



MOLECULAR GENETIC VARIATION (POLYMORPHISMS) USING *IN-SILICO* METHODS IN DETERMINING EFFECT OF GDF8 GENE IN CATTLE, SHEEP AND GOATS

¹Garba, A. M.*, ²Mancha, Y. P, ²Bello K. M., Panshak, L. N. and Zakka, T. E.

¹Department of Animal Science, Federal University of Agriculture, Zuru, Kebbi State

²Department of Animal Production,

Abubakar Tafawa Balewa University, Bauchi, Bauchi State, Nigeria

*Corresponding Author's Email: amgarba47@gmail.com GSM: 07062324147

ABSTRACT

The *in-silico* research was carried out to explore the effect of molecular genetic variation of the myostatin (GDF8) gene in cattle, sheep, and goats. A total of twenty-three (23) GDF8 gene nucleotide sequences comprising cattle (4), sheep (8) and goats (11) were retrieved from NCBI, aligned, translated, and compared using ClustalW. The variation in sequence length base pairs (bp) of the GDF8-gene within and among species ranged between 381 bp and 7831 bp. The functional analysis of mutations using SIFT showed that ten amino acid substitutions (S7Y, M14T, N26K, L43M, F139Y, F94L, R48G, A167S, S125T, and P285E) in cattle, seven in sheep (F7Y, L16I, K36E, L43M, N48T, K49T, and T120S), and ten in goats (F7Y, F14I, L16I, H23Y, K36E, K39Q, T42A, L43M, N48T, and L53I) were beneficial. However, three in cattle (Y9C, E46D, and R161T), five in sheep (Y9C, E34G, R45G, T114A, and I119V) and six in goats (V26L, W27R, G37R, T40I, R50G, and G57R) were harmful. The significant value ($p < 0.01$) for the relative ratio of non-synonymous substitutions per non-synonymous site (dN) to the number of synonymous mutations per synonymous site (dS) known as omega ($\omega = dN/dS$), revealed positive selection for all the studied species (1.55 for cattle, 2.71 for sheep and 1.87 for goats). The Tajima's D indicates selective sweep in sheep (-2.39). The phylogenetic trees showed a close relationship among members of the Bovidae family. It can be inferred from the study that cattle, sheep, and goat GDF8 genes are polymorphic genes with numerous mutations being more beneficial. Further research should be conducted to explore the potential applications of these results in selecting desirable traits and the genetic improvement of cattle, sheep, and goats.

Keywords: Molecular, Polymorphisms, Myostatin, Sequence, Gene

INTRODUCTION

The Transforming Growth Factor-Beta (TGF- β) superfamily is a diverse gene collection that encodes various growth factors (Jason, 2020). Over thirty homologs within the TGF- β superfamily have been identified and categorised into different subfamilies, including Bone Morphogenetic Proteins (BMPs), Growth Differentiation Factors (GDFs), Inhibins/Activins, and Nodal/Lefty genes, based on structural and functional similarities (Dushyanth *et al.*, 2016; Satheesh *et al.*, 2016; Vishnu *et al.*, 2016; Guru *et al.*, 2017; Divya *et al.*, 2018). These proteins regulate and mediate fundamental cellular processes (Bhattacharya *et al.*, 2022). The GDF8 gene, situated on bovine chromosome 2 (BTA2) and chromosome 2 in sheep, has been associated with single nucleotide polymorphisms (SNPs) in its coding region and linked to double muscling due to mutation (Aiello *et al.*, 2018; Osman *et al.*, 2021). GDF8 gene functions as a negative regulator



of skeletal muscle mass, muscle growth, and development, leading to various metabolic changes (Osman *et al.*, 2021; Wu *et al.*, 2022). The association of GDF8 gene polymorphisms with several muscle-related traits has been documented in other livestock species, such as cattle (Casas and Kehrl, 2016), chickens (Zhang, 2019) and rabbits (Guo *et al.*, 2016).

Economic traits in livestock are primarily determined by quantitative trait loci (QTL), which explain a portion of the variation in traits of interest. The QTL plays a role in developing increasingly dense genome maps for farm livestock, and studies investigating associations between molecular genetic markers and traits of interest (Haskell *et al.*, 2014). Some quantitative trait genetic investigations in animal populations focus on detecting single nucleotide polymorphisms (SNPs) and their polymorphisms (Yang *et al.*, 2017). It stresses identifying associations between genetic polymorphisms and their impact on specific species that contribute to the economic growth of livestock breeders (Al-Khudhair *et al.*, 2020). Wet laboratory studies investigating mutations in genes of interest are often time-consuming, expensive, and laborious (Ewuola *et al.*, 2018). Computational studies, on the other hand, can efficiently provide useful information to justify and guide further experimental studies (Rasal *et al.*, 2016). Thus, computational predictions have become imperative for evaluating the disease-related consequences of nonsynonymous single-nucleotide variants discovered through exome sequencing (Liu and Kumar, 2013). The study aims to explore the molecular genetic variation of the myostatin (GDF8) gene in cattle, sheep, and goats using in-silico methods.

