ISSN:
Print - 2277 - 0755
Online - 2315 - 7453

Journal of Agricultural Science and Environment

PHYLOGENETICS OF ELONGATION FACTOR-G MITOCHONDRIAL PROTEIN GENE (*GFM1*) IN TEN SELECTED SPECIES

A.T. ABOLUDE[1], O. OLOWOFESO*[1], S.O. PETERS[1, 3], M.O. OZOJE[1], O.F. SMITH[2] AND A.S. ADENAIKE[1]

[1]Department of Animal Breeding & Genetics,
[2]Department of Animal Physiology,
Federal University of Agriculture, Abeokuta, Postcode 110001. Nigeria.
[3]Department of Animal and Range Sciences,
New Mexico State University, Las Cruces, NM 88003, USA
Corresponding author: olowofesoinunaab@gmail.com **Tel:**+2348060223098

ABSTRACT

The second stage of protein synthesis is elongation. One of the elongation factors in the elongation cycle of protein synthesis is the elongation factor-G (*GFM1*). *GFM1* is an ancient translational GTPase (trGTPase); the bacterial homolog of eukaryotic eEF2 and archaeal aEF2, respectively. It may interact with the transcriptional apparatus as a positive regulator of RNA synthesis in various species. Genetic variations in *GFM1* gene of ten species including cattle, human, chicken, mouse, rat, horse, zebra fish, honeybee, pig and rabbit based on availability were investigated using bioinformatic approach. Using a comparative genomic approach, 4,442 base pairs (bp) of the *GFM1* sequences were obtained. Alignment of the sequences within the region of 3,626 bp and containing 816 gaps was carried out using Clustal W software. A very close relationship between rabbit and pig was observed in the phylogenetic tree of *GFM1* gene which showed that the comparability of *GFM1* gene sequence was highest between the two species and they evolved from a most recent common ancestor with respect to *GFM1* gene. Cattle, human, rat and zebra fish were closest by their genetic distances to the ancestor, while mouse, horse, chicken, rabbit and honeybee were distant from the common ancestor. However, close phylogenetic relationship among species might be as a result of conservation of the sequence in the various species.

Keywords: Elongation factor-G, phylogenetic, protein gene, species

INTRODUCTION

Initiation, elongation and termination are the three important stages of protein synthesis. Each of the stages requires three factors. Initiation factors IF_1, IF_2 and IF_3; elongation factors EF-Tu, EF-Ts and EF-G, and termination or release factors RF_1, RF_2 and RF_3, respectively. The elongation step of mitochondrial translation, best documented in mammals, closely mimics the prokaryotic system (Hirokawa *et al.,* 2005). It is an essential protein with central roles in both the elongation and ribosome recycling phases of protein synthesis. The main function of the translation elongation factor-Tu (EF-Tu) and its eukaryotic counterpart eEF1A is to deliver aminoacyl-tRNA to the A-site on the ribosome during which energy is

consumed by the hydrolysis of GTP (Zipfel *et al.,* 2006). One of the goals of biology has been the creation of taxonomy for living things, a method of organizing species in terms of their relationships to one another. Early biologists classified species solely according to their morphology (the physical appearance of the organism) and later, as dissection became a more common practice, their anatomies were used (Benjamin, 2008). Molecular evolution extends the concept of evolution to the level of DNA and protein sequences. Although the replication of DNA sequence is a very accurate process, small replication errors accumulate over time, along with radiation damage and other mutations or alterations of the genomic sequence. The chemical processes of molecular evolution are responsible for more than just giving rise to species differences. Evolutionary change can occur within the genome of a single species as well. Orthologues are genes that are evolutionarily related, share a function, and have diverged by speciation. Paralogues, on the other hand, have a common ancestor but have diverged by gene duplication and no longer have a common functional role. In other words, orthologues have the same function but occur in different species, while paralogues exist in the same genome but have different functions (Veerassamy *et al.,* 2003). Hence, molecular phylogeny looks for similarities in the molecules of organisms to figure out relationships.

Bioinformatic analysis is the first step in an attempt to explicate the genetic variation associated with the expression of *GFM1* protein gene as a result of accumulation of biological data. Bioinformatic analysis when combined with sequencing and expression studies is a very powerful validation for previously detected quantitative trait loci (QTL) as recently elucidated by Adenaike (2011). The objective of this study was to perform bioinformatic analysis on elongation factor-G mitochondrial protein gene in cattle with nine other common existing species with the purpose of validating their genetic relationship at the elongation factor-G locus and constructing the phylogenetic tree among the species.

