Characterization and Prediction of Protein Structures of Interferon Regulatory Factor 3 (Irf 3) Gene in three Nigerian Cattle Breeds

Abubakar A. Mohammed[1]; Michael O. Ozoje[2]; Alaka F. A.[1]; Salami, S. O.[1]; Christian O. Ikeobi[2]; Adewale O. Talabi[3]; Samuel O. Durosaro[2]; John S. De Campos[4]

[1]*Department of Agricultural Education, Federal College of Education, (Special), Oyo, Oyo State, Nigeria.*

[2]*Department of Animal Breeding and Genetics, Federal University of Agriculture, Abeokuta, Ogun State, Nigeria.*

[3]*Department of Veterinary Medicine and Surgery, Federal University of Agriculture, Abeokuta, Ogun State, Nigeria.*

[4]*Department of Animal Breeding and Genetics, Federal University of Agriculture, Abeokuta, Ogun State, Nigeria.*

Corresponding author: mohammed.abubakar1796@fcesoyo.edu.ng**:**08066062852

Abstract

The study was conducted to characterize and predict the secondary and tertiary protein structures of exons 1to 2 and exons 5 to 6 of the IRF 3 gene in three Nigerian cattle breeds, namely; the White Fulani, Muturu and N'Dama breeds. Blood samples for deoxyribonucleic acid (DNA) and studies were collected from 90 animals (30 White Fulani, 30 Muturu and 30 N'Dama). DNA was extracted from the blood samples using the Zymo-spin extraction kit. Primers were designed and DNA amplification was carried out using the specified PCR protocols. The purified PCR products were subjected to sequencing using BigDye® terminator cycle sequencing kit on ABI 3730xl (Applied Biosystems) DNA analyzer. The nucleotide sequences of the IRF3 gene of each breed were translated to amino acid sequence using NCBI ORF finder, Primary and secondary protein structures were predicted using GOR IV and Phyre 2 software, respectively, while the tertiary structures were viewed with Rasmol. Findings of the study revealed that the predicted secondary and tertiary protein structures and physico-chemical properties of exons 1 to 2 of the candidate gene in the three cattle breeds were identical, while for exons 5 to 6, the structures and physico-chemical properties were different. It was concluded that the prediction of secondary protein structures of the loci under consideration showed mixed folding proteins, which contained alpha helices, extended strand and random coils. It was recommended that; the roles of IRF3 gene in animal reproduction and development should be further explored and Polymorphism in the IRF3 gene and their association with diseases of economic importance in Nigerian cattle should also be studied.

Key words: Cattle, IRF 3, Proteins, Nigeria, Exons 1 and 2, Exons 5 and 6

Caractérisation et Prédiction des Structures Protéiques du Gène du Facteur Régulateur de l'Interféron 3 (IRF 3) de trois Races de Bovins Nigérianes

Résumé

L'étude a été menée pour caractériser et prédire les structures protéiques secondaires et tertiaires des exons 1 à 2 et des exons 5 à 6 du gène IRF 3 chez trois races de bovins nigérianes, à savoir; les races White Fulani, Muturu et N'Dama. Des échantillons de sang pour l'acide ésoxyribonucléique (ADN) et les études ont été collectés auprès de 90 animaux (30 White Fulani, 30 Muturu et 30 N'Dama). L'ADN a été extrait des échantillons de sang en utilisant le kit d'extraction Zymo-spin. Des amorces ont été conçues et l'amplification de l'ADN a été réalisée en utilisant les protocoles de PCR spécifiés. Les produits de PCR purifiés ont été soumis à un séquençage utilisant le kit de séquençage par cycle BigDye® terminator sur l'analyseur d'ADN ABI 3730xl (Applied Biosystems). Les séquences nucléotidiques du gène IRF3 de chaque race ont été traduites en séquence d'acides aminés à l'aide de

l'outil NCBI ORF finder. Les structures protéiques primaires et secondaires ont été prédites en utilisant respectivement les logiciels GOR IV et Phyre 2, tandis que les structures tertiaires ont été visualisées avec Rasmol. Les résultats de l'étude ont révélé que les structures protéiques secondaires et tertiaires prédites et les propriétés physico-chimiques des exons 1 à 2 du gène candidat chez les trois races de bovins étaient identiques, tandis que pour les exons 5 à 6, les structures et les propriétés physico-chimiques étaient différentes. Il a été conclu que la prédiction des structures protéiques secondaires des loci étudiés montrait des protéines à pliage mixte, contenant des hélices alpha, des brins étendus et des bobines aléatoires. Il a été recommandé que les rôles du gène IRF3 dans la reproduction et le développement des animaux soient davantage explorés et que le polymorphisme du gène IRF3 et son association avec les maladies d'importance économique chez les bovins nigérians soient également étudiés.

Mots-clés: Bovins, IRF 3, Protéines, Nigeria, Exons 1 et 2, Exons 5 et 6

Introduction

Cattle are generally acceptable in various parts of Nigeria. They contribute significantly to the available meat supplies in the country. They are widely distributed across the sub-humid and humid zones of Nigeria, where they make significant contributions to the economic development of the areas. Genetic variation has also been reported to be an indicator of functional adaptations, promoting species diversification, especially, during evolutionary processes (Vamathevan *et al*., 2008). The White Fulani cattle are members of the Zebu (*Bos indicus*), which constitute majority of cattle types in Africa. It is the most numerous and widespread of all the Nigerian cattle breeds, representing about 37.2% of the national cattle population (DAGRIS, 2005). Phenotypically, their coat colour is reported to be usually white with black dots, but some animals have a black suit with blue, red, or white spots (Norezzine *et al.,* 2019). The skin is loose with pigmented soft hair. The ears are erected. The horns are medium or long and lyre-shaped. White Fulani is commonly used for meat, milk, and draught (Norezzine *et al*., 2019). Various authors have reported the Zebu as being susceptible to trypanosomosis (Murray *et al.,* 1982, Njogu *et al.,* 1985 and Mattioli *et al*., 2000). They are adapted to dry environmental conditions and high temperatures and are known to be more tolerant to tick infestation compared to *Bos taurus* cattle (Mattioli *et al*., 2000). The Muturu (humpless shorthorn) and N'Dama (humpless longhorns) cattle are both members of the African *Bos taurus*. They mostly inhabit West and Central Africa. They are both characterised by small size and lower productivity compared to most of the zebu cattle populations in tropical areas (Rege and Tawah, 1999). Various authors have reported their unique evolutionary adaptation to harsh climate and various endemic diseases (Murray *et al*., 1984 and Mattioli *et al*., 2000).

The Interferon regulatory factor 3 (IRF3) gene is a member of the interferon regulatory transcription factor (IRF) family, which encodes interferon regulatory factor 3. It is widely expressed and resides predominantly in the cytoplasm of uninfected cells (Jann *et al*., 2009). The IRF3 gene has been identified as one of the various genes of the immune system which plays an active part in the establishment of the innate immune response against viral infection in animals and it has also been implicated in cells reaction to stress situations (Tamura *et al*., 2008). The cattle IRF3 gene consists of 8 exons and 7 introns and encodes a 417-amino acid protein. It is a phosphor-protein and consists of an N-terminal DNA binding domain (DBD domain), a C-terminal IRF-associated domain (IAD) and a transactivation domain, coupled with its critical roles in host defense and cell survival, the activity of IRF3 is strictly controlled. This gene is one of the strongest positional candidate genes implicated in a host of health-related phenotypes such as general disease resistance not only in cattle but in humans and mice as well, likewise this gene has also been reported to play an important role

in tolerance to protozoan infections in cattle and mice (Jann *et al*., 2009).

Cattle populations have several polymorphisms at the IRF3 locus that change single amino acids. Some of the most polymorphic regions of this gene in cattle are exons 2, 5 and 6 (Ensembl cow release 92). The complete exon 1 of this gene in cattle is un-translated, while exons 2 to 8 are completely translated to proteins (Ensembl cow release 92), but there is no report on the characterization and predicted protein structures of the IRF3 gene in Muturu, N'Dama and White Fulani breeds of cattle, which would be of great benefit to animal breeders and farmers by giving a better insight into the structures and functions of the candidate gene for animal/breed improvement. The study was therefore carried out to characterize and predict the protein structure of this gene with a view to better understand its functions for animal improvement purposes.

Materials and Methods

Experimental Location

The cattle sampling was purposively carried out across Oyo and Ogun States in Nigeria. The N'Dama breed were sampled at two locations; the Institute of Agricultural Research and Training Moor Plantation, Ibadan, off-site ranch at Ilora in Oyo State and the Federal Department of Livestock N'Dama Conservation Programme Ranch at Fashola in Oyo State. The Muturu breed were sampled at three locations; Odeda Local Government in Ogun State, Ipokia Local Government in Ogun State and Institute of Food Security, Environmental Resources and Agricultural Research facility at the Federal University of Agriculture, Abeokuta. Samples of White Fulani cattle were collected at three locations; the Cattle Production Venture of the Federal University of Agriculture, Abeokuta, Ajani Farms, Ogbomosho and Odeda Local Government, Ogun State.

Sampling

A total number of 90 animals were purposively sampled in this study. These comprised 30 White Fulani cattle, 30 Muturu and 30 N'Dama cattle. Information on the history of the animals sampled was obtained from the ranchers, with a view to ensure samples were drawn from unrelated individuals. Pregnant and nursing cows, likewise sick animals and newly born calves were not sampled.

Collection of blood samples for DNA analyses

5ml of blood were collected from the jugular vein of the animals with a syringe. The blood samples were collected from the restrained animals through the assistance of livestock attendants and research assistants. Each blood sample was stored in a plain sample bottle and refrigerated for DNA extraction which was carried out within 24hrs post collection.

DNA Extraction

Genomic DNA was extracted from the blood samples using Zymo-Spin IIC^{TM} extraction kit at the Central Biotechnology Laboratory of the Federal University of Agriculture, Abeokuta, and strictly following manufacturer's protocols.

DNA Quantification

The extracted DNA was quantified for concentration and purity using Nanodrop spectrophotometer in congruent with protocol reported by Desjaldins and Conklin (2010). After quantification the samples were kept at -4^{o}C for further analyses.

Primer design and DNA amplification

Bovine exons 1to 2 and exons 5 to 6 IRF3 gene specific primers were designed using Fast PCR software at Stab Vida Genetic Laboratory in Caprica-Portugal. Primer length, Primer sequence, annealing temperatures and the product sizes of the amplicons are presented in Table 1. For amplification, 10-20ng of genomic DNA was added to the reaction containing 0.4mM of primers forward and reverse, 1mM of each dNTPs, 1.5mM of $MgCl_2$ and 1.5uL Taq polymerase and was amplified by Magnetic Beads carboxylate cycler at the following conditions; one cycle of initial denaturation of 15 minutes at 96°C, final denaturation of 30 seconds at 95°C. Optimum annealing temperature of 60°C for 30 seconds, extension at 70°C for 2 minutes in 35 cycles

with one cycle of the final extension performed at 70°C for 5 min.

Table 1: Primer sequences used in the amplification of Bovine Interferon Regulatory Factor 3 gene exons 1-2 and exons 5-6

Primer	Length	Primer sequence	Amplicon size	TmºC
Exons1-2 (f)	19 bp	5' TCGGAAAACCTAAGAAGGG -3'	1335 bp	60º
Exons1-2 (r)	19 bp	5' ACCAGCCAACACAAATACC -3'		
Exons5-6 (f)	22 bp	5' CTGTCTTTTACTGTGCTGGTGG -3'	1029 bp	60º
Exons5-6 (r)	22 bp	5' CAGGTAAGGAGAGGGAGGAGAC-3'		

Sequencing of PCR products

The PCR products were purified using commercial kit (Magnetic Beads Carboxylate MC Lab, USA). The purified products were subjected to sequencing using BigDye® terminator cycle sequencing kit on ABI 3730xl (Applied Biosystems) DNA analyzer.

Trimming and cleaning of sequence

The sequences were trimmed and edited using Bioedit and MEGA 5 to remove the noise and to trim the sequences respectively.

Protein structure prediction

The nucleotide sequences of the IRF3 gene of each breed were translated to amino acid sequence using NCBI ORF finder (https://www.ncbi.nlm.nih.gov/orffinder/).The protein secondary structures were predicted using the GORIV software (Garnier *et al.,* 1996), while the tertiary structure were predicted using Phyre2 software (Kelley and Sternberg, 2009). The tertiary structures were then viewed with Rasmol.

Determination of physico-chemical properties of bovine IRF3 protein

Physico-chemical properties of bovine IRF3 protein were determined using Swiss Prot and Expert Protein Analysis System (ExPASy) which is proteomic server of Swiss Institute of Bioinformatics (SIB) (http://web.expasy.org). FASTA format of amino acid sequence of bovine IRF3 gene were used for the analysis. Various tools in the proteomic server (ProtParam,Protein Calculator, Compute pI/Mw and protscale) were used to deduce different physico-chemical properties such as number of amino acids residues, molecular mass, theoreticalpI, total number of negatively-charged residues (aspartate + glutamine), total number of positively-charged residues (arginine + lysine), extinction coefficient (M-1Cm-1) at 280 nm, instability index, aliphatic index and grand average hydropathy (GRAVY).

Results

Secondary protein structure for exons 1to 2 of the Interferon Regulatory Factor 3 gene in the three breeds of cattle.

Table 2 shows the secondary protein structure for exons 1 to 2 of the IRF 3 gene in the three breeds of cattle. The three breeds had the same alpha helix (25.49%), extended strand (12.75%) and random coil (61.76%) values.

Table 2- Secondary Protein Structures for exons 1 to 2 of the Interferon Regulatory Factor 3 gene in the three cattle breeds

Breed	Alpha helix (%)	Extended Strands (%)	Random coils (%)
White Fulani	25.49	12.75	61.76
N'Dama	25.49	12.75	61.76
Muturu	25.49	12.75	61.76

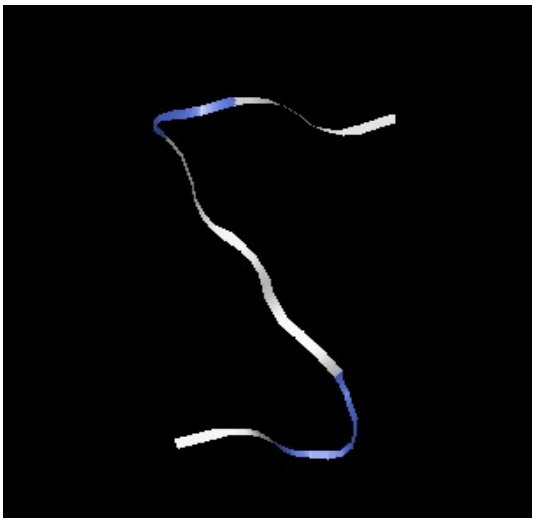

Plate 1: Tertiary protein structure of exons 1to 2 of the IRF3 gene in the White Fulani cattle in ribbon model.

