Original Research Article

Bioinformatic Analysis of Insulin-Like Growth Factor-1 Gene of Three Avian Species

Egom, M.A., *Ndofor-Foleng, H.M., Iloghalu, O.G., Ikeh, N.E., Onodugo, M.O., Okenyi, N. and Agu, C.I.

Department of Animal Science
University of Nigeria
Nsukka, Nigeria

*Corresponding author: harriet.ndoforfoleng@unn.edu.ng; +2349025855148

Received 6th December, 2018; Accepted 6th February, 2019; Corrected 10th February, 2019

Abstract

In recent times, attention has been focused on the study of Insulin-like Growth factor-1 (*IGF-1*) due to its biological functions such as; stimulating systemic body growth, regulating cell growth and cell development. A bioinformatics study was done to investigate the Insulin-like Growth Factor-1 gene of some avian species (turkey, chicken and quail). Nucleotide sequences and their corresponding proteins of the insulin-like growth factor-1 gene were obtained from the Genebank (a public domain protein database) and were analyzed using various software tools (Clustal W, MEGA 6, dnaSP, and BLAST) to determine the percent identity and similarities in function, genetic diversity and evolutionary relationship of the *IGF-1* gene. From the results seen in this study, percent identity and similarity of the *IGF-1* gene in avian ranged from 86-99%, thus indicating similarity in function in the species. Also genetic diversity was high within each avian (1.000 in turkey, 0.900 in chicken and 1.000 in quail). However chicken had the highest haplotype number value (4), an indication that chicken has more variation than turkey and quail in *IGF-1* gene sequence. Phylogenetic study showed that the *IGF-1* gene sequence of avian were grouped into the same taxon, chicken and quail shared a most recent common ancestor and were closely related than the *IGF-1* gene sequence of turkey. In conclusion, the high percent identity and similarity in function, high genetic diversity, and a close relative relatedness in the phylogentic tree of *IGF-1* gene seen in this study make the gene highly effective in improving growth and regulating cellular activities.

Keywords: Bioinformatics, Insulin-like Growth factor-1, GeneBank, avian

Introduction

The insulin-like growth factors (IGFs) are synthesized by almost all tissues, and are important mediators of cell growth, differentiation, transformation and many biological effects. They increase the absorption of glucose, stimulate myogenesis, inhibit cell cycle genes, increase the synthesis of lipids, and stimulate the production of progesterone in the synthesis of DNA, RNA and protein

(Etherton, 2004; Hegarty *et al.,* 2006). As a result of the wide range of their biologic effects and their therapeutic potential, the IGFs have become the focus of research by an increasing number of investigators.

Circulating *IGF-1* is produced by the liver under the control of growth hormone as an endocrine hormone as well as in target tissues in a paracrine or autocrine manner (Kemp, 2007). The binding of growth hormone with its hepatic receptor stimulates expression and release of *IGF-1* peptide in the circulation, which has high affinity for IGFBPs (insulin growth factor binding proteins), and represents the endocrine form of *IGF-1* (Sakai *et al.,* 2001). The *IGF-1* released stimulates systemic body growth and has growth-promoting effects on almost every cell in the body system (Yilmaz *et al.,* 2011). Deficiency of either growth hormone or *IGF-1* therefore results in diminished stature (Akinfenwa *et al.,* 2011). Different researchers have established a link between circulating *IGF-1* concentration and growth trait in many livestock species and laboratory animals (Bertlett and Toms, 2005; Bunter *et al.,* 2005; Hegarty *et al.,* 2006).

Bioinformatics involves discovery, development and implementation of computational algorithms and software tools that help to facilitate an understanding of various biological processes with the goal to serve primarily agriculture and health care sectors with several spinoffs (Bruce *et al.,* 2002). In a developing country like Nigeria, bioinformatics plays an important role in agriculture where it can be used to analyze livestock genomic and proteomic data that can be very useful in making genetic improvements. Computational analysis greatly helps in understanding the molecular basis of the biological functions of proteins through the use of available information to understand the biological function of unknown proteins. Technical progress in computational methods offers the potential to make many improvements far faster and more efficient than would be possible by laboratory methods (Zimin *et al.,* 2009). According to Mahmoud *et al.* (2014), chicken *IGF-1* have been seen to serve as a better candidate gene for growth and other metabolic process (proliferation and cellular differentiation) when compared to most species.

The study of *IGF-1* gene of avian through bioinformatics in Nigeria is important to ascertain if the variation and polymorphism among *Gallus gallus, Meleagris gallopavo,* and *Coturnix coturnix* are as a result of convergent or divergent evolution or by chance. Predicting the secondary and tertiary structure of the insulin-like growth factor-1 gene of avians, and also knowing if mutation in *IGF-1* gene that encodes for *IGF-1* protein can lead to changes in the behavior of the protein among the different species. The present study was designed to obtain information on the insulin-like growth factor-1 gene of three avian species using bioinformatics.

