IN SILICO ANALYSIS OF BETA-LACTOGLOBULIN GENE IN SOME SELECTED MAMMALIAN SPECIES

O.M. Momoh[1], S.T. Vincent [1], A. Yakubu[2]

[1]Department of Animal Breeding and Physiology, College of Animal Science, University of Agriculture, Makurdi, Benue State, Nigeria.
[2]Department of Animal Science, Faculty of Agriculture, Nasarawa State University, Keffi, Shabu-Lafia Campus, Nasarawa State, Nigeria.
*Corresponding author: vincent.samuelter@gmail.com
Original Scientific paper

Abstract: This study investigated in silico, the genetic diversity of Beta-Lactoglobulin (β-Lg) and their evolutionary and differentiation within and among selected mammalian species; and also examined the attendant effects of polymorphism on the functionality of the gene. A total of 21 β-Lg gene sequences with corresponding amino acids belonging to 6 species [cattle (4), buffalo (4), sheep (3), goat (3), pig (3) and horse (4)] were retrieved from GenBank (www.ncbi.nlm.nih.gov). All sequences were trimmed to equal length (500bp) corresponding to the same region. Sequences' alignment, translation and comparison were done with ClustalW using IUB substitution matrix, gap open penalty of 15 and gap extension penalty of 6.66. The alignment revealed high polymorphism of sequences among extant species. The Dxy inferred using p-distance revealed that sheep and goat had the lowest distance of 0.05 with a maximum distance of 0.65 between goat and horse. The hypothesis of strict neutrality ($d_N = d_S$) was rejected for all extant species as allelic sequence evolution was driven by both purifying and positive selection. Only those of pig and buffalo were driven by positive selection. In-silico functional analysis of non-synonymous mutations using PANTHER revealed that, all the 12 amino acid substitutions (10 in cattle and 2 in sheep) did not impair protein function. The Neighbour-Joining phylogeny revealed trans-species evolution, but a species-wise phylogeny was obtained for UPGMA with consensus sequences. Thus, all probed SNPs from this study have no deleterious effect and can be tolerated by breeders when selecting stocks for milk improvement.

Key words: single nucleotide polymorphism, beta-lactoglobulin, in silico, phylogeny.

Introduction

To select for greater milk production, breeders have traditionally relied upon phenotypic evaluation and selection. In developed countries, phenotypic selection has yielded appreciable degree of successes in predicting genetically superior animal (*Staiger, 2007*). In United State, average milk production per lactation has doubled in the last forty (*Dekkers and Hospital, 2002*) while in United Kingdom, production has tripled in the last seventy years (*Simm, 1998*). More than half of this increased in milk production has been due to improvement in genetics. Despite these increases, phenotypic selection could be improved to give better accuracy by utilizing molecular genetics and bioinformatics selection techniques. Because milk production is a quantitative trait, heritability value is not a perfect predictor of genetic merit of an individual (*Staiger, 2007*). In addition, milk production can only be measured in females that have reached sexual maturity or calved; this makes it very difficult to analyze males and prepubescent animals. However, with recent development and advances in DNA technology to identify genes, QTL and SNPs, many of these pitfalls can be overcomed because DNA extraction and analysis is not limited by age, sex or time.

Milk proteins have been grouped into casein and whey. Casein accounts for about 80% while whey accounts for 20% of milk protein (*Hoffman and Falvo, 2004*). Beta-lactoglobulin (β-Lg) is a lipocalin, a widely diverse family, most of which bind small hydrophobic ligands and thus may act as specific transporters, as does serum retinol binding protein. In bovine, β-Lg gene is located on chromosome 11 (*Berry et al., 2010*).

Milk proteins exhibit genetic polymorphism at both protein and DNA levels and this polymorphism could be due to amino acid substitution or deletion of small peptides along the polypeptide chain (*Chin, 1998*). Today, a number of studies have indicated that milk production, composition and quality are affected by genetic variants of milk proteins (*Chin, 1998*). For instance, it has been reported that β-casein A2 and A3 and K-casein are associated with higher milk yield when compared with other variants (*Ng-Kwai-Hang et al., 1986*). Therefore, identification and probing of milk protein variants provides an important tool for complementing traditional breeding methods in improving the yield and quality of milk and dairy production (*Chin, 1998*). This study aimed at examining the genetic diversity of β-Lg gene *in-silico* especially on its evolution and differentiation within and among species and also examining the attendant effects of polymorphism on its functionality in the selected mammalian species.

