



EFFECTS OF MUTATIONS OF MYOSTATIN (MSTN) GENE IN BREEDS OF SHEEP IN NIGERIA

Zakka, T. E., Mancha, Y. P., Bello, K. M., Garba, A. M. and Luka, P. N.

Department of Animal Production, Faculty of Agriculture and Agricultural Technology
Abubakar Tafawa Balewa University, Bauchi, Nigeria.

Corresponding Author's Email: znasamu@gmail.com, **GSM:** 07040507547

ABSTRACT

Growth Differentiation Factor 8 (GDF8), also known as myostatin, is a member of the Transforming Growth Factor β (TGF β) superfamily. It has been reported to act as a negative extracellular regulator and a significant negative regulator of skeletal muscle mass, muscle growth, and development. This research was carried out to identify the effects of molecular genetic variation of the myostatin (MSTN) gene in breeds of sheep in Nigeria, using, in silico tools. A total of eight myostatin (MSTN) nucleotide sequences were retrieved from NCBI (National Centre for Biotechnology Information) website (www.ncbi.nlm.nih.gov) it was aligned, translated, and compared using Clustal W from MEGA to find genetic distances and evolutionary relationships. BLAST tool was used to align and identify the mutations of the sequences considered. The variation in sequence length base pairs (bp) were 511 and 580 bp. Functional analysis of missense mutations using SIFT identified twenty-one (21) amino acids substitutions in the myostatin (MSTN) gene sequences. (18 were tolerant or beneficial while 3 were deleterious or harmful). There were 7 Essential to Essential substitutions, 5 Essential to Non-Essential substitutions, 4 Non-Essential to Non-Essential substitutions, 5 Non-Essential to Essential substitutions. The beneficial mutations were more prevalent than the harmful mutations. There should be Implementation of selective breeding strategies to propagate the beneficial mutations identified, enhancing muscle growth and meat yield in breeds of sheep in Nigeria.

Keywords: Myostatin, In-silico, Growth, Factor, Mutations.

INTRODUCTION

Meat products from ruminants like sheep play vital role in the agricultural sector of most countries (Selvaggi *et al.*, 2015). The Food and Agriculture Organization (FAO, 2023) reported that the economic importance of sheep in developing nations cannot be overemphasized. Sheep, with their small body size, high productive capacity, and rapid growth rates, are ideally suited to production by resource-poor smallholders. FAO (2023) further highlighted that in sub-Saharan Africa, sheep provide almost 30 % of the meat consumed and around 16 % of the milk produced. Sheep contribute about 50 % of the total domestically produced meat in Nigeria, which possesses about 22.1 million sheep (FAO, 2023). However, the lack of genetic information has prevented farmers from optimizing the full potential of sheep breeds in Nigeria.

Decisions on applying marker-assisted selection (MAS) for genetic improvement of any breed must consider traits of economic value and the specific identification of genes associated with those traits (Li *et al.*, 2006). The myostatin (MSTN) gene, also known as the growth differentiation factor-8 gene, is expressed in skeletal muscle and inhibits muscle growth (Bellinge *et al.*, 2005). Located on chromosome 2 in sheep (Archibald *et al.*, 2010), loss of functional myostatin causes the "double muscled" phenotype in various species, including sheep (McPherron and Lee, 1997; Grobet *et al.*, 1998; Broad *et al.*, 2000). Mutations in the



myostatin gene are linked to the muscular hypertrophy allele in double-muscled breeds (Kambadur *et al.*, 1997), presenting an opportunity to enhance processing yields using gene knockout technology (Kocabas *et al.*, 2002). Mutations in the myostatin gene have been reported in sheep breeds such as Kemieniec, Pomeranian, Medras, and Mecheri, with heterozygotes showing higher body weight (Amiya *et al.*, 2017). Most studies on Nigerian sheep breeds are morphological (Agaviezor *et al.*, 2012; Asamoah and Sam, 2016), resulting in a scarcity of molecular markers and making genetic improvement programs challenging. Thus, examining the myostatin (MSTN) gene and its variants in breeds of sheep in Nigeria is crucial for identifying animals with desirable genotypes for future selection programs, particularly in MAS for economic traits. These genetic markers will be valuable for selection and breeding programs, as variations in the non-coding regions of the myostatin (MSTN) gene relate to muscle growth and meat quality traits in sheep, impacting the gene.

MATERIALS AND METHODS

The Study Area

This study focuses on, in-silico analysis of the myostatin (MSTN) gene in the breeds of sheep in Nigeria. The research utilizes several bioinformatics tools to understand the evolutionary relationships, genetic distances, sequence similarities, and potential effects of mutations in the myostatin (MSTN) gene sequences. All analysis were done with bioinformatics software programs namely; MEGA (molecular evolutionary genetics analysis tool); BLAST (Basic local alignment search tool); and SIFT (sorting intolerant from tolerant) which were installed in the computer and optimized.

Myostatin (MSTN) Sequences of Breeds of Sheep in Nigeria

Eight myostatin (MSTN) gene sequences were analyzed, with two sequences each from West African Dwarf (W.A.D), Uda, Balami, and Yankasa breeds. These sequences were retrieved from the NCBI GenBank (www.ncbi.nlm.nih.gov) with the following accession numbers; MZ065344.1, MZ065345.1, for WAD; MZ065339.1, MZ065340.1, for Uda; MZ065334.1, MZ065335.1, for Balami; and MZ712587.1, MZ712588.1, for Yankasa.

Sequence Alignment and Translation

Sequence alignment, translation, and comparison were performed using ClustalW (Larkin *et al.*, 2007) from MEGA with the IUB substitution matrix, a gap open penalty of 15, and a gap extension penalty of 6.66.

Identification of Mutations of Myostatin (MSTN) Gene Sequences Selected

To identify mutations of myostatin (MSTN) gene sequences, the Basic local alignment search tool (BLAST)x algorithm on the NCBI BLAST website (www.ncbi.nlm.nih.gov) was accessed. Myostatin gene sequence was pasted into the "Enter query sequence" field maintaining default parameters, then standard non-redundant protein sequence databases were selected, and the BLASTx algorithm was ran to compare the query sequence to the selected database. The pairwise with dots alignments was selected to identify mutations and similarities between the query sequence and matched sequences. One of the most common means to search sequence databases is through sequence similarity with known sequences. The mechanisms for the sequence similarity search tools available include Blast (Altschul *et al.*, 1990).

Characterization of Variations of Myostatin (MSTN) Gene Sequences Selected

To characterize gene variants, we used the SIFT (Sorting Intolerant from Tolerant) tool (<https://sift.bii.a-star.edu.sg>). The protein sequence was entered into the SIFT input field, and the appropriate database and parameters were selected. SIFT predicted the effects of amino acid substitutions on protein function, using a threshold value of ≤ 0.05 . Substitutions scoring less than 0.05 were deemed deleterious, while those scoring higher were considered tolerant.



SIFT, which uses UniProt as protein reference database, is a widely used algorithm for predicting the functional effects of amino acid substitutions (Ng and Henikoff, 2001; Thomas *et al.*, 2003).

RESULTS AND DISCUSSION

Results Retrieved Sequences and Base pairs

The variation in sequence length of was not significant between the breeds of sheep, with all the breeds having the same base pair (511), except for Yankasa, which had 580bp per sequence (Table 1). All these sequence lengths partial coding sequences (partial CDS) because they had less than six thousand base pairs (<6000 bp) each. Variability might arise from DNA duplication, rearrangement, short tandem repeats (STR), insertions, or deletions (indels). Sequences MZ065344.1, MZ065345.1, MZ065339.1, MZ065340.1, MZ065334.1, and MZ065335.1 have 511 bp and come from exon 1, while MZ712587.1 and MZ712588.1 have 580 bp and are from the promoter region. Such variability can lead to unique structures and biological activities among individuals (Bhattacharya *et al.*, 2008).

Table 1: Breeds, Accession Number, and Base Pairs of Myostatin (MSTN) Gene

Breeds	Ascension number	Base Pairs
WAD	MZ065344.1	511
	MZ065345.1	511
UDA	MZ065339.1	511
	MZ065340.1	511
BALAMI	MZ065334.1	511
	MZ065335.1	511
YANKASSA	MZ712587.1	580
	MZ712588.1	580

Sequence Alignment

The myostatin gene sequences were aligned without gaps, indicating high similarity and conservation among the sequences. This lack of gaps suggests minimal evolutionary insertions or deletions, allowing for seamless alignment (Durbin *et al.*, 1998; Mount, 2004). Proper alignment parameters, such as gap penalties, also contribute to minimizing gaps (Thompson *et al.*, 1994). For short sequences, fewer variations necessitate gaps, further leading to smooth alignments (Higgins and Sharp, 1988).

Genetic Distances

The genetic distances revealed the relationship between the four breeds of sheep in Nigeria. It showed that Balami and Uda were closer with a distance of 0, which was attributed to the fact that Balami and Uda had a common ancestor. This result was similar to Yunusa on morphology of breeds of sheep in Nigeria who reported that Balami and Uda were the closest (Yunusa *et al.*, 2013). Yankasa and Balami had a maximum distance of 0.568693 even though they had a common ancestor. This result is similar to Abubakar on genetic diversity of breeds of sheep in Nigeria who reported that Yankassa and Balami had the farthest distance (Abubakar *et al.*, 2020) (Table 2). Nei (1987) expounded the concept by stating that genetic distance is “the extent of gene differences between populations or species that is measured by some numerical quantity.” A second definition was given by Beaumont *et al.*, (1998). They defined

genetic distance as a “quantitative measure of genetic difference, be it at the sequence level or the allele frequency level that is calculated between individuals, populations or species.”

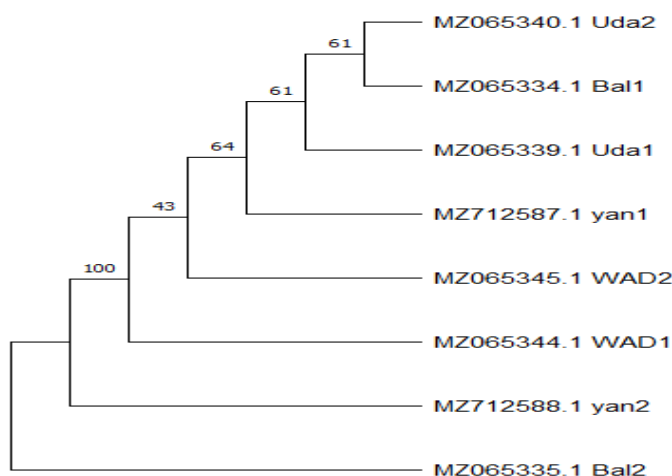


Table 2: Genetic Distances in the Sequences Considered

	WAD1	WAD2	UDA1	UDA2	BAL1	BAL2	YAN1	YAN2
WAD1								
WAD2	0.006							
UDA1	0.012	0.006						
UDA2	0.012	0.006	0.000					
BAL1	0.012	0.006	0.000	0.000				
BAL2	0.376	0.367	0.376	0.376	0.376			
YAN1	0.018	0.012	0.006	0.006	0.006	0.385		
YAN2	0.222	0.214	0.208	0.208	0.208	0.569	0.214	

Figure 1: Phylogenetic Relationships of Sequences Considered using UPGMA Bootstrap Concession Tree.

Nei Genetic Distance: West African Dwarf (Wad), Balami (Bal), Yankasa (Yan) Evolutionary Relationships

The UPGMA consensus tree in Figures 2 and 3 showed that the first node, which carried Balami and Uda, were closer to a common ancestor than the others. The last node, which carried Yankasa and Balami, also had a common ancestor. The evolutionary relationships in Figures 2 and 3 indicated that breeds coming from the same nodes had a common ancestor. The sequences considered for the myostatin (MSTN) gene showed that seven nodes stemmed from the root. Balami 1 and Uda 2, West African Dwarf 1 and West African Dwarf 1, Yankasa 1 and Uda 1, and Balami 2 and Yankasa 2 all had a common ancestor. However, the node of interest was the first node, which carried Balami and Uda, indicating they were more homologous to the ancestors of breeds of sheep in Nigeria, similar to the report by Yunusa *et al.*, (2013).

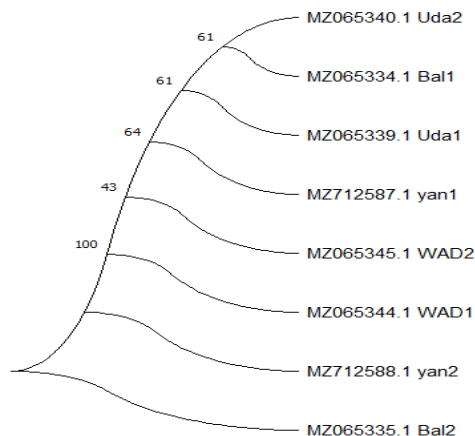


Figure 2: Phylogenic Relationships of the Sequences Considered using UPGMA Bootstrap Concession Tree

Characterization of Mutations in the Sequences Considered

There were twenty-one (21) amino acids substitutions of myostatin (MSTN) gene sequence considered. Eighteen (18) were tolerant, namely: Lysine 49 Threonine, Phenylalanine 13 Isoleucine, Phenylalanine 7 Tyrosine, Leucine 16 Isoleucine, Lysine 36 Glutamic, Leucine 43 Methionine, Asparagine 48 Threonine, Asparagine 49 Lysine, Leucine 51 Serine, Arginine 31 Glutamic, Proline 55 Alanine, Glycine 56 Isoleucine, Threonine 58 Isoleucine, Proline 59 Glutamic, Serine 60 Isoleucine, Lysine 36 Glutamic, Leucine 43 Methionine, and Arginine 57 Lysine were tolerant. Three (3) substitutions, namely; Glutamic 34 Glycine, Arginine 45 Glycine, Tyrosine 9 Cysteine, were deleterious (Table 3).

Table 2: Characterization of Myostatin (MSTN) Sequences Considered

Variants	SIFT Score	Effects
Lysine 49 Threonine	0.17	Tolerated
Glutamic 34 Glycine	0.00	Deleterious
Phenylalanine 13 Isoleucine	0.06	Tolerated
Arginine 45 Glycine	0.00	Deleterious
Phenylalanine 7 Tyrosine	1.00	Tolerated
Tyrosine 9 Cysteine	0.00	Deleterious
Leucine 16 Isoleucine	1.00	Tolerated
Lysine 36 Glutamic	1.00	Tolerated
Leucine 43 Methionine	0.12	Tolerated
Asparagine 48 Threonine	1.00	Tolerated
Asparagine 49 Lysine	1.00	Tolerated
Leucine 51 Serine	0.17	Tolerated
Arginine 31 Glutamic	1.00	Tolerated
Proline 55 Alanine	1.00	Tolerated
Glycine 56 Isoleucine	1.00	Tolerated
Threonine 58 Isoleucine	1.00	Tolerated
Proline 59 Glutamic	1.00	Tolerated
Serine 60 Isoleucine	1.00	Tolerated
Lysine 36 Glutamic	1.00	Tolerated
Leucine 43 Methionine	1.00	Tolerated
Arginine 57 Lysine	1.00	Tolerated

Sift score of < 0.05 is deleterious, while score of > 0.05 is tolerated.



The SIFT algorithm predicts the impact of amino acid substitutions on protein function based on the conservation of amino acids within a protein family. It assumes that well-conserved positions are crucial for protein function, thus substitutions at these positions are likely to be deleterious. For instance, substitutions from hydrophobic to hydrophilic amino acids or vice versa may be tolerated or deleterious depending on the protein's requirements. SIFT assigns a score to each substitution, with scores below 0.05 indicating a prediction of deleterious impact, though some suggest using a threshold of 0.1 for better sensitivity. This computational approach aids in identifying functional mutations in genetic studies, distinguishing beneficial mutations from those that reduce fitness, and is valuable in both human and animal genetics research (Tariq *et al.*, 2013).

CONCLUSION AND RECOMMENDATIONS

It was concluded from the study that; the sequence length of the myostatin (MSTN) gene in breeds of sheep in Nigeria varies between 511 and 580bp and within these breeds of sheep in Nigeria, there were both beneficial and harmful mutations in the myostatin (MSTN) gene. The beneficial mutations were more prevalent than the harmful ones. The study recommended selective breeding strategies to propagate the beneficial mutations identified, enhancing muscle growth and meat yield in breeds of sheep in Nigeria.

REFERENCES

- Abubakar, G. R., Ezewudu, E. A., Egena, S. S. A., and Usman, A. (2020). Genetic diversity of Nigeria indigenous sheep breed from β -Lactoglobulin gene locus. *Journal Biodiversity Genetics*, **4**(2): 40-47.
- Agaviezor, B. O., Peters, S. O., Adefenwa, M. A., Yakubu, M., Adebambo, and O. A., Ekundayo. (2022). Genetic diversity of Nigeria indigenous sheep breed from DRB1 gene locus. *Journal Biodiversity*, **5**(3): 50-58.
- Altschul, S. F, Gish, W., Miller, W., Myers, E. W, Lipman, D. J. (1990). Basic local alignment search tool. *Journal Molecular Biology*, **215**(7): 403-410.
- Amiya, R. S., Jeichitra, V., Rajendran, R., and Raja, A. (2017). Polymorphism in exon 3 of myostatin (MSTN) and its association with growth rate in Indian sheep breeds. *Small Ruminants Research*, **149**(3): 81-84.
- Archibald, A. L., Cocket, N. E., Dalrymple, B. P., and Faranaut. (2010). *Journal Animal Science Veterinary Medicine*, **4**(5): 449-453
- Asamoah, M., and Sam, E. K. (2016). Morphological characterization of breeds of sheep: a discriminant analysis approach. *Springer Plus*, **5**(69): 60-67.
- Beaumont, M. A., Ibrahim, K. M., Boursot, P., and Bruford, M. W. (1998). Measuring genetic distance. *Molecular tools for screening biodiversity*, **6**(4): 315-325.
- Bellinge, R. H., Liberles, D. A., Iaschi, S. P., O'Brien P, A., and Tay, G. K (2005). Myostatin and its implications on animal breeding: a review. *Animal Genetics*, **36**(4): 1-6.
- Bhattacharya, T., Kumar, P. and Sharma, A. (2008). Molecular cloning and characterization of CSN2 in Indian riverian buffalo (*Bubalis bubalis*). *Buffalo Bulletin*, **27**(5): 222-230.
- Broad, T. E., Glass, B. C., Greer, G. J., Robertson, T. M., Bain, W. E., Lord, E. A., McEwan, J. C., and Peterson, S. W. (2000). Search for a locus near to myostatin that increases muscling in Texel sheep in New Zealand. *In Proceedings-New Zealand Society of Animal Production*, **60**(5): 110-112.
- Durbin, R., Eddy, S. R., Krogh, A., and Mitchison, G. (1998). *Biological Sequence Analysis: Probabilistic Models of Proteins and Nucleic Acids Research*, **8**(4): 33-39.



