JOURNAL OF SCIENCE TECHNOLOGY AND EDUCATION 14(1), JANUARY, 2026
E-ISSN: 3093-0898, PRINT ISSN: 2277-0011; Journal homepage: www.atbufstejoste.com

Bioinformatics Analysis of Non-Synonymous Single Nucleotide Polymorphisms Substitution in PRL Gene of Goats

[1]Halilu, A., [2]Yakubu, A. [3]Musa, I. S., [4]Gambo, D., [5]Henry, A. J.
[1&5]*Department of Animal Science, Faculty of Agriculture, University of Calabar, Calabar, Nigeria.*
[2,2&4]*Department of Animal Science, Faculty of Agriculture, Nasarawa State University, Keffi, Shabu-Lafia Campus, Nigeria*

ABSTRACT

The present investigation aimed at identifying deleterious non-synonymous single nucleotide polymorphisms (nsSNPs) in prolactin gene of goats using a bioinformatics analysis. Amino acid sequence data of the PRL proteins of goats were retrieved from the National Centre for Biotechnology Information (NCBI) database. Bioinformatics prediction algorithms used for the detection of deleterious nsSNPs were SIFT, PANTHER, SNPs & Go and PhD-SNP program. A total of 5 nsSNPs were obtained from the aligned sequences of PRL of goats, out of which three (M111K, H127Y and H127P) variants were predicted to be deleterious by three out of the four algorithms. The substitution (M111K) in goats was found to decrease protein stability while H127Y and H127P were found to increases stability. Further confirmatory analysis also revealed that these variants including the Cmutant were highly deleterious as there were marked differences between them and the native protein in terms of physico-chemical properties, total free energy, interacting residues and secondary structure. These may distort PRL protein structural landscape and function. The phylogenetic tree revealed species-wise clustering of the PRL sequences. Impact of the identified deleterious nsSNPs (M111K, H127Y and H127P) on goat productivity and health should be investigated in further studies. The deleterious nsSNPs when validated in large populations using wet lab experimental protocols could be important biological markers for to increase meat and milk production.

ARTICLE INFO
Article History
Received: August, 2025
Received in revised form: October, 2025
Accepted: January, 2026
Published online: January, 2026

KEYWORDS
Bioinformatics Analysis, Non-Synonymous Single Nucleotide, Polymorphisms Substitution, PRL Gene, Goats

INTRODUCTION

Goats (Capra hircus) are recognized as a significant livestock species. They are extensively dispersed throughout various ecological zones and are sometimes referred to as "poor men's banks" since they provided rural residents with a source of income and served as a sign of prosperity. In Nigeria, traditional farmers hold the majority of goats (80%) (Bitto, 2008). In developing nations like Nigeria, they make up a sizable amount of the impoverished farmers' financial resources (Kurnianto *et al*., 2013). Goats have a higher rate of multiplication in terms of twinning and short generation intervals when compared to large ruminants like cattle. Goats are a popular choice for reducing poverty because of their capacity for reproduction. Products from goats include fiber, milk, meat, and skin. Additionally, biomedical research uses goats as models (Kon *et al*., 2013; Faisal *et al*., 2013). Dogs are raised for milk, whereas bucks are bred for meat, making them dual-purpose animals (Latif *et al*., 1987).

Goat production and productivity are still at a low level in Nigeria, despite their significance. For the animals to perform better, this necessitates coordinated efforts. Genetic improvement by molecular approaches is one such method (de Lima *et al*., 2020). Current developments in molecular genetics are making it

Corresponding author: Halilu, A.
abdhalilu@gmail.com
Department of Animal Science, Faculty of Agriculture, University of Calabar, Calabar, Nigeria.

JOURNAL OF SCIENCE TECHNOLOGY AND EDUCATION 14(1), JANUARY, 2026
E-ISSN: 3093-0898, PRINT ISSN: 2277-0011; Journal homepage: www.atbufstejoste.com

possible to identify specific genes or candidate genes that have a significant impact on traits of economic importance. Studying the genetic make-up of individuals at the DNA level and molecular genetics has given scientists the tools for opportunities for genetic improvement.

Molecular genetics and DNA analysis of an individual's genetic composition have provided scientists with instruments for potential genetic advancement. According to Tambasco *et al*. (2003), molecular genetics techniques have already led to the identification of genetic markers associated with QTL and a number of genes that have a significant impact on several intriguing quantitative traits. When used in conjunction with conventional selection techniques, marker assisted selection (MAS) may hasten the evolution of economically significant traits (Womack, 2005).

The prolactin (PRL) gene affects the production of milk during lactation. According to research, goat milk yield and content are related to genetic changes in prolactin. This gene plays a role in pituitary gland development and hormone regulation, impacting growth and lactation. In Nigerian goats, studies have identified genetic variability in these genes, which could be useful for selective breeding programs aiming to improve milk and meat production. PRL gene polymorphisms are linked to milk yield, lactation length, and milk composition. Nigerian goat breeds show significant genetic variability, which can be leveraged for breeding programs. By screening these proteins for harmful amino acid alterations, scientists will be able to remove SNPs (single nucleotide polymorphisms) that have detrimental impacts on biological function and may have a number of detrimental implications on milk production.

MATERIALS AND METHODS

The Study Area

This experiment was carried out at the Faculty of Agriculture, Nasarawa State University, Keffi, Shabu – Lafia Campus. Lafia is located on latitude 08^0 35″ and longitude 08^0 33″. It is geographically located in Guinea Savanna Zone of North Central Nigeria. It has the mean maximum monthly temperature of 35.06^0 C and mean minimum monthly temperature of 20.16^0 C with a mean monthly relative humidity of 74%. The annual rainfall is about 168.90 mm (NIMET, Faculty of Agriculture, Lafia, 2022).

Sequence retrieval

The amino acid sequence data on goat PRL gene was retrieved from the website of the National Centre for Biotechnology Information (NCBI). The amino acid sequence alignment of the species was carried out using Multalin software (http://multalin.toulouse.inra.fr/multalin/) to obtain nonsynonymous amino acid substitutions.

Bioinformatics Analyses

Sequence Alignment and Interpretation

Sequences arrangement and interpretation of these genes' quality of the selected goat breeds were completed with ClustalW as depicted by Larkin *et al*. (2007) utilizing IUB substitution grid, hole open punishment of 15 and hole augmentation punishment of 6.66.

Substitution Examination

General extent of non-synonymous substitution per non synonymous site (dN) and the quantity of synonymous substitutions per synonymous site (dS) of the concluded amino successions of these genes was registered by bootstrap technique (1000 duplicates) utilizing the changed Nei-Gojobori (expected progress/transversion inclination = 2) strategy (Zhang *et al*., 1998). The proportion of nonsynonymous to synonymous disparity (dN/dS) was tried for takeoff from the impartial desire for solidarity utilizing the codon-based Z-appropriation as by adjusted Nei-Gojobori, applying Jukes-Cantor rectification.

Functional Prediction of Nonsynonymous Amino Acid Substitutions

The functional effects of the non-synonymous SNPs (nsSNPs) of PLR protein in goats was predicted computationally using

Corresponding author: Halilu, A.
✉ abdhalilu@gmail.com
Department of Animal Science, Faculty of Agriculture, University of Calabar, Calabar, Nigeria.

JOURNAL OF SCIENCE TECHNOLOGY AND EDUCATION 14(1), JANUARY, 2026
E-ISSN: 3093-0898, PRINT ISSN: 2277-0011; Journal homepage: www.atbufstejoste.com

PANTHER and PhD-SNP models according to Yakubu *et al*., 2017 and Ugbo *et al*., 2015).

SIFT (Sorting Intolerant from Tolerant): SIFT (https:// sift. bii.a- star. edu. sg/) (Sim *et al*., 2012) is a power tool used to determine whether a change in amino acid substitution alters the protein function based on sequence homology and the physical characteristics of amino acids. The rsIDs of nsSNPs from NCBI's dbSNP database were submitted as query sequences to SIFT and multiple alignment information was used to analyze tolerated and deleterious substitutions in every position of the query sequence. The result provides nsSNPs as deleterious or tolerated with a SIFT score. A score ≤ 0.05 indicates deleterious and those > 0.05 indicates tolerated. The deleterious nsSNPs were further analyzed to identify the damaging ones.

SNPs & Go server: The disease relationship with the studied SNPs was analyzed using this online web server (http:// snps. Biofo ld. org/ snps- and- go/ snps- and- go. html) (Calabrese *et al*., 2009). The result is based on the combination of Panther result, PHDSNP result, and SNPs&GO result. It predicts whether the mutation is disease-related or neutral, the reliability index (RI), and disease probability.

Protein Stability Prediction

The prediction of the effects of nsSNPs on protein stability was done using I-Mutant2.0 (Capriotti *et al*., 2005). The free energy change (DDG) between the mutant and native proteins was predicted by I-Mutant2.0.

Physical and Chemical Properties of the PRL Protein of Goat

PEPSTATS (Madeira *et al*., 2019) was used for the computation of various physical and chemical properties of these genes using amino acid sequences. The parameters that were computed included molecular weight, average residue weight, isoelectric point, charge, A280 molar extinction coefficients and A280 extinction coefficients 1mg/ml.

Prediction of secondary structure of PRL gene

ExPASy's SOPMA apparatus was utilized to foresee secondary structure of amino corrosive groupings of these genes. SOPM was improved to SOPMA and can anticipate "69.5% of amino acids for a 3-state portrayal of the optional structure" (a helix, b sheets and loop). SOPMA predicts auxiliary structure by agreement forecast from various arrangements.

Tertiary Structure Prediction

Phyre2 server (Kelley *et al*., 2015) was used to construct the structural models of PLR gene protein of goat. The server was use for the alignment of hidden Markov models via HHsearch (Soding, 2005) to improve the accuracy of alignment and rate of detection. Tm-Align was used to calculate Tm-scores and Root Mean Square Deviation (RMSD) (http://zhanglab.ccmb.med.umich.edu/TM-align/) as earlier reported (Yakubu *et al*., 2018). Structural similarities of alternative protein models of both species were quantified by the template used in the prediction. After viewing the models using PyMOL (DeLano, 2006), the position of each mutation on the protein structure was labelled using the Mutagenesis option.

Validation of Protein Structures

The proposed PRL gene protein structures of goats was validated with ERRAT statistical software (Sippl, 1995).

Molecular Dynamic Simulation

Generalized Born (GB) models, which was executed through The Bluues server, was used to find electrostatic differences in structures between the native and the mutant alleles of goats. The server was employing the program Bluues to execute electrostatic calculations for single-atomic structures and provides options for point mutations (Walsh *et al*., 2012). The conditions set were: Solvent probe radius (1.4); Minimum atomic radius (1.0); Outer dielectric constant (78.5); Salt radius (2.0); Inner dielectric constant (4.0); Ionic strength (0.150) and Temperature (Kelvin) (298.95).

Corresponding author: Halilu, A.
abdhalilu@gmail.com
Department of Animal Science, Faculty of Agriculture, University of Calabar, Calabar, Nigeria.

JOURNAL OF SCIENCE TECHNOLOGY AND EDUCATION 14(1), JANUARY, 2026
E-ISSN: 3093-0898, PRINT ISSN: 2277-0011; Journal homepage: www.atbufstejoste.com

Protein-Protein Interaction

In order to predict the solvent accessibility and secondary structures in the 3D structure (Porollo and Meller, 2007) of PRL protein of goats, the web-based SPPIDER Solvent accessibility-based Protein-Protein Interface Identification and Recognition server was used. SOPMA (Combet *et al*., 2000) was used to predict the secondary structure of the PRL gene protein. In order to predict the functional partners of PRL gene protein, the STRING v11 software was be used (Szklarczyk *et al*., 2019).

Phylogenetic Trees Analysis

Neighbor-Joining NJ trees was constructed on the basis of genetic distances, depicting phylogenetic relationships among the PRL gene amino sequences of the goats. The evolutionary distances was computed using the Poisson correction method. The reliability of the trees was calculated by bootstrap confidence values, with 1000 bootstrap iterations using MEGA 11 software (Tamura *et al*., 2013).

RESULTS AND DISCUSSION

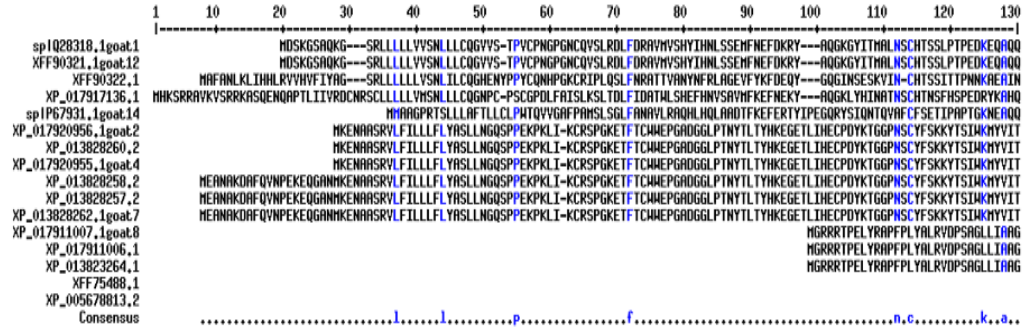
Alignment of PRL Sequence

Results of the alignment showed that five correct polymorphic sites (M111K, H127Y, H127P, V138A and R155L) were obtained from the alignment of the deduced amino acid sequences of PRL gene of goats (Figure 1). The use of polymorphisms within genes is fast gaining attention as a complement to the current methods of selection because of their association with traits of interest in animals (Yakubu *et al*., 2017; Zou *et al*., 2020). Biological insights into the effect of pathogenic missense amino acid variants can be gained by analyzing the relationship between point mutations and protein structures (Iqbal *et al*., 2020).

Prediction of Deleterious nsSNPs by SIFT, PANTHER, SNPs & Go and PhD-SNP Program

A total of 5 nsSNPs (Table 1) were selected for SIFT, PANTHER, SNPs & Go and PhD-SNP analysis for PRL. According to SIFT, the considered deleterious nsSNPs score is 0.05 or below. Among the 5 nsSNPs, 3 nsSNPs were predicted as damaging by SIFT tool whereas the remaining 2 nsSNPs were predicted as "tolerated." According to SNPs&GO, 1 nsSNP was associated with diseases while the remaining 4 were neutral. Moreover, via the PANTHER software tool, 3 nsSNPs were predicted as probably damaging and 1 was probably benign and 1 was not scored. The PhD-SNP tool predicted 3 neutral nsSNPs and 2 nsSNPs were diseases associated (Table 1).

There was a consensus by majority of the algorithms, in the prediction of variants M111K, H127Y, and H127P as being deleterious. These three variants were therefore collectively referred to as 'Cmutant' for further confirmatory analysis. The differences in prediction capabilities of the four algorithms used in the present study may be connected with their differing alignment procedures. The consensus in the prediction of some of the variants as being deleterious may be a pointer to their pathological phenotypic consequences.



Corresponding author: Halilu, A.
abdhalilu@gmail.com
Department of Animal Science, Faculty of Agriculture, University of Calabar, Calabar, Nigeria.

JOURNAL OF SCIENCE TECHNOLOGY AND EDUCATION 14(1), JANUARY, 2026
E-ISSN: 3093-0898, PRINT ISSN: 2277-0011; Journal homepage: www.atbufstejoste.com

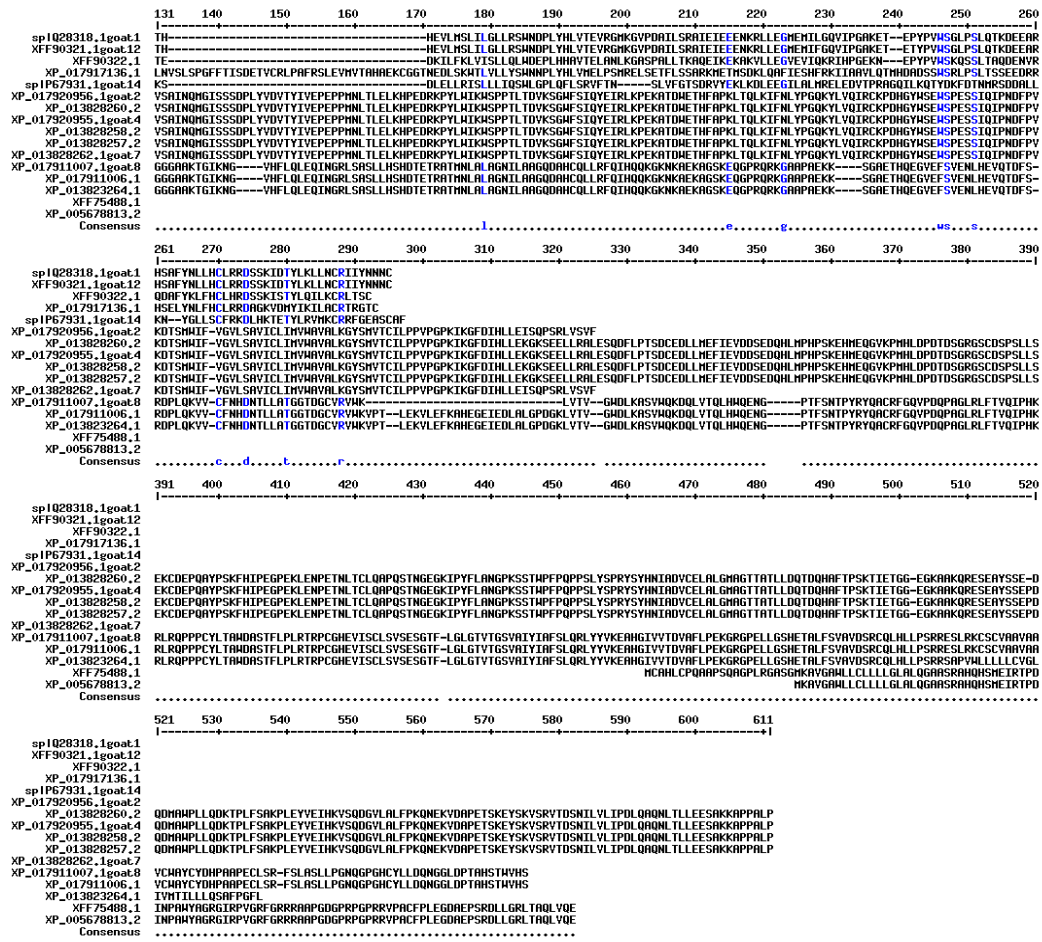


Figure 1: The aligned PRL gene sequences

Table 1: The effect of amino acid variants on the function of PRL protein of goats using SIFT, PANTHER, SNPs & Go and PhD-SNP.