MATERIALS AND METHODS

DNA and protein sequences representing the *GFM1* protein gene belonging to the following species were retrieved from Genebank. Nucleotide sequences of *Bos taurus* (cattle*) Gallus gallus* (chicken) *Oryctolagus cuniculus* (rabbit) *Danio rerio* (zebra fish) *Apis mellifera* (honeybee) *Equus caballus* (horse) *Homo sapiens* (man) *Mus musculus* (mouse) *Sus scrofa* (pig) and *Rattus norvegicus* (rat) were retrieved. This was done first, by obtaining the FASTA format (a text-based format representing either nucleotide sequences or peptide sequences, in which nucleotides or amino acids are represented using single-letter code) of the nucleotide and protein sequence of *GFM1* protein gene at the National Centre for Biotechnology Information database (www.ncbi.nlm.nih.gov) and Basic Local Alignment Search Tool, BLAST which is an algorithm for comparing primary biological sequence information, such as the amino-acid sequences of different proteins or the nucleotides of DNA sequences. DNASP (version 5.0) was used to estimate genetic distance among these species based on their DNA sequences of the GFMI.

Selection of the ten species was based on the availability of the species possessing the *GFM1* gene. Phylogenetic tree was constructed with the GeneBee tool available at http://www.genebee.msu.su/services/phtree_reduced.html.

Table 1: Genetic distance matrix among ten species based on DNA and protein sequences of the *GFM1* protein gene retrieved from Genebank*

Species	Bos	Gallus	Oryctolag	Danio	Apis	Equus	Homo	Mus	Sus	Rattus
Bos	0.0000									
Gallus	1.0000	0.0000								
Oryctolag	1.0000	1.0000	0.0000							
Danio	0.1870	1.0000	1.0000	0.0000						
Apis	1.0000	1.0000	1.0000	1.0000	0.0000					
Equus	0.9610	1.0000	1.0000	0.9670	1.0000	0.0000				
Homo	0.0620	1.0000	1.0000	0.1870	1.0000	0.9670	0.0000			
Mus	0.9130	1.0000	1.0000	0.9220	1.0000	1.0000	0.9170	0.0000		
Sus	1.0000	1.0000	0.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.0000	
Rattus	0.1220	1.0000	1.0000	0.2420	1.0000	0.9670	0.1430	0.9260	1.0000	0.0000

*Names of species are defined within the text.

The phylogenetic tree of the *GFM1* protein gene among the ten species is shown in Figure 1.

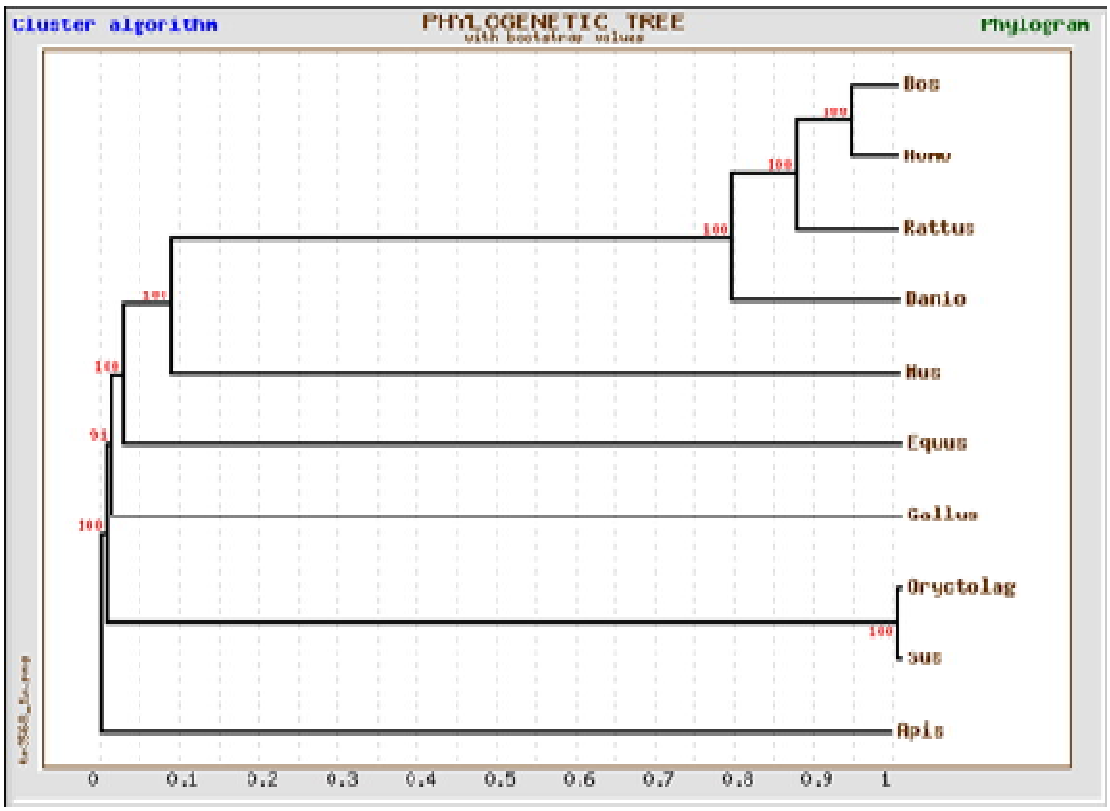

Figure 1: Phylogenetic tree showing the level of relationship among the ten selected species based on *GFM1* protein gene data retrieved from Genebank

RESULTS AND DISCUSSION

The distance matrix generated using the DNA of the *GFM1* protein gene data among the ten selected species is presented in Table 1.

The tree shows evolutionary relationship among the *GFM1* protein sequences of the species. The phylogenetic findings confirmed the existence of high genetic polymorphism which was in accordance with the results of Farrell *et al.* (2004) and Alinaghizadeh *et al.* (2007) who reported that polymorphism in the gene is due to the genetic variation and its variants are transmitted by simple Mendelian inheritance with no dominance. The species joined together in the tree have common ancestor. The nodes with descendants show most recent common ancestor on the locus of the *GFM1* gene. Species with short distance values imply a recent relationship with the common ancestor at the *GFM1* gene locus with the highest distance value being 1. However, short genetic distance could imply conservation of the sequence over time. Cattle with chicken, rabbit, honeybee and pig had a distance value of 1.0000 which shows that they have no recent common ancestor. Cattle with horse and mouse had distance values of 0.9610 and 0.9130, respectively. This also signifies that either they do not have a recent common ancestor or the sequence was not conserved while cattle with rat and zebra fish have distance values of 0.1220 and 0.1870; meaning that they have very recent common ancestor and with human, a distance value of 0.0620 having the most recent common ancestor or most conserved region. Chicken and honeybee with all the other species had a distance value of 1.0000 indicating that they have no recent common ancestor with any of the selected species or the sequence had totally changed. Rabbit also had distance value of 1.0000 with all the species except pig with distance value of 0.0000; thus signifying that they are of the same origin or the sequence was totally conserved. Rat showed a high level of distance from horse (0.9670) and mouse (0.9260) but showed a close relationship with cattle (0.1220), zebra fish (0.2420) and human (0.1430). Zebra fish with horse and mouse have very high distance values (0.9670 and 0.9220) respectively. This implies that they do not have any recent common ancestor at that very locus. Human with horse and mouse have high distance values of 0.9670 and 0.9170; while with zebra fish, a low distance value of 0.1870, all with respect to the *GFM1* protein gene sequence available. Some relationships shown by the dendrogram are expected although, the most striking result is the clustering of rabbit and pig and the contrasting result in mouse and rat despite the fact that they are both rodents. It can be seen that sequences of rabbit and pig *GFM1* gene shared a common ancestor or their gene sources are likely to be same. This agreed with theory of genealogical processes developed by Tajima (1983), predicting that if the time of divergence among species is short, there is a higher probability that the evolution of sequences under mutation and random genetic drift will produce paraphyletic rather than monophyletic relationships. A similar observation of this nature has also been reported by Adenaike (2011).

REFERENCES

Adenaike, A.S. 2011. Bioinformatic analyses of kappa casein (CSN3) gene in twenty-four mammalian species. Unpublished M. Agric Dissertation, Federal University of Agriculture, Abeokuta, Nigeria, pp. 126.

Alinaghizadeh, R., Mohammad, A.M., Moradnasab, B.S. 2007. Kappa Casein Gene Study in Iranian sistani cattle Breed using PCR-RFLP. *Pakistan Journal of Biological Science,* 10(23): 4291-4294.

Benjamin, A.P. 2008. *Genetics: A Conceptual Approach*. Freeman and Company, New York**,** pp.710.

Farrell, H.M., Jimenez-Florez, R., Bleck, G.T., Brown, E.M., Butler, J.E., Creamer, L.K., Hicks, C.M., Hollar, C. L., Ng-Kwai-Hang, K.F., Swaisgood, H. E. 2004. Nomenclature of the proteins of cows' milk. *Journal of Dairy Science,* 87: 1641–1674.

Hirokawa, G., Nijman, R.M., Raj, V.S., Kaji, H., Igarashi, K., Kaji, A. 2005. The role of ribosome recycling factor in dissociation of 70S ribosomes into subunits, *RNA,* 11:1317-1328.

Tajima, F. 1983. Evolutionary relationship of DNA sequences in finite populations. *Genetics,* 105: 437-460.

Veerassamy, S., Smith, A., Tillier, E. R. 2003. A transition probability model for amino acid substitutions from Blocks. *Journal of Computational Biology,* 10**:** 997-1010.

Zipfel, C., Kunze, G., Chinchilla, D., Caniard, A., Jones, J. D., Boller, T., Felix, G. 2006. Perception of the bacterial PAMP EF-TU by the receptor EFR restricts agrobacterium-mediated transformation. *The Cell* 125: 749-760.

(Manuscript received: 8th December, 2011; accepted: 1st June, 2012).