- Grobet, L., Martin, L. J. R., Poncelet, D., Pirottin, D., Brouwers, B., Riquet, J., Schoeberlein, A., Dunner, S., Menissier, F., Massabanda, J., Fries, R., Hanset, R., and Georges, M. (1997). A deletion in the bovine myostatin gene causes the double muscled phenotype in cattle. *Nature Genetics*, **17**(7): 71-74.
- Higgins, D. G., and Sharp, P. M. (1988). CLUSTAL: a package for performing multiple sequence alignment on a microcomputer. *Gene*, **73**(1): 237-244.
- Kambadur, R., Sharma, M., Smith, T. P., and Bass, J. J. (1997). Mutations in myostatin (GDF8) in double-muscled Belgian Blue and Piedmontese cattle. *Genome Research*, **7**(5): 910-916.
- Kocabas, A. M., Kucuktas, H., Dunham, R. A., and Liu, Z. (2002). Molecular characterization and differential expression of the myostatin gene in channel catfish (*Ictalurus punctatus*). *Biochimical Biophysical Acta (BBA)-Gene Structure and Expression*, **15**(6): 99-107.
- 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**(8): 2947-8.
- Li, X. L., Wu, Z. L., Gong, Y. F., Liu, Y. Q., Liu, Z. Z., Wang, X. J., Xin, T. R., and Ji, Q. (2006). Single-nucleotide polymorphism identification in the caprine myostatin gene. *Journal of Animal Breeding and Genetics*, **123**(2): 141-144.
- McPherron, A. C., and Lee, S. J. (1997). Double muscling in cattle due to mutations in the myostatin gene. *Proceedings of the National Academy of Science of the United States of America*, **94**(33): 12457-12461.
- Mount, D. W. (2004). *Bioinformatics: Sequence and Genome Analysis*. Cod Spring Harbor Laboratory Press, **9**(7): 73-78.
- Nei, M. (1987). *Molecular evolutionary genetics*. New York: Columbia University Press, **7**(4): 33-45.
- Ng, P. C., and Henikoff, S. (2003). SIFT: Predicting Amino Acid Changes that Affect Protein Function. *Nucleic Acids Research*, **31**(13): 3812-3814.
- OECD-FAO Agricultural Outlook 2023-2032. OECD I Library, 2023. Accessed July 7, 2024. https://www.oecd-ilibrary.org/agriculture-and-food/oecd-fao-agricultural-outlook-2023-2032_agr_outl-2023-en.
- Tariq, A. M., Al-Shammari, A. S., Al-Muammar, M. N. and Alhamdan, A. A. (2013). Evaluation and identification of damaged SNPs in COL1A1 gene involved in osteoporosis. *Archive Medical Science*, **9**(6): 899-905.
- Thomas, M., Langley, B., Berry, C., Sharma, M., Kirk, S., Bass, J., and Kambadur, R. (2000). Myostatin, a negative regulator of muscle growth, functions by inhibiting myoblast proliferation. *The Journal of Biological Chemistry*, **275**(45): 40235-40243.
- Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic acids research*, **22**(22): 4673-4680.
- Yunusa, A. J., Salako, A. E., and Oladejo, O. A. (2013). Morphometric Characterization of the Nigerian Indigenous Sheep using multi factorial discriminant analysis. *International Journal of Biodiversity and Conservation*, **5**(10): 661-665.
- Selvaggi, M., Laudadio, V., and Tufarelli, V. (2014). Meat quality and safety: A review of the current situation in the EU. *Journal of Food Science and Technology*, **51**(5): 833-843.