	SIFT		PANTHER		SNPs & Go		PhD-SNP	
Variant	Prediction[1]	Score	Prediction[2]	Time	Prediction[3]	Score	Prediction[4]	Score
M111K	D	0.00	P	726	N	0.340	D	0.507
H127Y	D	0.00	P	591	N	0.433	D	0.541
H127P	D	0.01	P	456	D	0.862	D	0.872
V138A	B	1.00	N	-	N	0.053	N	0.106
R155L	B	0.15	B	1	N	0.199	N	0.243

M = methionine; K = lysine; H= histidine; Y = tyrosine; P = proline; A = alanine; V = valine; R = arginine; L = leucine.

1 D = affect protein function; B = tolerated

2 P = probably damaging; B = probably benign; N = not score

3 N = neutral; D = deleterious

4 N = neutral; D = deleterious

Corresponding author: Halilu, A.
abdhalilu@gmail.com
Department of Animal Science, Faculty of Agriculture, University of Calabar, Calabar, Nigeria.

JOURNAL OF SCIENCE TECHNOLOGY AND EDUCATION 14(1), JANUARY, 2026
E-ISSN: 3093-0898, PRINT ISSN: 2277-0011; Journal homepage: www.atbufstejoste.com

Protein Stability Prediction by I-Mutant 3.0

I-Mutant 3.0 analysis of the nsSNPs for PRL revealed that two (H127Y and H127P) of the three deleterious nsSNPs increased the stability of the PRL protein as shown by its score, which was > 0 for every mutation while M111K with DDG < 0 decreased the stability of the PRL protein. Table 2 displays the free energy change (ΔΔG) values, along with predictions and relative indexes. Three variants were evolutionarily conserved indicating their role in protein structural stability except H127Y and H127P. Only M111K nsSNP was found to reduce the stability of PRL revealed by the negative free energy change values as predicted in I-Mutant 3.0.

This indicates they may have an impact on the folded structure of the protein. According to literature evidence, both deleterious SNPs and mutations are frequently found in the helix and coil regions and not usually in turns (Kucukkal *et al*., 2015). The high and average reliability index obtained in the present study when these variants were subjected to further stability and pathological tests indicates that they could be disease-related mutations.

Physico-Chemical Properties of the Native and Mutant Proteins

The physico-chemical properties of the native protein and substitutions M111K, H127Y, H127P and Cmutant are shown in Table 3. The native proteins varied with others in molecular weight, average residue weight, isoelectric point, charges and A280 extinction coefficient. The physico-chemical properties revealed reduction in the molecular weights and average residue weight of substitutions M111K and H127P while the H127Y and Cmutant revealed increased molecular weights and average residue weight in PLR protein. This reduction molecular weights and average residue weight of substitutions may destabilize and alter the conformational dynamics of the PLR protein. Knowledge of protein characteristics is imperative.

Table 2: I-Mutant prediction based on DDG value and binary classification

Variants	DDG value (kcal/mol)	Predictor	Relative index
M111K	-0.13	Decrease	1
H127Y	0.26	Increase	4
H127P	0.15	Increase	6

Table 3: Physico-chemical properties of the native and mutant proteins

	Variants				
Parameters	**Native protein**	**M111K**	**H127Y**	**H127P**	**Cmutant**
Molecular weight	25773.64	25770.62	25799.68	25733.62	25796.66
Average residue weight	112.55	112.54	112.66	112.37	112.65
Isoelectric point	6.6805	6.8859	6.5929	6.5930	6.8169
Charge	1.0	2.0	0.5	0.5	1.5
A280 molar extinction coefficients	22920	22920	24410	22920	24410
A280 extinction coefficients 1mg/ml	0.889	0.889	0.946	0.891	0.946
TNNCR	-26	-26	-26	-26	-26
TNPCR	23	24	23	23	23
Instability index (II)	57.72	55.43	57.72	57.72	53.26
Aliphatic index (AI)	95.33	95.33	95.33	95.33	96.20
Gravy	-0.246	-0.272	-0.238	-0.239	-0.239

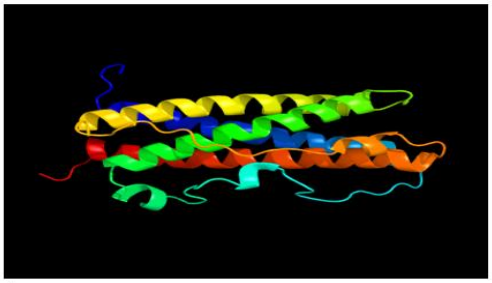
Corresponding author: Halilu, A.
abdhalilu@gmail.com
Department of Animal Science, Faculty of Agriculture, University of Calabar, Calabar, Nigeria.

JOURNAL OF SCIENCE TECHNOLOGY AND EDUCATION 14(1), JANUARY, 2026
E-ISSN: 3093-0898, PRINT ISSN: 2277-0011; Journal homepage: www.atbufstejoste.com

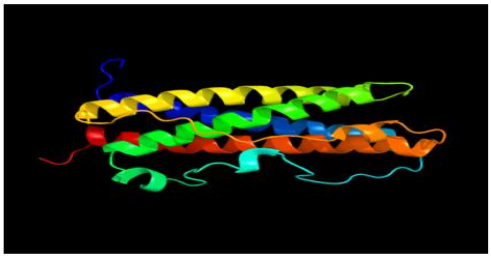
Grand average of hydropathicity; Total number of negative charge residues, Total number of positive charge residues considering the fact that specific changes in gene sequences need to be mapped onto multiple levels of the phenotypic space, from the physico-chemical properties of proteins, biological processes and pathways, up to organismal level (Bershtein *et al*., 2017, Yakubu *et al*., 2021). This will permit proper understanding of the evolutionary consequences of genetic variation through established connection between mutations, phenotypes and fitness (Bershtein *et al*., 2021; Yakubu *et al*., 2021).

The predicted 3D structures of the native and mutant PRL protein of goats

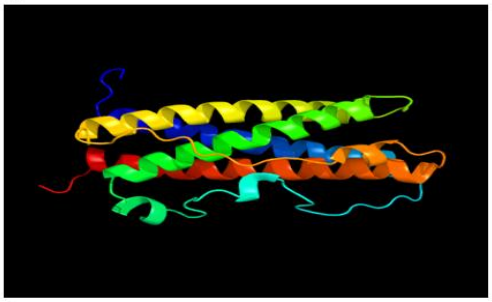
Figure 3 and 4 showed the predicted 3D structures of the native and mutant PRL protein of goats. 190 residues (83% of PRL amino acid sequence) with 100.0% confidence level have been modelled by the single highest scoring template for the native protein. Also, 190 residues of M111K, H127Y and H127P variants including the Cmutant have been modelled with 100.0% confidence level (Figure 2).



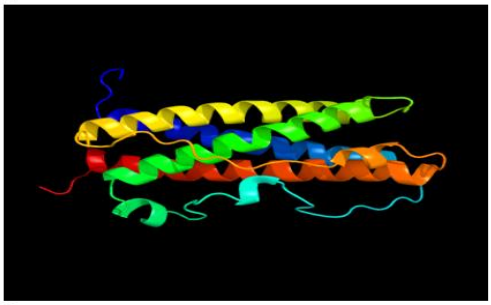
A



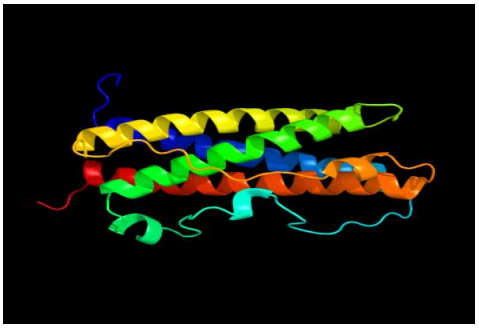
B



C



D



E

Figure 3. The predicted 3D structures of the native and mutant PRL proteins of goats

a: native Prl protein
b: prl protein due to mutant M111K (Methionine to Lysine at position 111)
c: prl protein due to mutant H127Y (Histidine to Tyrosine at position 127) d: prl protein due to mutant H127P (Histidine to Proline at position 127)
d: Prl protein due to Cmutant (Methionine, Tyrosine and Proline combined at positions 111 and 127) with 99.5 and 99.4% confidence. There was an indication of the correct prediction of the PRL 3D structures of the native and three PRL mutants.

Corresponding author: Halilu, A.
abdhalilu@gmail.com
Department of Animal Science, Faculty of Agriculture, University of Calabar, Calabar, Nigeria.

JOURNAL OF SCIENCE TECHNOLOGY AND EDUCATION 14(1), JANUARY, 2026
E-ISSN: 3093-0898, PRINT ISSN: 2277-0011; Journal homepage: www.atbufstejoste.com

TM Score for PRL

Results of TM alignment are presented in Table 4. A TM score of 1.000 and RMSD score of 0.00 was recorded for all the mutants of PRL protein. Based on TM and RSMD values, mutants' model had propensity for more alteration forces in PRL. However, there is every tendency that these substitutions as well as their Cmutants could affect protein conformation and biological roles by virtue of their less negative total free energy. According to Yue *et al*. (2005) and Studer *et al*. (2013), any mutation that adds energy (i.e. more than +2 kcal·mol−1; 1 kcal=4.184 kJ) to the folded state is likely to destabilize the structure and make the protein more likely to aggregate in its unfolded form, which could be a factor in some diseases. In a related study, Alberts *et* al. (2002) reported that the native structure or conformation of a protein generally is enhanced by free energy minimization.

Protein quality

The overall quality factor as revealed by ERRAT was 77.095 for the native protein, 71.51% for M111K, 78.21% for H127Y and 78.77% for H127P. However, the quality factor for Cmutant was lower at 73.18%.

Table 4: Tm score for PRL variants

Variants	TM-score	RMSD	Aligned length	Sequence identity
M111K	1.0000	0.00	190	0.995
H127Y	1.0000	0.00	190	0.995
H127P	1.0000	0.00	190	0.995
Cmutant	1.0000	0.00	190	0.979

Energy differences between the Native Protein and the Mutants

The molecular dynamic simulation revealed that the native protein and the variants M111K, H127Y, H127P and Cmutant differed in born self-energy, coulomb energy, electrostatic solvation energy and total energy for PRL (Table 5).

Protein Interaction

The interacting residues of the native protein and substitutions (M111K, H127Y, and H127P) including the Cmutant are shown in Table 6. The varying interacting residues of the native protein and substitutions in both proteins may suggest their differential reactions with other molecules with regard to the biological functions of protein. It has been reported that if a mutation alters the flexibility of protein in comparison with its native structure, there is every probability that the capacity of the protein to regulate slight conformational changes will be affected. This may invariably affect protein function (Islam *et al*., 2019). A crucial component to study the structural behaviour of the protein is the propensity of the secondary structural content. Abrusán and Marsh (2016) stated that missense mutations that change secondary structure are more likely to be pathogenic; and concluded that the inclusion of predicted secondary structure changes shows significant utility for improving upon state-of-the-art pathogenicity predictions.

Table 5: Energy differences between the native protein and the mutants

	Born self-energy (kJ)	Coulomb energy (kJ)	Electrostatic solvation energy (kJ/mol)	Total energy (kJ/mol)
Native	-842566	-52725.87	-1235.12	-52692.96
M111K	-8434.76	-52820.83	-1157.75	-52710.53
H127Y	-8417.17	-52712.42	-1157.75	-52710.53
H127P	-8402.21	-52565.70	-1240.26	-52536.56
Cmutant	-8426.26	-52807.38	-1159.99	-52699.35

Corresponding author: Halilu, A.
abdhalilu@gmail.com
Department of Animal Science, Faculty of Agriculture, University of Calabar, Calabar, Nigeria.

JOURNAL OF SCIENCE TECHNOLOGY AND EDUCATION 14(1), JANUARY, 2026
E-ISSN: 3093-0898, PRINT ISSN: 2277-0011; Journal homepage: www.atbufstejoste.com

Table 6: Interacting residues of the native protein and mutants

Variants	Interacting residues
Native	N40,C41,Q42,V43,S44,L45,R46,D47,D50,R51,V53,M54,Q76,G79,I81,T82,M83,A84,L85,N86,S87,S91,S92,R133,G134,M135,V138,D140,K172,E175,P176,Y177,V179,S181,N220,C221,I223,I224,Y225,N226,N227,N228,C229
M111K	N40,C41,Q42,V43,S44,L45,R46,D47,D50,R51,V53,M54,G79,I81,T82,M83,A84,L85,N86,S87,S91,S92,R133,G134,M135,V138,D140,K172,P176,Y177,V179,S181,C221,I224,Y225,N226,N227,N228,C229
H127Y	N40,C41,Q42,V43,S44,L45,R46,D47,D50,R51,V53,M54,Q76,G77,G79,I81,T82,M83,A84,L85,N86,S87,S91,S92,R133,G134,M135,V138,D140,K172,E173,E175,P176,Y177,V179,W180,S181,N220,C221,I223,I224,Y225,N226,N227,N228,C229
H127P	N40,C41,Q42,V43,S44,L45,R46,D47,D50,R51,V53,M54,G79,I81,T82,M83,A84,L85,N86,S87,S91,S92,R133,G134,M135,V138,D140,K172,E173,E175,P176,Y177,V179,W180,S181,N220,C221,I223,I224,Y225,N226,N227,N228,C229
Cmutant	N40,C41,Q42,V43,S44,L45,R46,D47,D50,R51,V53,M54,G79,I81,T82,M83,A84,L85,N86,S87,S91,S92,R133,G134,M135,V138,D140,K172,P176,Y177,V179,S181,C221,I224,Y225,N226,N227,N228,C229

Secondary Structure Prediction using SOPMA

Table 7 summarizes the secondary structures prediction of PRL protein using SOPMA. There were observed variations in the secondary structure of the native protein and the mutants in both proteins. The number of alpha helix in the native (129) differed from those of substitutions M111K (132), H127P (128) and Cmutant (177). The native protein had higher number of extended strand (6) compared to M111K (0), H127Y (0), H127P (2), H127P (0) and Cmutant (127), respectively. However, higher values of random coil were obtained in H127P, (101), H127Y (98), H127P (97), M111K (97), Cmutant (127) compared to the native protein (94).

The results from String 10 server show that the PRL gene interacts with a variety of genes, mainly participating in the control of several cellular activities (Figure 3). PRL is related to PRLH (Prolactin Releasing Hormones) that stimulates the pituitary gland to release prolactin, crucial for milk production, metabolism, and immune function; PRP6 (Pre-mRNA Processing Factor 8); a crucial protein involved in pre-mRNA splicing, acting as a component of the spliceosome's U4/U6-U5 tri-snRNP complex, essential for removing introns from messenger RNA; CSN2 (Casein Beta) is a major milk protein encoded by the CSN2 gene, crucial for infant nutrition, providing amino acids, calcium, and phosphate, and forming casein micelles with kappa-casein; IL2RG (Interleukin 2 Receptor Gamma) crucial for immune cell development and function; IL20 (Interleukin 20) immune responses and cell growth, produced by epithelial cells and leukocytes; GHR (Growth Hormone Receptor) a cell surface protein that binds growth hormone (GH).

Initiating crucial signals for growth, metabolism, and other bodily functions by activating intracellular pathways like JAK2/STAT, primarily in the liver and GH1 (Growth Hormone 1) is the gene and protein responsible for producing human growth hormone (somatotropin), a crucial hormone for normal body growth, bone development, and metabolism, primarily made by the pituitary gland; it stimulates the liver to make IGF-1, which promotes cell growth, and affects protein synthesis and glucose levels, with gene mutations leading to growth deficiency. With respect to the functional association analysis using STRING, the network obtained indicates that substitutions may alter cell cycle controls by disrupting genetic interactions involving PRL. This is particularly with its role in milk production and reproduction.

The STRING network revealed that LHX4 interacts with 21 proteins which include is related to CCDC136 (Coiled-Coil Domain 136) a

Corresponding author: Halilu, A.
abdhalilu@gmail.com
Department of Animal Science, Faculty of Agriculture, University of Calabar, Calabar, Nigeria.

JOURNAL OF SCIENCE TECHNOLOGY AND EDUCATION 14(1), JANUARY, 2026
E-ISSN: 3093-0898, PRINT ISSN: 2277-0011; Journal homepage: www.atbufstejoste.com

protein involved in crucial biological processes, known for its role in embryonic development and male fertility; ISOC1 (Isochorismatase Domain 1) is a conserved protein involved in various cellular processes, notably in DNA repair, inflammation, and cancer development; MNX1 (Motor Neuron And Pancreas Homeobox 1) is a transcription factor gene crucial for motor neuron and pancreas development, influencing cell fate and function, particularly in insulin-producing beta cells; PTF1A (Pancreas Transcription Factor 1A) is essential for the development and differentiation of the pancreas and cerebellum; ISL2 (ISL LIM Homeobox 2), ISL1 (ISL LIM Homeobox 1) involved in neural development, defining subtypes of motor neurons and guiding their connectivity in the spinal cord, and also playing roles in the retina, pineal gland, and potentially cancer and endometriosis, by controlling gene expression and cell differentiation.

SSBP2 (Single-Stranded DNA Binding Protein 2) involved in maintaining genome stability, DNA damage response, and DNA replication, acting as a subunit in an ssDNA-binding complex; NHLH2 (Nesvacumab also called Nerve Growth Factor), LHX3 (LIM Homeobox 3), LDB1 (LIM Domain Binding 1), LDB2 (LIM Domain Binding 1), ALX3 (ALX Homeobox 3), SSBP3 (Single-Stranded DNA Binding Protein 3), SSBP4 (Single-Stranded DNA Binding Protein 4), GRM6 (Glutamate Metabotropic Receptor 6), POU1F1 (POU Class 1 Homeobox 1) and ZBTB34 (Zinc Finger And BTB Domain 34) (Figure 11). These proteins showed higher interaction based on the confidence score generated by experimental validation and text mining. Due to the nsSNP variants in LHX4, amino acid alterations may also have an impact on the function of the interacting molecules.

Table 7: Secondary structure prediction of PRL using SOPMA

Secondary Structure	Variants				
	Native protein	M111K	H127Y	H127P	Cmutant
Alpha helix	129 (56.33%)	132 (57.64%)	129 (56.33%)	128 (55.90%)	127 (54.07%)
3_{10} helix	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)
Pi helix	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)
Beta bridge	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)
Extended strand	6 (2.62%)	0 (0.00%)	2 (0.87%)	0 (0.00%)	2 (13.00%)
Beta turn	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)
Bend region	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)
Random coil	94 (41.05%)	97 (42.36%)	98 (42.79%)	101 (44.10%)	93 (33.02%)
Ambiguous states	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)
Other states	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)

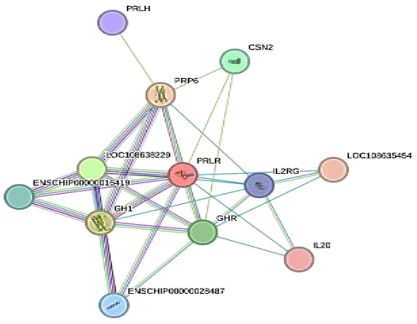


Figure 3: Functional association network of PRL with other important proteins in the cellular system

Phylogeny Trees for PRL

The phylogenetic tree revealed species-wise clustering. The goat sequences were more related to one another (Figure 4). The phylogenetic tree derived in the present study clearly revealed that clustering was largely species-wise. The species-wise clustering may be explained by species-specific residues (Takahashi and Nei, 2000), and such sequence patterns may be the result of gene conversion and balancing selection. The phylogenetic tree also

Corresponding author: Halilu, A.
abdhalilu@gmail.com
Department of Animal Science, Faculty of Agriculture, University of Calabar, Calabar, Nigeria.

suggested that trans-species evolution had not occurred in the both genes.

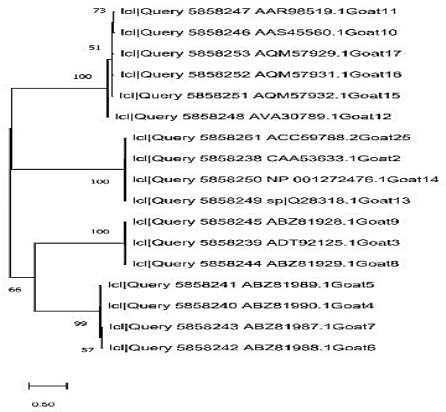


Figure 4: Neighbour joining tree of the PRL amino acid sequences of goats and other ruminant species

CONCLUSION

The bioinformatics analysis of SNPs study identified three potential high risk deleterious nsSNPs of PRL as well as their Cmutants, and the variants are likely to have an effect on the protein structure and/or function. When this is experimentally confirmed in future web lab and pathogenic population-based association studies, such information may be exploited in selection against M111K, H127Y and H127P substitutions to increase meat and milk production.

REFERENCES

Abrusán, G. and Marsh, J.A. (2016). Alpha helices are more robust to mutations than beta strands. *PLOS Computational Biology*, 12(12): e1005242. https://doi.org/10.1371/journal.pcbi.1005242

Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K. and Walter, P. (2002). Molecular Biology of the Cell. 4th edition. New York: Garland Science; 2002.

Bershtein, S., Daniel, K. and Dan, M. (2021). Predicting 3D protein structures in light of evolution. *Department of Life Sciences, Ben-Gurion University of the Negev, Beer-Sheva, Israel.* https://doi.org/10.1038/s41559-021-01519-8.

Bershtein, S., Serohijos, A. W. and Shakhnovich, E. I. (2017). Bridging the physical scales in evolutionary biology: from protein sequence space to fitness of organisms and populations. *Current Opinion in Structural Biology*, 42: 31–40.

Bitto, I. I. (2008). Seasonal variation in the histometric characteristics of reproductive organs of pubertal West African dwarf bucks in their native tropical environment. *Journal of Animal Science Morphology*, 26 (2):397 – 401.

Calabrese, R., Capriotti, E., Fariselli, P., Martelli, P. L. and Casadio, R. (2009) Functional annotations improve the predictive score of human disease-related mutations in proteins. Hum Mutat 30(8):1237–1244.

Capriotti, E., Fariselli, P. and Casadio, R. (2005). I-Mutant2.0: Predicting stability changes upon mutation from the protein sequence or structure. *Nucl. Acids Res.*, 33, 306-310.

Combet, C., Blanchet, C., Geourjon, C. and Deléage, G. (2000). Network Protein Sequence Analysis. *TIBS*, 25:147-150.

de Lima, L. G., de Souza, N. O. B., Rios, R. R., de Melo, B. A., dos Santos, L. T. A., Silva, K. de M., Murphy, T. W. and Fraga, A. B. (2020). Advances in molecular genetic techniques applied to selection for litter size in goats (*Capra hircus*): a review. *Journal of Applied Animal Research*, 48(1), 38–44. https://doi.org/10.1080/09712119.2020.1717497.

DeLano, W. L. (2006). The PyMOL Molecular Graphics System (http://www.pymol.org).

Faisal, S. M., Yan, F., Chen, T. T., Useh, N. M., Guo, S., Yan, W., Wang, S., Glaser, A. L., McDonough, S. P., Singh, B. and

Corresponding author: Halilu, A.
abdhalilu@gmail.com
Department of Animal Science, Faculty of Agriculture, University of Calabar, Calabar, Nigeria.

Chang, Y. (2013). Evaluation of a Salmonella vectored vaccine expressing Mycobacterium avium subsp. paratuberculosis antigens against challenge in a goat model. *PLoS ONE*, 8:e70171.

Iqbal, S., Pérez-Palma, E., Jespersen, J. B., May, P., Hoksza, D., Heyne, H. O., Ahmed, S. S., Rifat, Z. T., Rahman, M. S., Lage, K., Palotie, A., Cottrell, J. R., Wagner, F. F., Daly, M. J., Campbell, A. J. and Lal, D. (2020). Comprehensive characterization of amino acid positions in protein structures reveals molecular effect of missense variants. *Proceedings of the National Academy of Sciences*, 117: 28201-28211; DOI: 10.1073/pnas.2002660117.

Islam, M. J., Parves, M. R., Mahmud, S., Tithia, F. A. and Reza, M. A. (2019). Assessment of structurally and functionally high-risk nsSNPs impacts on human bone morphogenetic protein receptor type IA (BMPR1A) by computational approach. *Computational Biology and Chemistry*, 80: 31-35.

Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N. and Sternberg, M. J. E. (2015). The Phyre2 web portal for protein modeling, prediction and analysis. *Nat. Protoc*., 10, 845-858.

Kon, E., Filardo, G., Robinson, D., Eisman, J. A., Levy, A., Zaslav, K., Shani, J., Altschuler, N. (2013). Osteochondral regeneration using a novel aragonite-hyaluronate bi-phasic scaffold in a goat model. *TraumatolArthrosc*., 22 (6): 1452–1464.

Kucukkal, T. G., Petukh, M., Li, L. and Alexov, E. (2015) Structural and physicochemical effects of disease and non-disease nsSNPs on proteins. *Curr Opin Struct Biol*., 32:18–24.

Kurnianto, E., Suto po, S., Purbowati, E., Setiatin, E. T., Samsudewa, D. and Permatasari, T. (2013). Multivariate Analysis of Morphological Traits of Local Goats in Central Java, Indonesia. *Iranian Journal of Applied Animal Science*, 3(2): 361 - 367.

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. and Higgins, D. G. (2007). Clustal W and Clustal X version 2.0. *Bioinformatics*, 23 (21): 2947 – 8.

Latif, M. G., Abdelsalam, M. M. and Abdelaziz, N. M. (1987). Meat production characteristics of Egyptian Baladi and Angora goats. *Meat Sciences*, 20:211–6.

Madeira, F., Park, Y., Lee, J., Buso, N., Gur, T., Madhusoodanan, N., Basutkar, P., Tivey, A. R. N., Potter, N. C., Finn, R. D. and Lopez, R. (2019). The EMBL-EBI search and sequence analysis tools APIs. *Nucleic Acids Research*, 47 (1): 636 – 641. https://doi.org/10.1093/nar/gkz268.

NIMET (2022). Nigerian Meteorological Agency. Seasonal climate prediction (SCP) for Nigeria, Faculty of Agriculture, Lafia.

Porollo, A. and Meller, J. (2007). Prediction-based fingerprints of protein-protein interactions. PublMed, 66 (3): 630-45. doi: 10.1002/prot.21248.

Sim, N. L., Kumar, P., Hu, J., Henikoff, S., Schneider, G. and Ng, P. C. (2012) SIFT web server: predicting effects of amino acid substitutions on proteins. *Nucleic Acids Resources,* 40(W1):W452–W457.

Sippl, M. J. (1995). Knowledge-based potentials for proteins. *Current Opinion in Structural Biology,* 5: 229-235.

Söding, J. (2005). Protein homology detection by HMM-HMM comparison. *Bioinformatics*, 21: 951-960.

Studer, R. A., Dessailly, B. H.and Orengo, C. A. (2013). Residue mutations and their impact on protein structure and function:detecting beneficial and

Corresponding author: Halilu, A.
abdhalilu@gmail.com
Department of Animal Science, Faculty of Agriculture, University of Calabar, Calabar, Nigeria.

pathogenic changes. *Biochemical Journal*, 449, 581–594. doi:10.1042/BJ20121221.

Szklarczyk, D., Gable, A. L., Lyon, D., Junge, A., Wyder, S., Huerta-Cepas, J., Simonovic, M., Doncheva, N. T., Morris, J. H., Bork, P., Jensen, L. L. and von Mering, C. (2019). STRING v11: protein–protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets, *Nucleic Acids Research*, 47: 607–613. https://doi.org/10.1093/nar/gky1131.

Takahashi, K. and 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: 1251–1258.

Tambasco, D. D., Paz, C. C. P., Tambasco-Studart, M., Pereira, A. P., Alencar, M. M., Freitas, A. R., Coutinho, L. L., Packer, I. U. and Regitano, L. C. A. (2003). Candidate genes for growth traits in beef cattle cross Bos taurus × Bos indicus. *Journal of Animal Breeding and Genetics*, 120: 51 - 56.

Tamura, K., Stecher, G., Peterson, D., Filipski, A. and Kumar, S. (2013). MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. *Molecular Biology and Evolution*, 30: 2725- 2729.

Ugbo, S. B., Yakubu, A., Omeje, J. N., Bibinu, B. S., Musa, I. S., Egahi, J. O. and Dim, N. I. (2015). Assessment of genetic relationship and application of computational algorithm to assess functionality of non-synonymous substitutions in DQA2 gene of cattle, sheep and goats. *Open Journal of Genetics,* 5: 145-158. DOI: 10.4236/ojgen.2015.54011.

Walsh, I., Minervini, G., Corazza, A., Esposito, G., Tosatto, S. C. E. and Fogolari, F. (2012).Bluues Server: electrostatic properties of wild-type and mutated protein structures. *Bioinformatics*, 28 (16): 2189-2190.

Womack, J. E. (2005). Advances in livestock genomics: opening the barn door. *Genome Resources*, 15:1699-1705.

Yakubu, A., De Donato, M. and Imumorin, I. G. (2017). Modelling functional and structural impact of non-synonymous single nucleotide polymorphisms of the DQA1 gene of three Nigerian goat breeds. *South African Journal of Animal Science,* 47: 146-156. http://dx.doi.org/10.4314/sajas.v47i2.6.

Yakubu, A., Ogbanwo, F., Akinsola, O. M. and Bunjah Umar, D. S. (2021). *In silico* prediction of the deleterious missense variants in the Ovalbumin gene of ducks

Yakubu, A., Salako, A. E., De Donato, M., Isidahomen, C. E. and Imumorin, I. G. (2018). The effects of non-synonymous single nucleotide polymorphisms (nsSNPs) of DQB1 gene of tropical goats on protein function and stability. *Nigerian Journal of Genetics,* 32: 30-39.

Yue, P., Li, Z. and Moult, J. (2005). Loss of protein structure stability as a major causative factor in monogenic disease. *Journal of Molecular Biolology*, 353: 459–473.

Zhang, Q., Boichard, D., Hoeschele, I., Ernest, C., Eggen, A., Murkve, B., Pfister-Genskow, M., Witte, L. A., Grignola, F. E., Uimari, P., Thaller, G. and Bishop, M. D. (1998). Mapping quantitative trait loci for milk production and health of dairy cattle in a large outbred pedigree. *Genetics*, 149: 1959 – 1973.

Zou, A., Nadeau, K., Wang, P. W. , Lee, J. Y., Guttman, D. S., Sharif, S., Korver, D. R., Brumell, J. H. and Parkinson, J. (2020). Accumulation of genetic variants associated with immunity in the selective breeding of broilers. *BMC Genetics,* 21: 5.

Corresponding author: Halilu, A.
abdhalilu@gmail.com
Department of Animal Science, Faculty of Agriculture, University of Calabar, Calabar, Nigeria.