MATERIAL AND METHODS

GDF8-gene Sequences for cattle, sheep, and goats

A total of twenty-three (23) GDF8 nucleotide sequences comprising cattle (4), sheep (8), and goats (11) were retrieved on 15, 05, 2024 from the National Centre for Biotechnology Information (NCBI) GenBank (www.ncbi.nlm.nih.gov). The sequences have the following GenBank accession numbers: AF320998.1, EF432556.1, JQ711180.1, and AB076403.1 (cattle), EF069435.1, EF069434.1, EF069433.1, AF266758.1, FJ809783.1, FJ809782.1, FJ809781.1, and FJ809780.1 (sheep), AY827576.1, EF432557.1, JN688159.1, KY790422.1, KX171679.1, KU991727.1, KU980201.1, KY828431.1, KY828430.1, KY828429.1, and AY032689.1 (goats).

Sequence Alignment and Translation

Using ClustalW, the GDF8-gene of cattle, sheep, and goats was sequenced, translated, and compared using an International Union of Biochemistry (IUB) substitution matrix, a gap open penalty of 15 and a gap extension penalty of 6.66, as detailed by Larkin *et al.* (2007).

Functional Analysis

A set of homologous and distantly related sequences was collected from the NCBI non-redundant (nr) protein database using BLASTX with an E-value threshold of 0.1 to identify the variants described by Altschul, *et al.*, (1997). *In silico* functional analysis of missense mutations was obtained with (SIFT) as described by Vaser, *et al.* (2016), with a threshold value for <0.5 to predict the functional effect of amino acid substitution as harmful, and neutral of small insertions and deletions. Any variant having a SIFT score below the threshold value of <0.5 is termed DELETERIOUS and a SIFT score above the same threshold value is considered TOLERATED.

Substitution Analysis

The relative proportion of non-synonymous substitution per non-synonymous site (dN) and the number of synonymous substitutions per synonymous site (dS) of the deduced amino acid



sequences were computed by the bootstrapping method (1000 replicates) with the modified Nei-Gojobori (assumed transition/transversion bias = 2) method (Zhang *et al.*, 1998).

The (dN/dS) ratio was tested for departure from the neutral expectation of unity using the codon-based Z-distribution by modified Nei-Gojobori, applying Jukes-Cantor correction. The analysis was conducted in MEGA11 software, as described by Tamura *et al.* (2021).

Tajima's Test of Neutrality

Tajima's test statistics (Tajima, 1989) were estimated. Genes with a Tajima's *D* value below -2 have an excess rare allele, indicating positive selection or a selective sweep. However, genes with a Tajima's *D* value above 2 have an excess of common alleles, suggesting balancing selection as opined by Eckshtain-Levi *et al.* (2018)

Phylogeny Tree Construction

UPGMA trees for the GDF8-gene were constructed using the maximum composite likelihood method and pairwise deletion gap/missing data treatment described by Saitou and Nei (1987). The reliability of the trees was calculated by bootstrapping confidence values (Felsenstein, 1985) with 1000 bootstrap iterations using MEGA 11 software (Tamura *et al.*, 2021). All the nucleotide sequences were trimmed to an equal length of 381 bp, corresponding to the same region of the gene, before generating the trees.

RESULTS AND DISCUSSIONS

Sequence Length Variation of the Cattle, Sheep, and Goats GDF8 gene

The variation in sequence length in base pairs (bp) of the GDF8-gene within and among species ranged between 381 bp and 7831 bp (Table 1).