Materials and Methods

Retrieval of IGF-1 gene sequences

A total of 15 nucleotide and amino acid sequences of *IGF-1* gene belonging to turkey, chicken, quail, goose, dove and duck were obtained from GenBank, the National Institutes of Health (NIH) genetic sequence database. This was done by obtaining the FAST Alignment (FASTA) format of

nucleotide and amino acid sequences of turkey *IGF-1* gene at the National centre for Biotechnology Information (NCBI, USA) and using Basic local Alignment Search Tool (BLAST) to obtain similar sequences in other organisms. The GenBank accession numbers and sequence lengths of each species were retrieved and tabulated.

Multiple sequence alignment and determination of genetic diversity

Multiple sequence alignment was carried out on all the obtained sequences using Clustal W software (Thompson *et al.*, 1994 incorporated in Molecular Evolution and Genetic Analysis software (MEGA, version 6). The genetic diversity indices such as number of polymorphic sites, number of monomorphic site, haplotype number, haplotype diversity, nucleotide substitution per site, parsimony informative site, singleton variable site and conservation of *IGF-1* gene of turkey, chicken and quail were determined using dnaSP software.

Determination of evolutionary relationship

The percent identity and similarity among the amino acid sequence of *IGF-1* gene of turkey, chicken, and quail were determined by conducting a pairwise comparison of the sequences using *comparing two or more sequences* option of Basic Local Alignment Search Tool (BLAST) incorporated in NCBI website. Phylogenetic relationship among the *IGF-1* gene of turkey, chicken, quail, goose and duck were determined using Molecular Evolution and Genetic Analysis (MEGA6) software.

Results and Discussion

Retrieval of nucleotide and amino acid sequences of IGF-1 gene

The retrieved nucleotide sequences of the *IGF-1* gene of the selected species are shown in Table 1. A total of 15 nucleotide and amino acid sequences of the *IGF-1* gene belonging to the selected species were obtained from GenBank. The number of available sequences and sequence lengths varied for each species. The length of the nucleotide sequence of the *IGF-1* gene varied from 562

Table 1: Retrieved nucleotide and amino acid sequences of the *IGF-1* gene of the selected species with their accession numbers and sequence lengths

Species	GenBank accession number	Nucleotide sequence length	Amino acid sequence length
Chicken	NM_001004384.2	797	153
Turkey	NM_001303149.1	626	153
Quail	AF260131.1	758	153
Goose	DQ662932.1	562	153
Dove	XM_005500280.2	1070	153
Duck	XM_005022553.2	1704	153

base pairs to 1704 base pairs while the length of the amino acid sequence were the same (153 amino acid residues). The gene sequence for chicken, turkey and quail *IGF-1* gene contained 797, 626 and 758 base pairs respectively. The length of the amino acid sequence of chicken, turkey and quail contained 153 amino acids for each of the species. The shortest *IGF-1* nucleotide sequence (562 base pairs) was observed in the goose while the longest *IGF-1* nucleotide sequence (1704 base pairs) was observed in the duck.

Percentage identity of IGF-1 gene among avian species

The percent identity of chicken, turkey and quail *IGF-1* gene and the *IGF-1* gene of other avian species are shown in Table 2. From the result obtained using comparative sequence analysis, it was revealed that the *IGF-1* gene of chicken, turkey and quail shared percent identity ranging from 55 to 99%. It was obtained that chicken and turkey shared 98% identity, while turkey and quail shared 99% identity. This high identity suggest that the protein is highly conserved among the species and might be functionally similar, meaning that the gene can retain its essential amino acids which are responsible for the performance of its biological role from one generation to another (Obetoh *et al.,* 2011). The authors also reported that there are similarities in the *IGF-1* gene of chicken and other avian species. Although their protein is highly conserved, difference in percentage identity and similarity between turkey (98 and 99%), and quail (99 and 100%) respectively, suggest that neutral or synonymous mutation might have occurred in turkey and quail *IGF-1* protein sequence.

Table 2: Percent identity among turkey, chicken and quail *IGF-1* nucleotide sequence and the *IGF-1* nucleotide sequences of other avian species

Species	Chicken	Turkey	Quail
Chicken			
Turkey	98		
Quail	99	99	
Dove	99	98	99
Duck	99	98	98
Goose	99	98	98

This supports neutral mutation hypothesis of evolution that species with different protein variants have equal fitness and most variants are functionally equivalent (Masatoshi *et al.,* 1976).

Percentage similarities of IGF-1 gene on avian species

The result in Table 3, revealed that chicken, turkey and quail values of percentage similarities in function of the *IGF-1* gene falls within the range of 86-99%, meaning that the *IGF-1* protein within which each avian evolved is as a result of divergence. Divergence is the process in which ancestral species or lines accumulate independent genetic changes over time, after the populations have become reproductively isolated. This result is in line with Graur and Li (2000) who reported that

the higher the similarity value in function between two (2) sequences, the lower the probability that they have originated due to convergent evolution, and Obetoh *et al.* (2011) who reported that there are similarities in the *IGF-1* gene of chicken and other avian species. Hence the *IGF-1 gene* of individual avian species are similar in their biological function (increases the absorption of glucose, stimulates myogenesis, inhibits cell cycle genes, increases the synthesis of lipids, and stimulates the production of progesterone in the synthesis of DNA, RNA and protein). The percentage similarities between chicken and turkey, and that between turkey and quail was the same (99%) respectively. This high level of similarities in *IGF-1* gene of avian species is an indicator that the gene is highly effective in regulating a variety of cellular processes (Hegarty *et al.,* 2006).

Phylogenetic study of IGF-1 gene on three avian species

The evolutionary relationship among chicken, turkey and quail *IGF-1* gene and *IGF-1* of other avian are shown in Figure 1. The result showed an optimal tree with the sum of branch length = 1.10475583. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) is shown next to the branches. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree.

Phylogenetic studies can be used to detect gene conservation events because conservation between paralogous genes, which reveal history of gene family, will often cause them to group together rather than with their orthologous genes members in other related species (Drouin *et al.*, 1999; Graur and Li, 2000). It was revealed in this study that the evolutionary relationship of chicken, turkey and quail clustered together with a bootstrap probability of 100%, which is equivalent to 99% confidence level (Xiong, 2006). This implies that the species evolved from a common ancestor. The result confirms morphological classification of chicken, turkey and quail as non-ruminant animals and agrees with the taxonomy of NCBI **(**National Centre for Biotechnology Information). From Figure 1, although chicken, turkey and quail are in the same taxon (any named group of organisms but not necessarily a clade), chicken and quail are classified under the same clade (groups of organisms or genes that include the most recent common ancestor of all of its members and all of the descendants of that most recent common ancestor, it is a monophyletic taxon) (Andreas *et al.*, 2001). According to Kang *et al.* (2008), chicken and turkey share a high similarities value but do not share a most recent common ancestor, the reason being due to evolutionary pressure (amount of change occurring in processes investigated by evolution) and slight changes in the coding sequences of the *IGF-1* gene. A high level of similarity is between the sequences of animals sharing a most recent common ancestor. Evolution tree reveals how two or more sequences are derived, and not how similar they are in function. This result is in line with Obetoh *et al.* (2011), who stated that turkey and chicken share common ancestors when compared to other avian, but in contrast to results from Kang *et al.* (2008) who reported that turkey and chicken shared a most recent common ancestor. The result therefore shows that avian *IGF-1* gene may readily become adapted to similar conditions, and thus assume a close external resemblance

but such resemblances will not reveal but will rather tend to conceal their blood-relationship to their proper lines of descent.

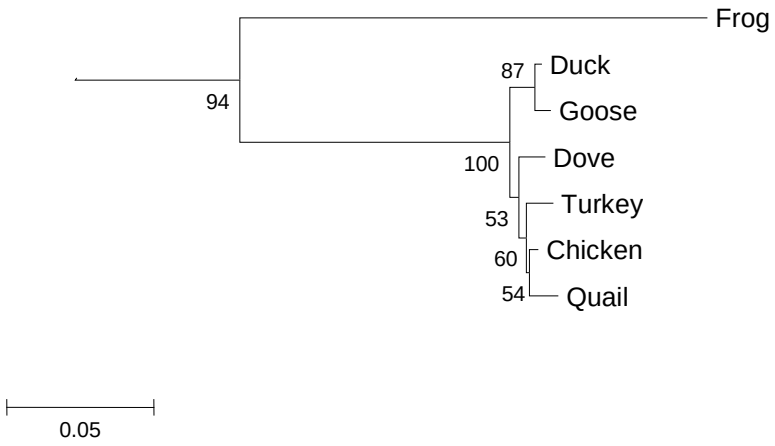

Figure 1: A phylogram showing the evolutionary relationship of avian *IGF1*gene

Genetic diversity study of IGF-1 gene on three avian species

Genetic diversity refers to the variety of genetic information in all the individual organisms in an area. It helps ensures the survival of species because it is what gives rise to the variation between individuals (National DNA Day, 2008). Nucleotide sequence diversity is useful for evolution and natural selection. Genetic diversity in *IGF-1* gene is very useful especially for selecting animals with good growth rate and carcass quality which are preferred by Animal scientists.

From the study of genetic diversity presented in Table 4, it was observed that chicken had the highest haplotype number (4) while quail and turkey had (3 and 2 respectively), of *IGF-1* gene. A haplotype is a specific group of gene or allele that progeny inherits from one parent.

Table 4: Genetic diversity estimation of *IGF-1* gene among chicken, quail and turkey

					Genetic	Diversity	Parameters				
Animals	**N**	**S**	**H**	**H_d**	**II**	**S_c**	**C_t**	**M**	**π_n**	**SP**	**PIP**
Chicken	5	4	4	0.900	0.00476	0.993	1	458	797	2	2
Turkey	2	2	2	1.000	0.00319	0.997	1	0	751	2	0
Quail	2	2	3	1.000	0.00475	0.996	1	0	754	2	0

N: Number of sequences; S: Polymorphic sites; H: Number of haplotype; H_d: Haplotype diversity; π: Nucleotide polymorphism per site; S_c: Sequence conservation; C_t: Conservation threshold; M: monomorphic sites: π_n: Number of nucleotide site; SP: Singleton variable site; PIP: Parsimony informative site.

The result suggests that genetic variation will be more in chicken when compared to other avian studied (Akinfenwa *et al.*, 2011). This variation may allow species to change over time and thereby survive changing environmental conditions. This variation is important because selection is based on variation as different breeders select for different traits of interest. Therefore without variation there will be no selection which will give rise to genetic gain (amount increase in performance that is achieved through artificial genetic improvement). In other words chicken *IGF-1* gene can easily adapt to changing environmental condition, making them fitter than the *IGF-1* gene of turkey and quail.

Turkey, chicken and quail had high haplotype diversity (1, 0.9 and 1.0 respectively), which suggest abundance of genetic diversity in turkey, chicken and quail *IGF-1* gene. Toro and Maki-Tanila (2007), reported that high genetic diversity within a population could arise from overlapping generation and population mixtures, with natural selection favouring heterozygosity or sub-division by genetic drift. The high level of genetic diversity in *IGF-1* gene of turkey, chicken and quail shows that the gene functions properly in promoting growth and regulating other cellular processes which will result to increased meat production and quality, making it a better candidate gene in the subjected avian (Akinfenwa *et al.*, 2011). Also high level of genetic diversity helps ensure survivability and increases variation among species. The result also showed that the highest number of nucleotide sites (797), number of polymorphic sites (4) and pasimony informative sites (2) of *IGF-1* gene. A haplotype is a specific group of gene or allele that progeny inherits from one were estimated in chicken. This suggests that the species has been subjected to natural selection (Borghese *et al.*, 2000).

Conclusion

From the results obtained in this study, it is evident that the high genetic diversity, high percentage similarity and identity in function, and high relative relatedness of *IGF-1* gene in avian reveals that the gene is effective in improving growth and regulating other cellular activities. *IGF-1* can therefore be used as a candidate gene for marker-assisted management strategies in poultry species.

From the results obtained in this study, it is evident that the high genetic diversity, high percentage similarity and identity in function, and high relative relatedness of *IGF-1* gene in avian reveals that the gene is effective in improving growth and regulating other cellular activities. *IGF-1* can therefore be used as a candidate gene for marker-assisted management strategies in poultry species.

References

Akinfenwa, M.O., Peter, S.O., Adenaike, A.S., Obetoh, J.O. and Ikeobi, C.O.N. (2011). Bioinformatics analysis of insulin-like growth factor-1 (IGFI) in five mammalian species. Proceedings of the 16th Annual Conference of the Animal Science Association (ASAN) September 12th, 2011, K.S.U, Anyigba, Nigeria P 19-21

Andreas, D., Baxevanis, B.F. and Francis, O. (2001). *Bioinformatics: A Practical Guide to the Analysis of Genes and Proteins.* Second Edition. Wiley ISBNs: 0-471-38390-2 (Hardback); 0-471-38391-0 (Paper); 0-471-22392-1 (Electronic)

Bertlett, J.C. and Toms, E.G. (2005). Developing a protocol for Bioinformatics analysis: an integrated information behaviour and task analysis approach. *Journal of the American Society for Information Science and Technology* 65: 469-482.

Borghese, A., Moioli, B. and Tripaldi, C. (2000). Buffalo milk: processing and product development in the Mediterranean countries. Third Asian Buffalo Congress, Kandy, Sri Lanka, 27 to 31 March: pp. 37-46.

Bunter, K.L., Hermesch, S., Luxford, B.G., Graser, H.U. and Crump, R.E. (2005). Insulin-like growth factor-1 (IGF-1) measured in juvenile pigs is genetically correlated with economically important performance traits. *Australian Journal of Experimental Agriculture* 45: 783-792.

Drouin, G., Prat, F., Ell, M. and Clarke, G.D.P. (1999). Detecting and characterizing gene conversions between multigene family members. *Molecular Biology and Evolution* 16:1369–1390.

Etherton, T.D. (2004). