants with PROVEAN score above −2.5 are considered "neutral". G = glycine, A = Alanine, L = leucine, M = methionine, F = phenylalanine, W = tryptophan,Q = glutamine, E = glutamic acid, S = serine, P = proline, V = valine, Y = tyrosine, R = arginine, N = asparagine, T = threonine, C = cysteine.

Table-4. Results from Tajima's Neutrality test

Species	M	S	Ps	Θ	Π	D
Cattle	7	171	1.000000	0.408163	0.825118	5.991808
Sheep	6	168	0.982456	0.430273	0.631969	3.040551
Goat	7	171	0.994186	0.405790	0.667220	3.778823

m = number of sites, S = Number of segregating sites, p_s = S/m, Θ = p_s/ a_1, and π = nucleotide diversity. D is the Tajima test statistic

Figure-1. Evolutionary relationships of goat, sheep and cattle nucleotide sequence of alpha casein s1 gene.



4. Discussion

Alpha casein S1 is a polymorphic gene [16]. Ruminant's milk proteins are coded by highly polymorphic genes, containing an unusually large number of polymorphisms [17]. In this study the Functional analyses of coding non-synonymous single nucleotide polymorphism (nsSNP) of the alpha casein S1 gene of sheep, goat and cattle are presented in Table 1, 2 and 3 respectively appeared both beneficial/neutral and deleterious/harmful. The beneficial amino acid substitution do not impaired with protein structure, function, chemical and physical properties of protein. According to Bibinu, *et al.* [18] noted that the neutral or beneficial amino acid substitutions are those substitutions that help in maintaining the structural integrity of cells and tissues. Also, they affect positively the functional roles of proteins involved in signal transduction of visual, hormonal, and other stimulants. The beneficial nsSNPs obtained in the present study, therefore, offer hope for future genetic improvement of sheep, goat and cattle at the Alpha casein S1 locus. This is due to the fact that nsSNPs have been reported to be linked to economically important traits and disease development [19]. However the deleterious/harmful amino acid substitution does the opposite. Selection for improvement on the milk protein quality in cattle, goat and sheep should take note that any attempt to increase the frequency of desirable traits will also lead to the risk of increasing the frequency of disease traits. The different genetic variants of milk proteins differ from each other by only a few amino acid substitutions or deletions within the polypeptide chain [20]. Therefore, selection of animals for the strong alpha casein s1 alleles could lead to improvements in milk composition and quality, particularly in protein content [21]. Since the main goal of animal breeding is to select best animal as parents of the next generation, thus, identification of new genes and/or mutations contributing to genetic variation can assist selection by reducing the generation interval, increasing fertility and genetic progress [22, 23]. Tajima's test for selection as shown in Table 4 is important to determine which type of selection the animal undergoes before now. A positive Tajima's D signifies low levels of low and high frequency polymorphisms, indicating a decrease in population size and/or balancing selection [24]. This imply that the population of the animal did not undergoes bottleneck selection or are in existence as a result of resistant to disease or loss their original purpose of existence. This might aid in purifying selection i.e. selection against non-synonymous mutations because of their deleterious effect. The nucleotide diversity of cattle is indicating that the alpha casein s1 is highly conserved. The evolutionary relationship shown in figure 1 of this study revealed close similarity of a gene among ruminants which may be as a result of recent separation in evolutionary process and/or similar selection pressure which the ruminants have suffered during evolution [25]. The phylogenetic relationship of alpha casein s1 gene is according to bovidae classification even though is express using amino acid nucleotide.

5. Conclusion

The study revealed that alpha casein s1 is polymorphic gene which showed nsSNP amino acid substitution of some variants indicating beneficial while some are deleterious. The Tajima's neutrality test for selection revealed positive value for the three species which is indicating purifying selection. The phylogenetic relationship indicated recent separation in evolutionary process or similar selection pressure during evolution. This study may provide insight for performing genotype-phenotype selection for milk quality in sheep, goat and cattle in Nigeria.

Reference

[1] Koczan, D., Hobom, G., and Seyfert, H. M., 1991. "Genomic organization of the bovine alpha-S1 casein gene." *Nuc. Acids Res.*, vol. 19, pp. 5591-5596.

[2] Aleandai, R., Buttazzoni, L. G., Schneider, J. C., Caroli, A., and Davoli, R., 1990. "The effects of milk protein polymorphisms on milk components and cheese-producing ability." *J. Dairy Sci.*, vol. 73, pp. 241-255.

[3] Bovenhius, H., Van, A. J. A. M., and Korver, S., 1992. "Associations between milk protein polymorphisms and milk production traits." *J. Dairy Sci.*, vol. 75, pp. 2549-2559.

[4] Prinzenberg, E. M., Weimann, C., Brandt, H., Bennewitz, J., Kalm, E., Schwerin, M., and Erhardt, G., 2003. "Polymorphism of the bovine CSN1S1 promoter: Linkage mapping, intragenic haplotypes, end effects on milk production traits." *J. Dairy Sci.*, vol. 86, pp. 2696-2705.

[5] André, E., 2012. "The development and application of genomic selection as a new breeding paradigm." *Anim. Front.*, vol. 2, pp. 10-15.

[6] Dekkers, J. C. M., 2012. "Application of genomics tools to animal breeding." *Current Genomics*, vol. 13, pp. 207-212.

[7] Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., Valentin, F., Wallace, I. M., Wilm, A., *et al.*, 2007. "Clustal W and Clustal X Version 2.0." *Bioinformatics*, vol. 23, pp. 2947-2948.

[8] Li, W. and Godzik, A., 2006. "Cd-Hit: A fast program for clustering and comparing large protein or nucleotide sequences." *Bioinformatics*, vol. 22, pp. 1658-1659.

[9] Choi, Y., Sims, G. E., Murphy, S., Miller, J. R., and Chan, A. P., 2012. "Predicting the functional effect of amino acid substitutions and indels." *PLoS ONE*, vol. 7, p. e46688.

[10] Saitou, N. and Nei, M., 1987. "The neighbor-joining method: A new method for reconstructing phylogenetic trees." *Molecular Biology and Evolution*, vol. 4, pp. 406-425.

[11] Felsenstein, J., 1985. "Confidence limits on phylogenies: An approach using the bootstrap." *Evolution*, vol. 39, pp. 783-791.

[12] Zuckerkandl, E. and Pauling, L., 1965. *Evolutionary divergence and convergence in proteins. Edited in evolving genes and proteins by V. Bryson and H.J. Vogel*. New York: Academic Press. pp. 97-166.

[13] Kumar, S., Stecher, G., and Tamura, K., 2016. "Molecular evolutionary genetics analysis version 7.0 for bigger datasets." *MEGA7: Molecular Biology and Evolution,* vol. 33, pp. 1870-1874.

[14] Tajima, F., 1989. "Statistical methods to test for nucleotide mutation hypothesis by DNA polymorphism." *Genetics,* vol. 123, pp. 585-595.

[15] Tajima, F., 1993. "Simple methods for testing molecular clock hypothesis. ." *Genetics,* vol. 135, pp. 599-607.

[16] Miluchová, M., Trakovická, A., and Gábor, M., 2009. "Analysis of polymorphism of alpha s1 casein of slovak pinzgau cattle by PCR-RFLP." *Lucrări Stiinţifice Zootehnie şi Biotehnologii,* vol. 42, pp. 284-287.

[17] Nilsen, R. M., Vollset, S. E., Gjessing, H. K., Skjaerven, R., Melve, K. K., Schreuder, P., Alsaker, E. R., Haug, K., Daltveit, A. K., *et al.*, 2009. "Self-selection and bias in a large prospective pregnancy cohort in Norway." *Paediatr Perinat Epidemiol.*, vol. 23, pp. 597-608.

[18] Bibinu, B. S., Yakubu, A., Ugbo, S. B., and Dim, N. I., 2016. "Computational molecular analysis of the sequences of bmp15 gene of ruminants and non-ruminants." *Open Journal of Genetics,* vol. 6, pp. 39-50.

[19] Vincent, S. T., Momoh, O. M., and Yakubu, A., 2014. "Bioinformatics analysis of beta casein gene in some selected mammalian species." *Research Opinions in Animal and Veterinary Sciences,* vol. 4, pp. 564-570.

[20] Eigel, W. N., Butler, J. E., Ernstrom, C. A., Farrell, H. M., Harwalkar, J., V. R., Jenness, R., and Whitney, R. M., 1984. "Nomenclature of proteins of cow's milk: Fifth revision." *J. Dairy Sci.*, vol. 67, pp. 1599–1631.

[21] Maga, E. A., Daftari, P., Kultz, D., and Penedo, M. C. T., 2009. "Prevalence of αs1-casein genotypes in American dairy goats." *Journal of Animal Science,* vol. 87, pp. 3464-3469.

[22] Johnston, D. J., Barwick, S. A., Corbet, N. J., Fordyce, G., Holroyd, R. G., Williams, P. J., and Burrow, H. M., 2009. "Genetics of heifer puberty in two tropical beef genotypes in northern Australia and associations with heifer and steer-production traits." *Anim. Prod. Sci.*, vol. 49, pp. 399-412.

[23] Fortes, M. R. S., Lehnert, S. A., Bolormaa, S., Reich, C., Fordyce, G., Corbet, N. J., Whan, V., Hawken, R. J., and Reverter, A., 2012. "Finding genes for economically important traits: Brahman cattle puberty." *Anim Prod Sci.*, vol. 52, pp. 143-150.

[24] Fu, Y. X. and Li, W. H., 1993. "Statistical tests of neutrality of mutations." *Genetics,* vol. 133, pp. 693-709.

[25] Sun, Y., Zhang, X., Xi, D., Li, G., Wang, L., Zheng, H., Du, M., Gu, Z., Yang, Y., *et al.*, 2015. "Isolation and cdna characteristics of mhc-dra genes from gayal (bos frontalis) and gaytle (bos frontalis × bostaurus)." *Biotechnology & Biotechnological Equipment,* vol. 29, pp. 33-39.