Table 1. Accession Number, Base Pair, and Sequence Length Variation of the Cattle, Sheep, and Goats GDF8-gene

Specie	Accession Number	Base Pair (bp)	Sequence Length Variation
Cattle	EF432556.1	589	589-7,831
	AB076403.1	6,660	
	AF320998.1	6,691	
Sheep	JQ711180.1	7,831	473-2,603
	EF069433.1	473	
	EF069434.1	473	
	EF069435.1	473	
	FJ809780.1	1,037	
	FJ809781.1	1,037	
	FJ809782.1	1,037	
	FJ809783.1	1,037	
	AF266758.1	2,603	
	Goat	KX171679.1	
KU980201.1		492	
KY790422.1		531	
KY828429.1		573	
KY828430.1		573	
KY828431.1		573	
KU991727.1		584	
EF432557.1		589	
JN688159.1		1,181	
AY827576.1		1,344	
AY032689.1		2,584	

In cattle species, three sequences AB076403.1, AF320998.1, and JQ711180.1 had complete coding sequences (CDS) from DNA each with sequence lengths that are greater than six thousand base pairs (>6000 bp), while the remaining sequences in cattle, sheep and goats are partial coding sequences (partial CDS) whose lengths are less than six thousand base pairs (<6000 bp). Across the species, two sequences (from goats and cattle) had 589 bp, while seven sequences, 1 in cattle and 6 in goats, had similar bp ranging from 531 to 589 bp. There was also a similarity in 2 sequences each in sheep and goat (1,037-1,344 bp). Within species, 4 sequences from sheep (1,037 bp) and 3 sequences from sheep and goats had the same bp of 473 and 573, respectively. The variation in sequence length observed within and across species in this study might be due to differences in the genomic regions, where the sequences were obtained and differences due to complete coding sequences (CDS) or partial CDS (Bibinu *et al.*, 2016). The sequence length variation found in the studied species agreed with the report of Ewuola *et al.* (2018) who reported MSTN gene sequence length variations within and across selected poultry species. Similarly, Bibinu *et al.* (2016) observed similar results with the BMP15 gene in ruminant and non-ruminant animals. The variations might result from evolution and differentiation, DNA duplication, DNA rearrangement, short tandem repeats (STRs), insertions, or deletion of sequences (Yakubu *et al.*, 2014; Vincent, *et al.*, 2014).

Sequence Alignment and Translation

The predicted amino acid sequences showed varying substitution degrees in the studied species. A portion of this is presented in Fig. 1. The dots (.) indicate identical amino acids, the numbers on the right-hand side represent the site number, the question mark (?) indicates missing amino acids, and the asterisk (*) indicates indels.

```
#AF320998.1_Bovine_A RCY----- [189]
#EF432556.1_Bovine_B ----- [189]
#JQ711180.1_Bovine_C KI*VYMLLLI KFLIFRMS?- ----- [189]
#AB076403.1_Bovine_D ...-QTTFAQ ----- [189]
#EF069435.1_Ovine_A ...-KTTFQA GSSTPGTD*S VRCPER*QQR RLLGRR*LP [189]
#EF069434.1_Ovine_B ...-KTTFQA GSSTPGTD*S VRCPER*QQR RLLGRR*LP [189]
#EF069433.1_Ovine_C ...-KTTFQA GSSTPGTD*S VRCPER*QQR RLLGRR*LP [189]
#AF266758.1_Ovine_D SRC-PVSHPM KTL-LIL?-- ----- [189]
#FJ809783.1_Ovine_E ----- [189]
#FJ809782.1_Ovine_F ----- [189]
#FJ809781.1_Ovine_G ----- [189]
#FJ809780.1_Ovine_H ----- [189]
#AY827576.1_Caprine_A KIH-YLLSHS KLF*D?NC*Q ?----- [189]
#EF432557.1_Caprine_B ----- [189]
#JN688159.1_Caprine_C DFH-YVL?-- NLSRETSF-- ----- [189]
#KY790422.1_Caprine_D ----- [189]
#KX171679.1_Caprine_E LIC-YILMAK NK*YMGRFQA W**IAVGAH? ----- [189]
#KU991727.1_Caprine_F TGS-NNLNPT *ALKSKL?-- ----- [189]
#KU980201.1_Caprine_G ...-KTTFQA GSSTPGTD*S VRCPER*QQR ?----- [189]
#KY828431.1_Caprine_H ----- [189]
#KY828430.1_Caprine_I ----- [189]
#KY828429.1_Caprine_J ----- [189]
#AY032689.1_Caprine_K SRC-PVSHPM KTL-LTLL-- ----- [189]
```

Figure 1: Multiple Sequence Alignment of the Reference Amino Acids of Cattle, Sheep, and Goats GDF8-gene Using ClustalW

Multiple sequence alignment (MSA) is a crucial computational process in molecular biology. It involves comparing sequences to conduct phylogenetic constructions, analyses of protein structures and predicts protein functions (Kemena and Notredame, 2009). Among the pioneering members of the Clustal family, ClustalW stands out as a widely favoured multiple sequence alignment algorithm (Jurate *et al.*, 2013). ClustalW, introduced by Thompson *et al.* (1994), quickly became the preferred method for generating multiple sequence alignments due to its significant improvements in alignment quality, sensitivity, and speed compared to other algorithms. (Jurate *et al.*, 2013). It employs a unique position-specific scoring scheme and a weighting scheme to address overrepresented sequence groups, with the "W" indicating "weights." Initially, the algorithm conducts a pairwise alignment of all sequences (nucleotide or amino acid) using the k-tuple method by Wilbur and Lipman (1983). The methods generate a similarity matrix for each pair of sequences, which is then converted to distance scores to construct a guide tree using the neighbour-joining (NJ) method (Saitou and Nei, 1987). The final step involves the

multiple sequence alignment (MSA) of all sequences by progressively aligning the most closely related sequences based on the guide tree generated by the NJ method. (Jurate *et al.*, 2013)

Functional Analysis (Missense Mutation) of the Cattle, Sheep, and Goat GDF8-Gene

The functional analysis results of coding non-synonymous single nucleotide polymorphisms (nsSNP) in the cattle and sheep GDF8-gene are presented in Tables 1 and 2. Table 1. showed that ten amino acid substitutions (S7Y, M14T, N26K, L43M, F139Y, F94L, R48G, A167S, S125T and P285E) in cattle, seven in sheep (F7Y, L16I, K36E, L43M, N48T, K49T and T120S) were tolerated indicating that the substitutions did not impair protein function (beneficial). however, three in cattle (Y9C, E46D and R161T), five in sheep (Y9C, E34G, R45G, T114A and I119V) were deleterious an indication that the substitutions were harmful.

Table 2: Functional Analysis of Coding nsSNP of the Cattle and Sheep GDF8-gene Using SIFT

Cattle			Sheep		
Variants	SIFT score	Prediction	Variants	SIFT score	Prediction
S7Y	1.00	Tolerated	F7Y	1.00	Tolerated
Y9C	0.01	Deleterious	Y9C	0.00	Deleterious
M14T	0.29	Tolerated	L16I	1.00	Tolerated
N26K	0.29	Tolerated	E34G	0.00	Deleterious
L43M	0.16	Tolerated	K36E	1.00	Tolerated
E46D	0.02	Deleterious	L43M	0.12	Tolerated
F94L	0.64	Tolerated	R45G	0.00	Deleterious
R98G	0.39	Tolerated	N48T	1.00	Tolerated
F139Y	0.06	Tolerated	K49T	0.17	Tolerated
R161T	0.02	Deleterious	T114A	0.04	Deleterious
A167S	0.43	Tolerated	I119V	0.04	Deleterious
S125T	1.00	Tolerated	T120S	0.27	Tolerated
P285E	0.07	Tolerated	-	-	-

The default threshold is <0.05, that is; Variants with a SIFT score < 0.05 are considered “deleterious” while Variants with a SIFT score above the threshold are considered “tolerated”. G = glycine, L = leucine, M = methionine, F = phenylalanine, K = lysine, E = glutamic acid, S = serine, P = proline, C = cysteine, Y = tyrosine, R = arginine, N = asparagine, D = aspartic acid, and T = threonine.

Table 3. Showed Sixteen amino acid substitutions of the wild-type were obtained. Out of these, ten amino acid substitutions (F7Y, F14I, L16I, H23Y, K36E, K39Q, T42A, L43M, N48T and L53I) were beneficial, the remaining six amino acid substitutions (V26L, W27R, G37R, T40I, R50G and G57R) were harmful.



Table 3: Functional Analysis of Coding nsSNP of the Goat GDF8-gene Using SIFT

Variants	SIFT score	Prediction
F7Y	1.0	Tolerated
F14I	0.12	Tolerated
L16I	1.00	Tolerated
H23Y	1.00	Tolerated
V26L	0.04	Deleterious
W27R	0.02	Deleterious
K36E	1.00	Tolerated
G37R	0.00	Deleterious
K39Q	0.57	Tolerated
T40I	0.02	Deleterious
T42A	0.45	Tolerated
L43T	1.00	Tolerated
N48T	1.00	Tolerated
R50G	0.00	Deleterious
L53I	1.00	Tolerated
G57R	0.01	Deleterious

The default threshold is <0.05 , that is; Variants with a SIFT score < 0.05 are considered “deleterious” while Variants with a SIFT score above the threshold are considered “tolerated.” G = glycine, A = alanine, L = leucine, F = phenylalanine, W = tryptophan, K = lysine, Q = glutamine, E = glutamic acid, V = valine, I = isoleucine, Y = tyrosine, H = histidine, R = arginine, N = asparagine, and T = threonine.