Materials and Methods

Sequences Retrieval

A total of twenty-one (21) β-Lg sequences from six (6) species [cattle (4), buffalo (4), sheep (3), goat (3), pig (3) and horse (4)] were retrieved from the GenBank (www.ncbi.nlm.nih.gov). The GenBank accession numbers of the sequences are: U60978, U60356, NM_001082494, NM_001082493 (Horse); AY962046, X54976, NM_213754 (Pig); Z19569, FR823504, X58471(Goat); FR821305, FR821309, AY515301 (Sheep); M19088, AJ296332, Z11996, X63139 (Cattle) and JF274007, FN377869, HQ398854, AJ005429 (Buffalo).

Multiple Sequence Alignment (MSA) and Translation

Sequences' alignment, translation and comparison were done with Clustal W as described by *Larkin et al*. (*2007*) using IUB substitution matrix, gap open penalty of 15 and gap extension penalty of 6.66.

Test of Selection

The relative proportion of non-synonymous substitution per non-synonymous site (d_N) and the number of synonymous substitutions per synonymous site (d_S) were estimated using the method of *Nei and Gojobori* (*1986*). The ratio of non-synonymous to synonymous divergence "ω" or (d_N/d_S) was also tested for departure from the neutral expectation of unity using the codon based Z-distribution model as modified by Nei-Gojobori, applying proportion correction.

Functional Analysis

In-silico functional analysis of non-synonymous mutation was estimated using PANTHER (*Thomas et al., 2003*). PANTHER tools estimates the likelihood of a given non-synonymous (mutation) Coding SNP to initiate a functional change in protein. The subPSEC (substitution position-specific evolutionary conservation) score was also estimated based on alignment of evolutionarily related proteins. The probability that a given variant will cause a deleterious effect on protein function is estimated by $P_{deleterious}$, such that a subPSEC score of -3 corresponds to a $P_{deleterious}$ of 0.5 (*Brunham et al., 2005*). The subPSEC score is the negative logarithm of the probability ratio of the wild-type and mutant amino acids at a particular position. PANTHER subPSEC scores, derived from the probabilities of observing the variant amino acids in a PANTHER Hidden Markov Model (HMM), are continuous values from 0 to -10. When subPSEC = 0, the substitution is interpreted as functionally neutral, whereas more negative values of subPSEC predict more deleterious substitutions *(Brunham et al., 2005*). The published SNPs used for this study include: 10 (cattle): Gly64Asp, Ala118Val, Gln59His, Glu45Gln, Glu158Gly, Pro50Ser, Asp130Tyr, Ile78Met, Glu108Gly, Pro126Leu and 2 (sheep): Tyr20His, Arg148Glu.

Phylogenetic Analysis

Neighbor-Joining NJ tree 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 β-Lg nucleotide sequences of the investigated species. The reliability of the tree was calculated by bootstrap confidence values (*Felsenstein, 1985*), with 1000 bootstrap iterations using MEGA 5.1 software *(Tamura et al., 2011*). Similarly, UPGMA tree for each gene was also constructed with consensus sequences; using same model as that of the NJ tree. All sequences were trimmed to (500bp) equal length corresponding to same region before generating the tree.

Results and Discussion

In Table 1 is shown the estimated distance matrix for β-Lg between consensus sequences of 6 mammalian species. Maximum Dxy was between goat and horse and minimum Dxy was between sheep and goat. The average genetic distance Dxy is an index of divergence within and across species; where Dxy=distance between sequence x and sequence y. The higher the value of Dxy the far apart the two species are. Amongst ruminants, the maximum Dxy value of 0.62 obtained between cattle and buffalo is also evident on the dendrogram which shows evidence of trans-species evolution. The high value between these two species might have arisen from the different regions of DNA used for this study or probably, the included sequences are the divergent regions of the two species. The minimum value between sheep and goat is consistent with recent molecular grouping of sheep and goat. Also the maximum value between goat and horse is consistent with classical grouping as goat is expected to be closer to sheep, cattle, buffalo and pig than horse. This is similar to the findings of *Vincent et al*. (*2014*) on β-CN, who reported maximum Dxy between goat and horse.