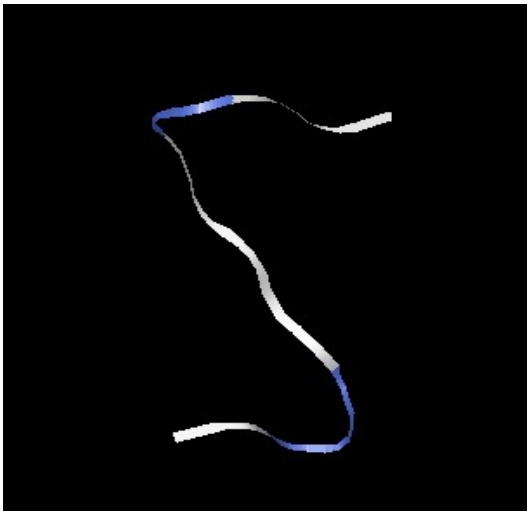

Plate 2: Tertiary protein structure of exons 1to 2 of the IRF3 gene in the Muturu cattle in ribbon model.

Physico-chemical properties of the translated protein of exons 1to 2 of the Interferon Regulatory Factor 3 gene in the three cattle breeds

Table 3 represents the physico-chemical properties of the protein translated by exons 1to 2 of the IRF 3 gene in the three cattle breeds. Due to the fact that the three cattle breeds shared identical secondary protein structures for this regions, their physico-chemical properties were also observed to be identical.

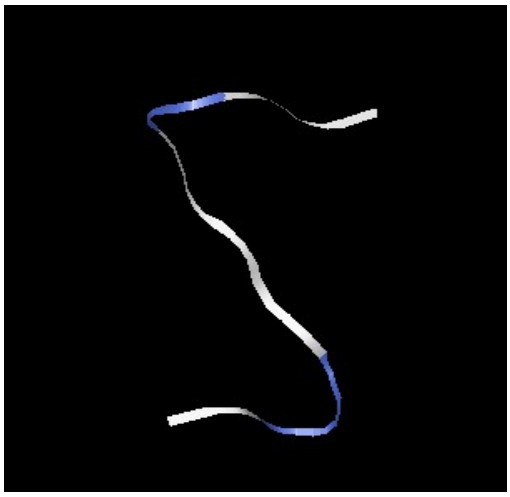

Plate 3: Tertiary protein structure of exons 1to 2 of the IRF3 gene in the N'Dama cattle in ribbon model.

The total number of amino acids translated in each breed was 102, with a molecular weight of 10562.87Da, they all had the same theoretical pI value of 7.75, their total number of positively charged residues (Asp+Glu) and negatively charged residues (Arg+Lys) were 11 and 12 respectively. They had a common molecular formular of $C_{451}H_{717}O_{136}S_{5,}$ with a total number of 1458 atoms. Their instability index (II) was 44.92 making them unstable. They also had a common aliphatic index and GRAVY value of 66.27 and -0.365, respectively.

Secondary protein structure for exons 5 to 6 of the Interferon Regulatory Factor 3 gene in the three breeds of cattle.

Table 4 shows the secondary protein structure for the exons 5 to 6 of the IRF3 gene in the three cattle breeds. The N'Dama breed had the highest value of 29.20% for the alpha helix, followed by the Muturu with 15.44% and the White Fulani cattle had the least value of 6.61%. With respect to the extended strand, the reverse was the case with the White Fulani cattle having the highest value of 25.62%, followed by the Muturu with 23.53% and the N'Dama breed having the lowest value of 5.31%. The White Fulani cattle still had the highest value for the random coil (67.77%), followed by the N'Dama (65.49%) and the Muturu had the lowest value of 61.03%.

Table 3: Physico-chemical properties of the exon 1to 2 protein in the three cattle breeds

Properties	Breed WF	MT	ND
No of amino acids	102	102	102
Molecular weight(Da) 10562.87	10562.87	10562.87	
Theoretical p.I	7.75	7.75	7.75
Total number of negatively charged residues (Asp+Glu)	11	11	11
Total number of positively charged residues (Arg+lys)	12	12	12
Formular	$C_{451}H_{717}O_{136}S_5$	$C_{451}H_{717}O_{136}S_5$	$C_{451}H_{717}O_{136}S_5$
Total number of atoms	1458	1458	1458
Instability index(II)	44.92	44.92	44.92
Aliphatic Index	66.27	66.27	66.27
Grand average of hydropathy(GRAVY)	-0.365	-0.365	-0.365

WF-White Fulani, MT-Muturu; ND-N'Dama

Table 4: Secondary protein structure for exons 5 to 6 of the Interferon Regulatory Factor 3 gene in the three Cattle breeds

Breed	Alpha helix (%)	Extended strand (%)	Random coil (%)
WF	15.44	23.53	61.03
ND	29.20	5.31	65.49
MT	6.61	25.62	67.77

WF-White Fulani, ND-N'Dama, MT-Muturu

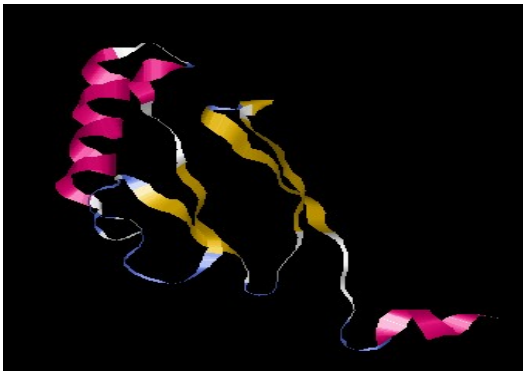

Plate 4: Tertiary protein structure of exons 5 to 6 of the IRF3 gene in Muturu cattle in ribbon model

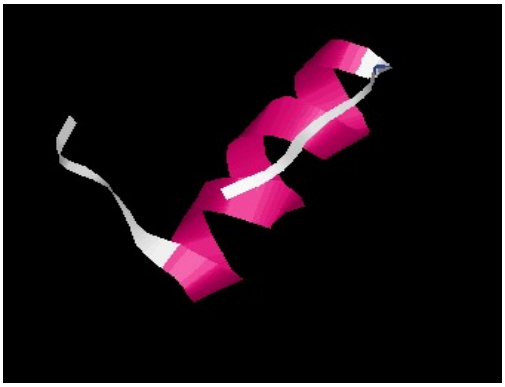

Plate 5: Tertiary protein structure of exons 5 to 6 of the IRF3 gene in N'Dama cattle in ribbon model.

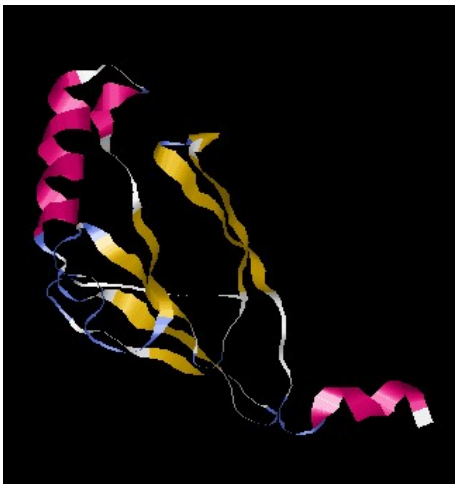

Plate 6: Tertiary protein structure of exons 5 to 6 of the IRF3 gene in White Fulani cattle in ribbon model

Physico-chemical properties of the protein translated by exons 5 to 6 of the Interferon Regulatory Factor 3 gene in the three cattle breeds

Table 5 represents the physico-chemical properties of exons 5 to 6 of the IRF 3 gene in the three cattle breeds. The Muturu cattle had the highest number of amino acids translated (136) with corresponding molecular weight (14,294.31Da), followed by the White Fulani cattle with 121 amino acids and a molecular weight of 12,932.67Da, the N'Dama cattle had the lowest values for both amino acids translated and molecular weight (113 and 12,065.64Da, respectively).

The Muturu cattle had the highest theoretical pI value of 9.12; the White Fulani had a value of 7.54, while the N'Dama had the lowest theoretical pI value of 4.72. With respect to the total number of positively-charged residues (Asp+Glu), the N'Dama had the highest number of such residues (14), the Muturu had 10 while the White Fulani had the lowest number of 9. The Muturu cattle had a total number of 15 negatively charged residues (Arg+Lys) which was the highest, followed by the White Fulani (10) and the N'Dama had the lowest number of 8 of such residues. From their respective protein formulae, the Muturu exon 5-6 translated protein had the highest number of atoms (1991), the White Fulani cattle had 1770 while the N'Dama had the lowest atoms number of 1678. The N'Dama translated protein in this region of the IRF 3 gene appeared to be the most unstable with an instability index (II) value of 68.10, this was followed by that of the Muturu with an II value of 50.28, the White Fulani had the least value of 40.97, though it was also considered to be unstable. With respect to the aliphatic index values, the N'Dama cattle translated protein had the highest value of 76.02; the Muturu was in second place with 71.69 while the White Fulani had the lowest value of 57.27.

Table 5 also revealed the GRAVY values for the various proteins translated in this region of the IRF 3 gene in the three cattle breeds with the N'Dama having the highest value of -0.432, the White Fulani had -0.254 while the Muturu recorded a GRAVY value of -0.163.

Discussion

Proteins are the most versatile macromolecules in living systems and serve crucial functions in essentially all biological processes. Exon 1 of the IRF3 gene was discovered to be an untranslated region (UTR), this could have been responsible not only for the sequence homology observed among the three breeds but also similar secondary and tertiary protein structures for the three breeds. Holm and Rosenstrum (2010) stated that shared structure between proteins was considered evidence of evolutionary relatedness between proteins and can be used to group proteins into super families. Vilela and McCarthy (2003) stated that though UTRs do not form the protein-coding region of the gene, upstream open reading frames (uORFs) located within the 5' UTR can be translated into peptides. Mutations in untranslated regions have

Table 5: Physico-chemical properties of the exon 5 to 6 Interferon Regulatory Factor 3 protein in the three cattle breeds

	Breed		
Properties	**WF**	**MT**	**ND**
No. of amino acids	121	136	113
Molecular weight(Da)	12932.67	14294.31	12065.64
Theoretical p.I	7.54	9.12	4.72
Total number of negatively charged residues (Asp+Glu)	9	10	14
Total number of positively charged residues (Arg+lys)	10	15	8
Formular	$C_{573}H_{862}N_{162}O_{164}S_9$	$C_{620}H_{990}N_{188}O_{185}S_8$	$C_{519}H_{836}N_{146}O_{170}S_7$
Total number of atoms	1770	1991	1678
Instability index(II)	40.97	50.28	68.10
Aliphatic Index	57.27	71.69	76.02
Grand average of hydropathy(GRAVY)	-0.254	-0.163	-0.432

WF-White Fulani, MT-Muturu, ND-N'Dama

been reported to be associated with increased risk for developing diseases (Zhu *et al.,* 2015).

There was no sequence homology observed in exons 5to 6 of the candidate gene, with the three breeds having different amounts of amino acids and subsequently different secondary and tertiary protein structures and functions at this region. The similar secondary protein structures for exons 1to 2 of the candidate gene in cattle earlier could be attributed to the UTR nature of this region of the gene, which could also be responsible for their identical physico-chemical properties. The theoretical pI (Isoelectric potential) of the exons 1to 2 protein of the IRF3 gene deduced from White Fulani, N'Dama and Muturu cattle were evidence that the protein of exon 1-2 in the IRF3 gene is basic in nature. The isoelectric potential of a protein has been described as an important property because it is at this point that the protein is least soluble (Rawiwan *et al*., 2022). It is a useful property of a protein to consider for optimized crystallization and a significant correlation between pI of a protein and pH of the buffer being used for crystallization as reported by (Kantardjieff and Rupp, 2004).

The higher fraction of positively charged residues in this region of the candidate gene in the three cattle breeds supported the fact that it is an exracellular protein. Extracellular proteins have been implicated in many important activities, such as, elimination of extracellular pathogens while minimizing damage to the host cells (Brinkman *et al.,* 2004). Schmitt *et al.* (2006) also, reported extracellular located or membrane-bound heat shock proteins (HSPs) involvement in mediating immunological functions, eliciting immune responses modulated either by the adaptive or innate immune system, therefore, the higher fraction of positively-charged residues in this region of the candidate gene in the three cattle breeds supported the immune system function of the IRF3 gene.

The instability index of the protein of exons 1 to 2 of the candidate gene in the three breeds suggested an unstable protein. Guruprasad *et al.* (1990) has earlier stated that instability index less than 40 predicts a stable protein, whereas, values higher than 40 denotes a potentially-unstable protein. The aliphatic index is a measure of the relative volume occupied by aliphatic side chains; alanine, valine, isoleucine, and leucine (Panda and Chandra, 2012). The aliphatic index values obtained for the protein of this region in the three breeds were slightly above average, a higher aliphatic index is an indicator of higher thermostability and also an indicator of solubility in a cell when the protein is over expressed (Ikai, 1980). The negative GRAVY score obtained for the protein of this same region of the candidate gene in White Fulani, N'Dama and Muturu were indications that it is polar and would most likely tend to be hydrophilic. The physico-chemical properties of exons 5 to 6 of the IRF3 gene in the three cattle breeds were unidentical unlike that of exons 1 to2. The theoretical pI (Isoelectric potential) of the exons 5 to 6 protein of the IRF3 gene deduced from White Fulani and Muturu cattle suggested that the protein is basic unlike that of the N'Dama which was acidic due to its low value. The White Fulani and Muturu also had a higher number of positively charged residues compared to their negatively charged residues, but the reverse was the case in the N'Dama, this implied that the proteins of the exons 5 to 6 in the IRF3 gene in the White Fulani and Muturu were extracellular in nature, while that of the N'Dama was intracellular. Intracellular proteins have also been indicated in many important activities; members of intracellular Galectin family have been shown to regulate cell growth and apoptosis and also regulate cell cycle (Liu *et al.,* 2002); intracellular heat shock proteins have also been reported to have a protective function; allowing cells to survive lethal conditions (Schmitt *et al.,* 2006). The instability index of the protein of exons 5 to 6 of the candidate gene in the three

breeds also suggests an unstable protein with all three having values higher than 40 which were still in line with Guruprasad *et al.* (1990). Unstable proteins have been reported to be prone to misfolding which has also been associated with some degenerative diseases in humans (Reynaud, 2010). The unstable nature of the IRF 3 exons 1 to 2 and 5 to 6 proteins could therefore, play a part in disease susceptibility in the three cattle breeds. Negative GRAVY scores were also obtained for the protein of this same region of the candidate gene in the three breeds which were indications that it was also polar and would most likely be hydrophilic. Hydrophilic proteins, also known as, late embryogenesis abundant (LEA) proteins have been reported to accumulate under conditions of extreme desiccation in higher plants, helping to cope with stressful situations (Bray, 1997; Chakrabortee *et al.,* 2007 and Battaglia *et al*., 2008). The hydrophilic nature of the proteins of exons 1 to 2, 5 to 6 may therefore play similar role in the three cattle breeds in the aspect of heat tolerance and adaptability to dry environmental conditions.

Protein Secondary structure refers to highly regular local sub-structures on the actual polypeptide backbone chain. Components of the secondary structure includes; the α-helix and the β-strand or β-sheets and random coils (Rehman *et al.,* 2022). These secondary structures are defined by patterns of hydrogen bonds between the main-chain peptide groups. The Protein structure of exons 1to 2 in the IRF3 gene in Muturu, N'Dama and White Fulani can be classified as a mixed folding structure which contains alpha helixes, extended strands and random coils. Random coils dominate among the secondary structure elements followed by the alpha helix and extended strand. These structural elements observed in Muturu, N'Dama and White Fulani at exons 1to 2 protein of the IRF3 gene were responsible for the folding, stability and the overall functions of the protein. Structural similarities observed between the secondary and tertiary protein structures of exons 1to 2 in the White Fulani, N'Dama, Muturu IRF3 gene implied that the genes shared similar molecular and biological functions like; immune response, in the three breeds. The predicted protein structure of exons 5to 6 in the IRF3 gene in Muturu, N'Dama and White Fulani can also be classified as a mixed folding structure. Random coils dominated among the secondary structure elements followed by the alpha helix and extended strand in the White Fulani and Muturu, but in the case of the N'Dama, a higher alpha helix value was obtained, when compared to the extended strand which may be attributed to the N'Dama having the least number of amino acids for the protein of this region. Tertiary structure refers to the three-dimensional structure of the entire polypeptide chain and this determines the function of the protein (Engelking, 2015), but unlike in the exons 1to 2 of the IRF3 gene in the three breeds which had identical secondary and tertiary structures, also, had different tertiary structures for the predicted exons 5 to 6 protein of the candidate gene suggesting likely different functions.

Conclusions

The results obtained from this study, indicated that the three cattle breeds used for the study have identical predicted secondary and tertiary protein structures and subsequently identical physic-chemical properties for exons 1to 2 but not for exons 5 to 6 of the IRF3 protein in cattle. Prediction of secondary protein structures of the loci under consideration showed mixed folding proteins, which contained alpha helices, extended strand and random coils. It is recommended that; the roles of IRF3 gene in animal reproduction and development should be further explored and Polymorphism in the IRF3 gene and their association with diseases of economic importance in Nigerian cattle should also be studied.

Acknowledgements
Appreciation goes to the cattle breeders and ranchers in Ipokia Local Government in Ogun State, the Federal Department of Livestock N'Dama Conservation Programme Ranch at Fashola in Oyo State, The Institute of Agricultural Research and Training Moor Plantation, Ibadan, off-site ranch at Ilora in Oyo State, cattle breeders and ranchers in Odeda Local Government in Ogun State, for their assistance in allowing us to sample their available Cattle. We also thank the directors and staffers of the Central Biotechnology laboratory for allowing us analyze the cattle blood samples, Cattle Production Venture(CPV) and Institute of food Security Environmental Resources and Agricultural Research (IFSERAR) all located at the Federal University of Agriculture, Abeokuta, Ogun State, Nigeria for allowing us use their available cattle for this research work.

Conflict of Interest Statement: The authors hereby declare no conflict of interest during and after the study was completed.

Funding: This work was supported by a grant from Tetfund 2022/2023 merged. But the funding body had no role in the planning, design and execution of the research work.

References

Battaglia, M.; Olvera-Carrillo,Y.; Garciarrubio, A.; Campos, F. and Covarrubias, A.A. 2008. The Enigmatic LEA Proteins and Other Hydrophilins. *Plant Physiol.* 148(1): 6–24

Bray, E. A. 1997. Plant responses to water deficit. *Trends Plant Sciences,* 2: 48–54.

Brinkmann, V.; Reichard, U.; Goosmann, C.; Fowler, B.; Uhlemann, Y.; Weiss, D.S.; Weinrauch, Y. and Zychlorsky, A. 2004. Neutrophil extracellular traps kills bacteria. *Science* 303(5663): 1532-1535.

Chakrabortee, S.; Boschetti, C.; Walton, L.J.; and Tunnacliffe, A. 2007. Hydrophilic protein associated with desiccation tolerance exhibits broad protein stabilization function. Proceedings of the National Academy of Science. 104 (46): 18073-18078

DAGRIS, 2005: http://dagris.ilri.cgiar.org

Engelking, L. R. 2015. Protein Structure. *Veterinary Physiological Chemistry* (Third Edition), Academic Press.18-25

Ensembl cow release 92.http://m.ensembl.org

Garnier, J., Gibrat, J. and Roson, B. 1996. GORIV method for predicting protein secondary structure from amino acid sequence. Method in enzymology 266: 540-553.

Guruprasad, K.; Reddy, B.V. and Pandit, M.W. 1990. Correlation between stability of a protein and its dipeptide composition: A novel approach for predicting in vivo stability of aprotein from its primary sequence. *Protein Engineering* 4: 155–161.

Holm, L. and Rosenstrum, P. 2010. Dali server: conservation mapping in 3D. *Nucleic Acids Research* 38: W545-549.www.ncbi.nlm.nih.gov/pubmed/20457744.

Ikai, A. 1980. Thermostability and aliphatic index of globular proteins. *Journal of biochemistry* (Tokyo) 88(6): 1895-1898

Jann, O.C.; King, A.; Corrales, N.L.; Anderson, S.I.; Jensen, K.;Ait-Ali, T.; Tang, H.; Wu, C.;Cockett, N.E., Archibald, A.L.;and Glass, E.J. 2009. Comparative genomics of toll like receptor signaling in five species. *BMC Genomics* 10:216.

Kantardjieff, K.A. and Rupp, B. 2004. Protein isoelectric point as a predictor for increased crystallization screening efficiency. *Bioinformatics* 20(14): 2162-2168.

Kelly, L.A. and Sternberg, M.J.E. 2009. Protein structure and prediction on the web; a case study using the Phyre sever. *Nature protocols-journal* 4:363-371.

Liu, F.T.; Patterson, R.J. and Wang, J.L. 2002. Intracellular functions of galectins. *Biochimica Biophysica Acta* 1572(2-3): 263-273.

Mattioli, R.C.; Pandey, V.S.; Murray, M. and Fitzpatrick, J.L. 2000. Immunogenetic influences on tick resistance in African cattle with particular reference to trypanotolerant N'Dama(*Bostaurus*) and trypanosusceptible Gobra zebu (*Bos indicus*) cattle. *ActaTropica*75: 263–277.

McTavish, E.J.; Decker, J.E.; Schnabel, R.D.; Taylor, J.F.; and Hillis, D.M. 2013."New World cattle show ancestry from multiple independent domestication events". Proc. Natl. Acad. Sci. U.S.A. 110: E1398–406

Murray, M.; Morrison, W., and Whitelaw, D. 1982. Host susceptibility to African trypanosomiasis: Trypanotolerance. *Advanced Parasitology*21:1–68.

Murray, M.; Trail, J.C.M.; Davis, C.E. and Black, S.J. 1984. Genetic resistance to African trypanosomiasis.*Journal of Infectious Diseases*149:311–319.

Njogu, A.; Dolan, R.; Wilson, A. andSayer, P. 1985. Trypanotolerance in East African Orma Borancattle.*Veterinay Record,*117:632–636.

Norezzine, A.; Duksi, F.; Tsvetkova, A.D.; Ulybina, E.A.; Gins, M.S.; Yacer, R.N.; Klenovitsky, A.A. ; Nikishov, A.A.; Amirshoev, F.; Digha, J. and Gladyr, E.A. 2019. Genetic characterization of White Fulani cattle in Nigeria: A comparative study. *J Adv Vet Anim Res.* 6(4): 474–480.

Panda, S. and Chandra, G. 2012. Physicochemical characterization and functional analysis of some snake venom toxin proteins and related non-toxin proteins of other chordates. Bioinformation. 8(18): 891–896.

Pauling, L.; Corey, R.B. and Branson, H.R. 1951."The structure of proteins; two hydrogen bonded helical configurations of the polypeptide chain". Proc Natl Acad Sci USA 37,(4): 205–211.

Rawiwan, P.; Peng, Y.; Paramayuda, I.G.P.B.; and Quek, S.Y. 2022. Red seaweed: A promising alternative protein source for global food sustainability. Trends in food science and technology.123:37-56

Rege, J.E.O. and Tawah, C.L. 1999. The state of African cattle genetic resources II. Geographical distributions, characteristics and uses of present-day breeds and strains. *FAO/UNEP Animal Genetic Resources Information Bulletin*.26**:**1-25.

Rehman, I.; Farooq, M. and Botelho, S. 2022. Biochemistry, Secondary Protein Structure. Stat pearl publishing. National Library of Medicine. www.ncbi.nlm.nih.gov/books/n/statpearls/

Reynaud, E. 2010. Protein Misfolding and Degenerative Diseases. *Nature Education* 3(9): 28.

Schmitt, E.; Gehrmann, M.; Multhoff, G. and Gerrido, C. 2006. Intracellular and extracellular functions of heat shock proteins: repercussions in cancer therapy. *Journal of Leukocyte Biology* 81(1): 15-27.

Tamura, T.; Yanai, H.; Savitsky, D. and Taniguchi, T. 2008.The IRF family transcription factors in immunity and oncogenesis. *Annual Review of Immunology,* 26: 535–58410.

Vamathevan, J.J.; Hasan, S.;Emes, R.D.;Amrine-Madsen, H.;Rajagopalan, D.; and Topp, S.D., 2008. The role of positive selection in determining the molecular cause of species differences in disease. BMC *Evolutionary Biology,* 8(273).

Vilela, C. and McCarthy, J. E. G. 2003. "Regulation of fungal gene expression via short open reading frames in the mRNA 5'untranslated region". *Molecular Microbiology,* 49(4): 859–867.

Zhu, Q.; Chen, Y.; Dai, J.; Wang, B.; Liu, M.; Wang, Y.; Tao, J.; Li, H. 2015. "Methylenetetrahydrofolate reductase polymorphisms at 3'-untranslated region are associated with susceptibility to preterm birth". *Translational Pediatrics,* 4(1): 57–62.

Date received: 5th February, 2024
Date accepted: 16th July, 2024.