In molecular biology, the main focus on nsSNPs and population genetics is identifying and describing functionally related nsSNPs (Doss *et al.*, 2012). According to Sim *et al.* (2012), SIFT (sorting intolerant from tolerant) is a bioinformatics tool that predicts how amino acid substitutions affect protein function. It can assess naturally occurring polymorphisms and lab-induced mutations. Harmful substitutions can change protein function, potentially increasing disease susceptibility, and affecting enzyme activity, structure, interactions, and muscle mass regulation. Meanwhile, neutral amino acid substitutions improve protein structure and assist in positive genetic modification of the species. (Ewuola *et al.*, 2018; Bibinu *et al.*, 2016). Therefore, the beneficial nsSNPs obtained in the present study offer hope for future genetic improvement of cattle, sheep, and goats at the GDF8-gene locus. This is because nsSNPs have been reported to be linked to economically important traits, and disease development (Vincent *et al.*, 2014). Therefore, the GDF8-gene as a genetic marker closely related to the myogenesis trait can be used in marker-assisted selection (MAS) for high muscle production and productivity in ruminants (cattle, sheep, and goats).

Estimates of Codon-based Evolutionary Divergence (Substitutional Analysis) between GDF8-gene Sequences of cattle, sheep, and goat

Table 3. showed the means of dS, dN, and dN/dS with their corresponding Z-scores and P-values are shown. The result showed the highest dN/dS value of 2.71 in sheep, while cattle had the lowest value of 1.55. The dN/dS pattern revealed positive selection for all the studied species.

Table 3. Mean Number of Nucleotide Substitutions Per Synonymous Site (dS) and Nonsynonymous Site (dN) with their Ratio Among Cattle, Sheep, and Goats

Specie	No. of codons	dS	dN	dN/dS	Z-score	P-value
Cattle	173	1.00	1.55	1.55	4.41	0.00
Sheep	191	0.14	0.38	2.71	2.50	0.00
Goat	140	0.79	1.49	1.87	5.40	0.00

dN/dS = ratio of nonsynonymous to synonymous divergence.

According to Yakubu *et al.* (2013), one of the useful estimates of gene selective pressure was the ratio of the number of nonsynonymous substitutions per nonsynonymous site (dN; amino acid altering) to the number of synonymous mutations per synonymous site (dS; silent mutation), also known as omega ($\omega = dN/dS$). An omega (ω) value greater than 1 ($\omega > 1$) is an indication of positive selection (Zhang, *et al.*, 2005; Bibinus, *et al.*, 2018). The omega values obtained in the present study were greater than one (>1). A value of omega ($\omega < 1$) indicates that there was an excess of non-synonymous over synonymous which increased allelic polymorphism (Bibinus, *et al.*, 2018), which in turn might induce a change in the conformation of protein structure, thereby affecting the signalling pathways (Cui, *et al.*, 2009). The dN/dS pattern revealed positive selection for all the studied species. This indicates that the non-synonymous site evolved faster than the synonymous sites, and the positive selection effect overshadows purifying selection. It is an implication that balancing selection (or a positive Darwinian section) favoured new variants (Bibinus, *et al.*, 2018).

Tajima's neutrality test for the cattle, sheep, and goat GDF8-gene

The Tajima's neutrality test is presented in Table 9. The results depicted that, two of the species cattle and goats showed a positive value of D, while sheep had a negative value of D. Cattle had the highest value of 0.964501.

Table 9: Tajima's Neutrality Test of Cattle, Sheep, and Goat GDF8-gene

Specie	M	S	Ps	Θ	π	D
Cattle	4	485	0.778491	0.424632	0.463617	0.964501
Sheep	8	265	0.417323	0.160951	0.090776	-2.390114
Goat	11	461	0.874763	0.298659	0.331965	0.540738

m = number of sequences, S = number of segregating sites, $p_s = S/n$, $\Theta = p_s/a_1$, π = nucleotide diversity, and D is the Tajima test statistic.

A positive Tajima's D in the selected species indicates low levels of low and high-frequency polymorphisms signifying a balanced selection. This might aid in purifying selection against nonsynonymous mutations due to their deleterious impacts (Ewoula *et al.*, 2018). However, a negative D value (-2.39) was observed in sheep, indicating that the GDF8-gene in the



selected species is not neutrally evolving. It has an excess of rare variation consistent with population growth or positive selection. This is in line with the submissions of Eckshtain-Levi *et al.* (2018), who found that a Tajima's *D* value below -2 or above 2 is generally considered a strong indication that a gene is not evolving neutrally. Genes with a Tajima's *D* value below -2 have excess rare alleles, indicating positive selection or a selective sweep. However, genes with a Tajima's *D* value above 2 have an excess of common alleles suggestive of balancing selection. The distribution of Tajima's *D* values (with a 95% confidence interval between -2 and $+2$) is based on the assumption that the ascertainment of polymorphism is independent of allele frequency (Nassar *et al.*, 2023).

Phylogenetic Relationship of the GDF8-gene Between Cattle, Sheep, and Goats

The phylogenetic analysis is based on nucleotide sequences of the GDF8-gene, as shown in Fig. 2. which showed inter-cluster among sequences in all the species with a high-level intermingling between them. The evolutionary pathway and the genetic relationship of GDF-gene of the studied species using UPGMA (unweighted pair-group method using an arithmetic average) showed that members of the Bovidae family (sheep, goats, and cattle) were closely related except for caprine A and C which shows a distant similarity with all the studied sequences. According to the report of Ewoula *et al.* (2018), the most widely used method in constructing phylogenetic trees or dendrograms is the unweighted pair group method using an arithmetic average (UPGMA). The UPGMA consensus trees based on nucleotide and amino acid sequences revealed that cattle, sheep, and goats were closer. This is by classical classification, as goats, sheep, and cattle are members of the *bovidae* family. The findings of this study agree with the submissions of Ugbo *et al.* (2015), who observed similar clustering of members of the *Bovidae* family and order *Artiodactyla*. Furthermore, Akinyemi *et al.* (2017) observed a similar clustering of members of the *Bovidae* family (cattle, sheep, and goats).

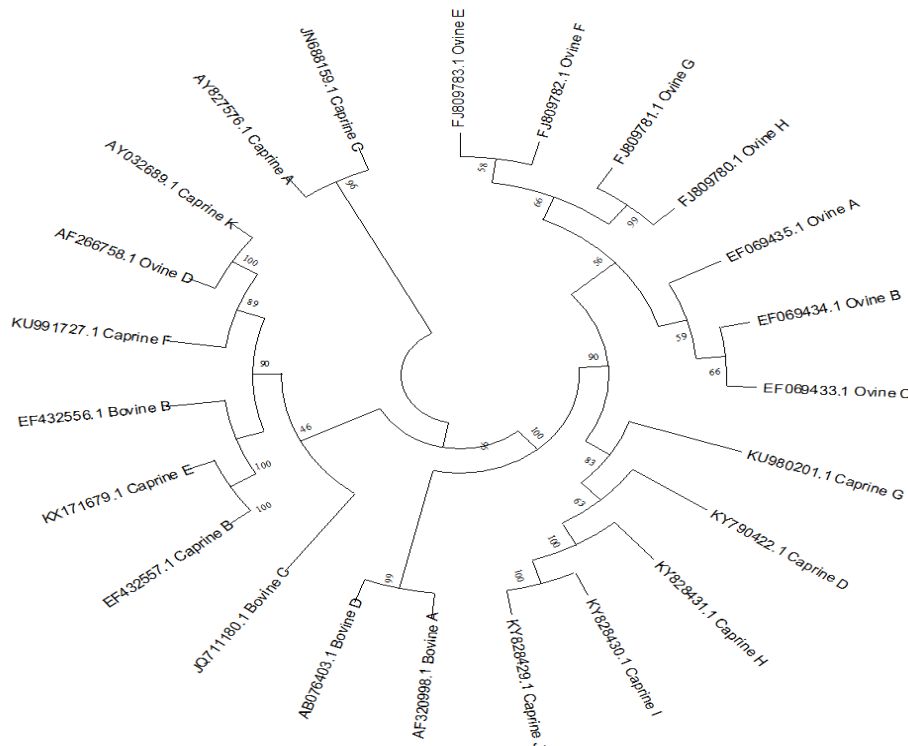


Figure. 2. Phylogenetic Relationship of Cattle, Sheep, and Goats GDF8-gene Nucleotide Sequences Using UPGMA

CONCLUSION AND RECOMMENDATIONS

It can be inferred from the study that there was variation in the sequence length of base pairs (bp) of the GDF8-genes (381 bp and 7831 bp) within and among species. The cattle, sheep, and goat GDF8-gene are polymorphic genes with numerous mutations that can be neutral (beneficial) or harmful (deleterious). The beneficial ones were more prevalent in the current study. The dN/dS pattern revealed positive selection for all the studied species. This indicates that the non-synonymous site evolved faster than the synonymous site, and the positive selection effect overshadows the purifying selection. The evolutionary pathway and genetic relationships of the GDF8-genes showed a close relationship among members of the Bovidae family (cattle, sheep, and goats). Based on this study, the following recommendations can serve as the groundwork for consideration; Utilizing the *in-silico* approach to predict mutations and their effects on biological mechanisms is crucial for gaining valuable insights into the role of mutations in biological processes. Further research should be conducted to explore the potential applications of the results in selecting desirable traits and the genetic improvement of cattle, sheep and goats.

REFERENCES

Aiello, D., Patel, K. and Lasagna, E. (2018). The myostatin gene: An overview of mechanisms of action and its relevance to livestock animals. *Animals Genetics*, **49**(6): 505–519.



- Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W. and Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic acids research*, **25**(17):3389–3402
- Akinyemi, M. O., Osaiyuwu, H. O. and Ismail, A. A. (2017). Computational molecular analysis of the sequences of the PAPP2 gene of selected ruminants and non-ruminants. *Biotechnology Journal International*, **19**(1):1-8
- Al-Khudhair, A., VanRaden, P. M., Null, D. J. and Li, B. (2021). Marker selection and genomic prediction of economically important traits using imputed high-density genotypes for 5 breeds of dairy cattle. *J. Dairy Sci*, **104**(4):4478–85.
- Bhattacharya, T. K., Rajith R., Bheemreddy, Chatterjee, R. and Ashwini, R. (2022). Myostatin (GDF8) gene and its intriguing role in regulating growth in poultry. *The Indian Journal of Animal Sciences*, **92**(10): 1141–1148.
- Bibinu, B. S., Yakubu, A., Ugbo, B. S. and Dim, I. N. (2016). Computational Molecular Analysis of the Sequences of BMP15 Gene of Ruminants and Non-Ruminants. *Open Journal of Genetics*, **6**(1):39-50.
- Casas, E. and Kehrl, M. E. Jr. (2016). A review of selected genes with known effects on the performance and health of cattle. *Front. Vet Sci.*, **3**(113):1-11.
- Cui, H.X., Zhao, S.M., Cheng, M.L., Guo, L., Ye, R.Q., Liu, W.Q. and Gao, S.Z. (2009). Cloning and Expression Levels of Genes Relating to the Ovulation Rate of Yunling Black Goat. *Biology of Reproduction*, **80**(2):219-226.
- Divya, D., Bhattacharya, T. K., Gnana, P. M., Chatterjee, R. N., Shukla, R., Vishnu, G. and Dushyanth, K. (2018). Molecular characterization and expression profiling of BMP 3 gene in broiler and layer chicken. *Molecular Biology Reports*, **45**(4):477–95.
- Doss, C.G. P. and Rajith, B. (2012). Computational refinement of functional single nucleotide polymorphisms associated with ATM gene. *PLoS ONE*. **7**(4): 1-12.
- Dushyanth, K, Bhattacharya, T K, Shukla R, Chatterjee, R N, Sitaramama, T, Paswan, C. and *et al.* (2016). Gene expression and polymorphism of Myostatin gene and its association with growth traits in chicken. *Animal Biotechnology*, **27**(4):269–77.
- Eckshtain-Levi, N., Weisberg, A. J. and Vinatzer, B. A. (2018). The population genetic test Tajima's D identifies genes encoding pathogen-associated molecular patterns and other virulence-related genes in *Ralstonia solanacearum*. *Molecular plant pathology*, **19**(9):2187–2192
- Ewuola, M. K., Omolara A. M. and H. Osamede, O. (2018). Insilico Analysis of Myostatin Gene in Selected Poultry Species. *Journal of Advances in Biology & Biotechnology*, **17**(2):1-10.
- Felsenstein J. (1985). Confidence limits on phylogenies: An approach using the bootstrap. *Evolution*, **39**:783-791.
- Guo, R., Wan, Y., Xu, D., Cui, L., Deng, M., Zhang, G. and *et al.* (2016) Generation and evaluation of Myostatin knock-out rabbits and goats using CRISPR/Cas9 system. *Sci. Rep.* **6**(29855):1-10.
- Guru, V. P B., Bhattacharya, T. K., Kumar, P., Chatterjee, R. N., Ravi-Kumar G. V. P. P. S., Paswan, C. and Rajendra, P. A. (2017). Expression profiling of activin type IIB receptor during ontogeny in broiler and indigenous chicken. *Animal Biotechnology* **28**(1):26–36.
- Haskell, M. J., Simm, G. and Turner, S. P. (2014). Genetic selection for temperament traits in dairy and beef cattle. *Frontiers in genetics*, **5**(368):1-14.