Table 2 shows the means of d_S and d_N, omega (ω), z-statistics and probability. The 'ω' value reveals that evolution is driven by positive selection for only pig and buffalo whereas all other species are under purifying selection. Insight into the mechanism by which natural selection drives gene functional diversification across different species and lineages is a key issue in biology (*Toll-Riera et al. 2011; Yakubu et al. 2013a*).

Table 1: Evolutionary Divergence of β-Lg between Species

	Cattle	Buffalo	Goat	Sheep	Pig	Horse
Cattle		0.02	0.03	0.03	0.03	0.03
Buffalo	0.62		0.03	0.03	0.03	0.03
Goat	0.53	0.57		0.01	0.03	0.03
Sheep	0.55	0.57	0.05		0.03	0.03
Pig	0.55	0.57	0.57	0.59		0.03
Horse	0.63	0.59	0.65	0.64	0.61	

Values above the diagonal represent standard error estimate(s) while those below the diagonal are the average genetic distances between species.

Table 2: Mean Numbers of Nucleotide Substitutions per Synonymous Site (d_S) and per Non-Synonymous Site (d_N) With Their Ratio In β-Lg Among Selected Extant Species

Species	Codons	d_S (±SE)	d_N (±SE)	ω	Z-Statistics	P-value
Buffalo	33	1.05±0.08	1.14±0.05	1.09	5.34	0
Cattle	27	1.96±0.12	1.29±0.06	0.66	-0.14	0.89
Sheep	45	0.80±0.07	0.56±0.03	0.7	-0.39	0.7
Goat	44	0.77±0.07	0.53±0.03	0.69	-0.01	0.98
Pig	39	0.47±0.06	0.50±0.03	1.06	3.26	0
Horse	48	0.44±0.08	0.24±0.03	0.55	-0.18	0.85

ω = omega or (d_N/d_S), d_N = relative proportion of non-synonymous substitution per non-synonymous site, d_S= the number of synonymous substitutions per synonymous sites

The varying substitutions of amino acids within and across species might be as a result of separate divergence from their common ancestor. According to *Marini et al*. (*2010*), as orthologs diverge from their most recent common ancestor, their different evolutionary trajectories lead to divergence in the selective constraints on homologous sites. The comparison of the number of non-synonymous substitution per non-synonymous sites (d_N; amino acid altering) to the number of synonymous mutations per synonymous sites (d_S; silent mutation) also known as omega ($\omega = d_N/d_S$); is a useful estimate of gene selective pressure (*Yakubu et al. 2013a*). *Zhang et al*. (*2005*) noted that omega (ω) >1 implies positive selection, that is, selection has caused some amino acid substitutions that are non-deleterious and that the operative effect of purifying selection is not strong enough to overcome the effect of positive selection. In this study, the null hypothesis of strict neutrality (d_N-d_S) was rejected. The estimated (ω) values which range from 0.55-1.09 symbolize the operation of both positive and purifying selection. Only pig and buffalo had ω >1 signifying positive selection; while other

species with ω<1 are under purifying selection. *Yakubu et al*. (*2013b*), in their study on MHC reported that, positive selection favours new variants thereby increasing allelic polymorphism *(Bergstronm and Gyllensten, 1995*) which in turn favours the ability of antigen-presenting cells to bind a wide array of self and non-self-peptides, thus conferring higher resistance to infectious diseases (*Hedrick and Kim, 2000*). Mutation in pig and buffalo increases fitness (*Kosakovsky Pond, 2011*) and this is expected to reflect on performance. The d_N-d_S shows the surplus of beneficial or deficit of deleterious non-synonymous substitution. The high 'ω' values might enhance retinol binding, immune response and metabolism of phosphate in the mammary gland (*Kontopidis et al., 2003*)

Table 3 depicts predicted effect of non-synonymous amino acid variants in β-Lg gene. Several SNPs have been reported to serve as biomarkers for exploring the genetic bases of disease conditions (*Tariq et al., 2013*) and production traits. The prediction of SNPs status is promising in modern genetics analysis and breeding programmes as they have been used to identify those animals with higher breeding value. But it is still a great challenge to identify functional SNPs in production or disease related gene. However, computational approach has helped in overcoming this challenge and this has increased the success rate of genetic association studies. Major interest in both human and animal genetics is to distinguish mutations that increase fitness from those that reduce fitness (*Tariq et al., 2013*).

Table 3: Prediction of the Effect of Non-Synonymous Amino Acid Variants in β-Lg Gene

Mutated amino acid	SubPSEC score	P deleterious	P Substituted	Wild-type
Cattle				
Gly64Asp	-	-	-	Lys
Ala118Val	-	-	-	Lys
Gln59His	-	-	-	His
Glu45Gln	-	-	-	Asp
Glu158Gly	-	-	-	Lys
Pro50Ser	-	-	-	Asp
Asp130Tyr	-	-	-	Pro
Ile78Met	-	-	-	Trp
Glu108Gly	-	-	-	Lys
Pro126Leu	-	-	-	Asn
Sheep				
Tyr20His	-	-	-	Ile
Arg148Glu	-	-	-	Asp

In this study, a total of 12 SNPs were sourced and probed, however, all the probed SNPs were beneficial i.e. protein function is not altered in anyway. Therefore, for milk improvement programmes, all probed SNPs should be noted.

Figure 1 depicts the deduced amino acid by ClustalW excluding sites with missing/ambiguous data and gaps. A greater level of polymorphism was shown within and across species. The topology of distance-based β-Lg NJ-tree is a clear reflection of trans-species evolution (figure 2) which could be attributed to the different regions used for the study. Random clustering of sequences was exhibited by all the studied species except for horse whose entire sequences aggregated. This distinct embranchment shown by horse is premised on the identical promoter region. Generally, the random clustering may largely be due to regional variability such that those of similar region clustered together. On the contrary, the UPGMA tree revealed species-wise evolutionary history (figure 3); a congruence with classical taxonomy.

```
#BUFFALO_JF274007    EAVPAWHSCV TCASPPPPQG PWRSGGGDSF SGGALSSHGW GRCPRPISPG SRRLPGRH [184]
#BUFFALO_FN377869    GLT..Q.AGP E.PEGAAHRR SC.M.NEARS LKKPSLVQAL TKSLCLALH. AKPV.D.R [184]
#BUFFALO_HQ398854    YFGYFQYKVF F.FVQGAGEK HLLLINSI.S LVWGPWGGLP MWVCSLMLGD NELAHQKQ [184]
#BUFFALO_AJ005429    GLT..Q.AGP E.PEGAAHRR SC.M.NEARS LKKPSLVQAL TKSLCLALH. AKPV.D.R [184]
#CATTLE_M19088       PLT.VQ.AGP Q.PEGAAHRR SC.M.NEARS LKKPSLVQAL TKSLCLALH. AKPV.D.R [184]
#CATTLE_AJ296332     GLS..Q.AGP E.PEAAAFRR SC.MVSKAHS PIWTPGVQAP NGAWSWGCSR FNFMLNKT [184]
#CATTLE_Z11996       PLA.SRLVSS V.ILCVVFKH .RP.SNETPL .DTP.LPYHP DEWFTLTPAS QGYELETP [184]
#CATTLE_X63139       .GSL.RELGL K.WEVLAFHW SSALSCRGFP ..LL.LCCLP SG.LQ.GARE P.LVTRY. [184]
#GOAT_Z19569         GLT..Q.AGP E.PEGAAHRR SC.M.NEARL LKKPSLVQAL TKSLCLALH. AKPV.D.G [184]
#GOAT_FR823504       MV.HELLPEW WHFLERACLP .CIPSRSA.L .LTPSKPC.L ..A.H.PLDQ HGQGS.AR [184]
#GOAT_X58471         GLT..Q.AGP E.PEGAAHRR SC.M.NEARL LKKPSLVQAL TKSLCLALH. AKPV.D.G [184]
#SHEEP_FR821305      VV.HELLPEW WHFLEQACLP .CIPSKSA.L .LTPSKPC.L ..ASH.PLDQ RG.GS.AC [184]
#SHEEP_FR821309      VV.HELLPEW WHFLEQACLP .CIPSKSA.L .LTPSKPC.L ..ASH.PLDQ RG.GS.AC [184]
#SHEEP_AY515301      .FGILLSREF KNFNNWVHYV Q.EVHERFP. YTCMAHDQLP AGA.GRTDAQ LDS.QDCA [184]
#PIG_AY962046        RLAG.RA.ER GEDGAVGWDW A.SWTSRVG. ..HSAPGRSS ILAANL.AE. PDLSSSIL [184]
#PIG_X54976          GHT..QRAGQ EQPEGAAHRR SC.A.NKACS LKKPSLVQAL TSSSCLAPH. PQPE.D.P [184]
#PIG_NM_213754       GHT..QRAGQ EQPEGAAHRR SC.A.NKACS LKKPSLVQAL TSSSCLAPH. PQPE.D.P [184]
#HORSE_U60978        GGT.VQ.PGR E.SEGAAHRR SCAR.NALTL WRRLTQV.YQ E.F.CLVLC. R.HV.D.G [184]
#HORSE_U60356        GST..Q.PGR E.SEEAVHRR LCVR.N.ARS LVRLALVQYL T.SL.LVPHE TEPV.DWG [184]
#HORSE_NM_001082494  RGT.VQ.PGR E.SEGAAHRR SCAI.DAVTL WRRLTQV.YQ D.F.CLVLC. R.HV.D.G [184]
#HORSE_NM_001082493  GST..Q.PGR E.SEEAAHRR SCVR.N.ARS LKRLAQVQYL T.SL.LVPHE TEPV.DWG [184]
```

Figure 1. Amino acid prediction of mammalian β-Lg, deduced by ClustalW excluding sites with missing and ambiguous data and gaps
Dots indicated identical amino acids and numbers on the right hand side represent site number.

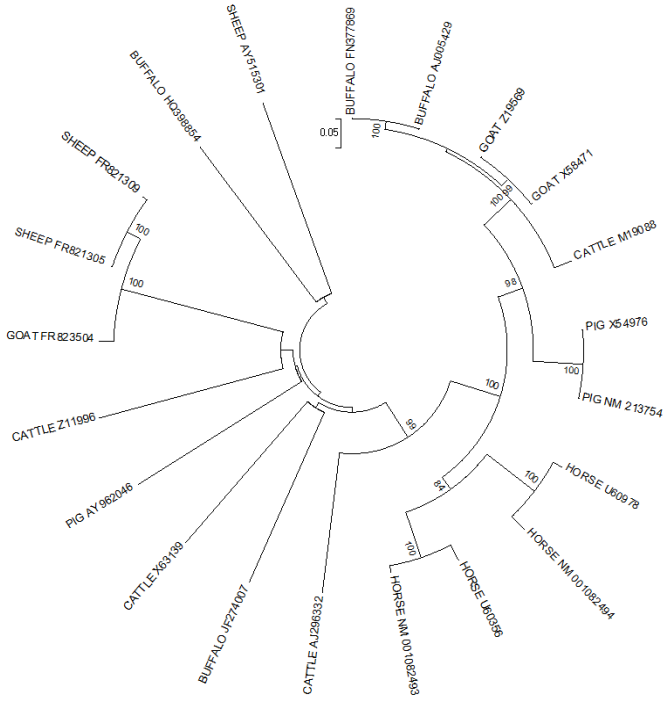


Figure 2. Phylogenetic tree of mammalian β-Lg computed using NJ-method

.

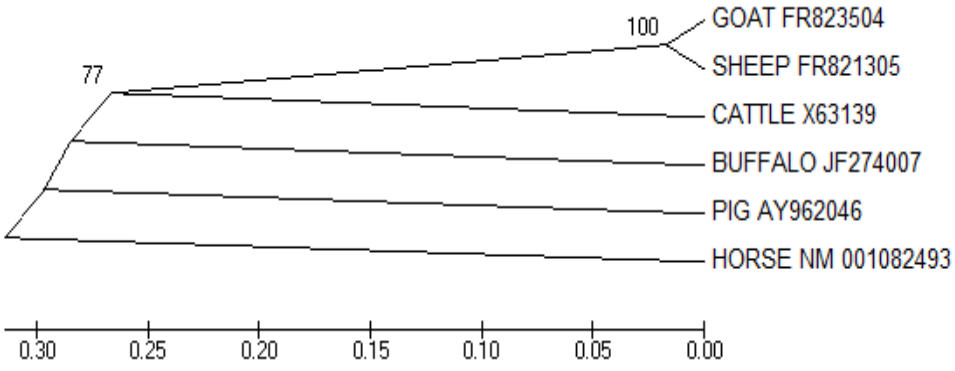


Figure 3. Phylogenetic tree derived from consensus sequences of each of the selected species using the UPGMA method.

Conclusion

In this study, the hypothesis of strict neutrality (d_N=d_S) was rejected for all extant species as allelic sequence evolution was driven by both purifying and positive selection. However, only those of pig and buffalo were driven by positive selection. *In-silico* functional analysis of non-synonymous mutations using PANTHER revealed that, all the 12 amino acid substitutions (10 in cattle and 2 in sheep) did not impair protein function. The phylogeny revealed trans-species evolution, but was species-wise on with consensus sequences.

Kompjuterska analiza gena beta-laktoglobulina kod nekoliko odabranih vrsta sisara

O.M. Momoh, S.T. Vincent, A. Yakubu

Rezime

Cilj ove studije je kompjutersko ispitivanje genetske raznolikosti β-Lg, njihove evolucije i diferencijacije unutar i između određenih vrsta sisara; takođe ispituje prateće efekte polimorfizma na funkcionalnost gena. Ukupno 21 sekvenci β-Lg gena sa odgovarajućim aminokiselinama koje pripadaju 6 vrsta [goveda (4), bivoli (4), ovce (3), koze (3), svinje (3) i konji (4)] su preuzeti iz GenBank (www.ncbi.nlm.nih.gov). Sve sekvence su obrađene do jednake dužine (500bp), i odgovaraju istom regionu. Poravnanje sekvenci, prevod i upoređivanje je urađeno pomoću ClustalW koristeći IUB supstitucije matrice, otvor od 15 i 6.66. Poravnjanje je otkrilo visok polimorfizam sekvenci. Dxy izveden korišćenjem p-distance otkrio je da su kod ovaca i koza postojale najmanje distance od 0.05 sa maksimalnim rastojanjem od 0,65 između koza i konja. Hipoteza o strogoj neutralnosti ($d_N = d_S$) je odbijena jer je evolucija alelske sekvence pokrenuta i vođena kako prečišćavanjem tako i pozitivnom selekcijom. Samo u slučaju svinja i bivola, su vođeni pozitivnom selekcijom. Kompjuterska funkcionalna analiza ne-sinonimnih mutacije korišćenjem PANTHER otkrila je da nijedna od 12 supstitucija amino kiselina (10 kod goveda i 2 kod ovaca) ne narušava funkciju proteina. Prema tome, nijedan analizirani SNP u ovoj studije nema štetan efekat i može se tolerisati od strane odgajivača pri izboru grla za poboljšanje proizvodnje mleka.

References

BERGSTROM T., GYLLENSTEN U. (1995): Evolution of the MHC class II polymorphism: the rise and fall of class II gene function in primates'. Immunological Reviews, 143, 14–31.

BERRY S.D., LOPEZ-VILLALOBOS N., BEATTIE E.M., DAVIS S., ADAMS L.F., THOMAS N.L., ANKERSMIT-UDY A.E., STANFIELD A.M., LEHNERT K., WARD H.E., ARIAS J.A., SPELMAN R.J., SNELL R.G. (2010): Mapping QTL for concentration of β-Lg in milk, and the effect of β-Lg variants on the composition of milk from Holstein-Friesian x Jersey crossbred cows. New Zealand Veterinary Journal, 58, 1, 1-5.

BRUNHAM L.R., SINGARAJA R.R., PAPE T.D., KEJARIWAL A., THOMAS P.D., HAYDEN M.R. (2005): Accurate prediction of the functional significance of single nucleotide polymorphisms and mutations in the ABCA1 gene. PLOS Genetics, 1: e83.

CHIN D. (1998): Characterisation of genetic variants of milk proteins that are not identified by electrophoresis. Ph.D thesis, Department of Animal Science, Montreal University, Quebec, Canada. 174pp

DEKKER J.C.M., HOSPITAL F. (2002): The use of molecular genetics in the improvement of agricultural population. Nature Reviews Genetics, 3, 22-32.

FELSENSTEIN J. (1985): Confidence limits on phylogenies: An approach using the bootstrap. Evolution, 39, 783–791

HEDRICK, P.W., KIM, T.J. (2000): Genetics of complex polymorphisms: parasites and the maintenance of the MHC variation. In: Singh, R.S., Krinbas, C.B. (eds). Evolutionary genetics: from molecules to morphology. Cambridge University Press, Cambridge, p 204 – 234.

HOFFMAN J.R., FALVO, M.J. (2004): 'Protein – Which is best?'Journal of Sport Science and Medicine, 3, 118 - 130

KONTOPIDIS G., HOLT C., SAWYER L. (2003): Invited review, β-Lg: binding properties, structure and functions. Journal of Dairy Science, 87, 785-786.

KOSAKOVSKY POND S.L. (2011): Natural selection in coding sequences manual. Division of disease and Biomedical Informatics, Department of Medicine, University of California san Diego. 119pp.

LARKIN M.A., BLACKSHIELDS G., BROWN N.P., CHENNA R., MCGETTIGAN P.A., MCWILLIAM H., VALENTIN F., WALLACE I.M., WILM A., LOPEZ R., THOMPSON J.D., GIBSON T.J., HIGGINS D.G. (2007): ClustalW and ClustalX version 2.0. Bioinformatics 23, 21, 2947-8

MARINI N J, THOMAS P D, RINE J (2010): The use of orthologous sequences to predict the impact of amino acid substitutions on protein function. PLOS Genetics 6, e1000968.

NEI M., GOJOBORI T. (1986): Simple methods for estimating the numbers of synonymous and non-synonymous nucleotide substitutions. Molecular Biology and Evolution, 3, 5, 418 – 26.
NG-KWAI-HANG K.F., HAYES J.F., MOXLEY J.E., MONARDES H.G. (1986): Relationship between milk protein polymorphism and major milk constituents in Holstein-Friesian Cows. Journal of Dairy Science, 69, 22-26
SIMM G. (1998): Genetic improvement of cattle and sheep. Farming Press, Ips witch, UK.
STAIGER E.A. (2007): Genetic markers for increasing milk production in sheep. M.Sc. thesis: Department of Animal Science, Faculty of Graduate College, Oklahoma State University. 112pp.
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, 2731-2739.
TARIQ A.M., AL-SHAMMARI A.S., AL-MUAMMAR M.N., ALHAMDAN A.A. (2013): Evaluation and identification of damaged SNPs in COL1A1 gene involved in osteoporosis. Archive of Medical Science, 9, 5, 899-905.
THOMAS P.D., CAMPBELL M.J., KEJARIWAL A., MI H., KARLAK B., DAVERMAN R., DIEMER K., MURUGANUJAN A., NARECHANIA A. (2003): PANTHER: a library of protein families and subfamilies indexed by function. Genome Research, 13, 9, 2129-41.
TOLL-RIERA M., LAURIE S., ALBA M.M. (2011): Lineage-specific variation in intensity of natural selection in mammals Molecular Biology and Evolution. 28, 383-398.
VINCENT S.T., MOMOH O.M., YAKUBU A. (2014): Bioinformatics analysis of beta-casein gene in some selected mammalian species. Research Opinions in Animal and Veterinary Sciences, 4, 10, 564 – 570.
YAKUBU A., SALAKO A.E., DE DONATA M., TAKEET M.I., PETERS S.O., ADEFENWA M.A., OKPEKU M., WHETO M., AGAVIEZOR B.O., SANNI T.M., AJAYI O.O., ONASANYA G.O., EKUNDAYO O.J., ILORI B.M., AMUSAN S.A., IMUMORIN I.G. (2013A): Genetic diversity in Exon 2 at the MHC DQB1 locus in Nigerian indigenous goats. Biochemical Genetics, 51, 954-966.
YAKUBU A., MUSA-AZARA I.S., YAKUBU B.N.S., DAIKWO S.I., VINCENT S.T., MOMOH O.M., DIM N.I. (2013B): Biocomputational Genome-Wide Analysis of micro RNA genetic variability in some vertebrates. GENETIKA, 45, 3, 799-810.
ZHANG J., NIELSEN R., YANG Z. (2005): Evaluation of an improved branch-site likelihood method for detecting positive selection at the molecular level. Molecular Biology and Evolution, 22, 2472–247.

Received 27 April 2015; accepted for publication 21 July 2015