- Jurate D., Aisling, O'. and Roy, S. (2013) An Overview of Multiple Sequence Alignments and Cloud Computing in Bioinformatics. *ISRN Biomathematics*, **6**(1):1-14.
- Kemena, C. & Notredame, C. (2009) Upcoming challenges for multiple sequence alignment methods in the high-throughput era. *Bioinformatics*, **25**(19):2455–2465
- Larkin, M.A., Blackshields, G., Brown, N.P., Chenna, R., McGettigan, P.A., McWilliam, H. and *et al.* (2007) Clustal W and Clustal X Version 2.0. *Bioinformatics*, **23**(1):2947-2948.
- Liu, L. and Kumar, S. (2013). Evolutionary balancing is critical for correctly forecasting disease-associated amino acid variants. *Molecular Biology*, **30**(1)1252-1257.
- Nassar, L. R., Barber, G. P., Benet-Pagès, A., Casper, J., Clawson, H., Diekhans, M. and *et al.* (2023). The UCSC Genome Browser database: 2023 update. *Nucleic acids research*, **51**(D1): D1188–D1195.
- Osman, N. M., Shafey, H. I., Abdelhafez, M. A., Sallam, A. M. and Mahrous, K. F. (2021) Genetic variations in the *Myostatin* gene affecting growth traits in sheep. *Veterinary World*, **14**(2):475-482.
- Rasal, K. D., Chakrapani, V. and Patra, S. (2016). Identification and prediction of the consequences of nonsynonymous SNPs in glyceraldehyde-3- phosphate dehydrogenase (GAPDH) gene of zebrafish *Danio rerio*. *Turkish Journal of Biology*, **40**(1):43–54.
- Saitou, N. and Nei, M. (1987). The Neighbor-Joining Method: A New Method for Reconstructing Phylogenetic Trees. *Molecular Biology and Evolution*, **4**(1)406-425.
- 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 Research*, **40**(W1): W452–W457.
- Satheesh, P., Bhattacharya, T. K., Kumar, P., Chatterjee, R. N., Dhara, S. K., Paswan, C. and Dushyanth, K. (2016). Gene expression and silencing of activin receptor type 2A (ACVR2A) in myoblast cells of chicken. *British Poultry Science*, **57**(6):763–70.
- Tamura K., Stecher G. and Kumar S. (2021). MEGA 11: Molecular Evolutionary Genetics Analysis Version 11. *Molecular Biology and Evolution*, **38**(7):3022–3027.
- Thompson, J. D., Higgins, D. G. and Gibson T. J., (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignments through sequence weighting, position-specific gap penalties, and weight matrix choice, *Nucleic Acids Research*, **22**(1)4673–4680
- 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**(1):145-158.
- Vaser, R., Adusumalli, S., Leng, S. N., Sikic, M. and Ng, P. C. (2016). SIFT missense predictions for genomes. *Nat Protocols* **11**: 1-9.
- 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*, **4**(1)564-570.
- Vishnu, P. G., Bhattacharya, T. K., Kumar, P., Chatterjee, R. N., Paswan, C. and Dushyanth, K. (2016). Single nucleotide polymorphism in the 5'-upstream regions of activin A receptor, type IIB in Aseel Chicken. *Journal of Animal Research*, **6**(2):221–25.



- Wilbur, W. J. and Lipman, D. J. (1983) Rapid similarity searches of nucleic acid and protein data banks, *Proceedings of the National Academy of Sciences of the United States of America*, **80**(3):726–730
- Wu, D., Gu, M., Wei, Z., Bai, C., Su, G., Liu, X. and *et al.* (2022). Myostatin knockout regulates bile acid metabolism by promoting bile acid synthesis in cattle. *Animals*, **12**(2):1–13.
- Yakubu, A., Salako, A.E., De Donato, M., Takeet, M.I., Peters, S.O., Adefenwa, M.A. and *et al.* (2013). Genetic Diversity in Exon 2 at the Major Histocompatibility Complex *DQB1* Locus in Nigerian Indigenous Goats. *Biochemical Genetics*, **51**, 954-966.
- Yakubu, A., Alade, E.D. and Dim, N.I. (2014). Molecular Analysis of Solute Carrier Family 11 a1 (Slc11a1) Gene in Ruminants and Non-Ruminants Using Computational Method. *Genetika*, **46**(1)925-934.
- Yang Y. L., Zhou, R. and Li, K. (2017). Future livestock breeding: precision breeding based on multi-omics information and population personalization. *J. Integr. Agric.*, **16**(12):2784–91.
- Zhang, X. X., Ran, J. S., Lian, T., Li, Z.Q., Yang, C.W., Jiang, X.S. and *et al.* (2019). The single nucleotide polymorphisms of myostatin gene and their associations with growth and carcass traits in Daheng broiler. *Braz. J. Poult. Sci.*, **21**(3):1-8.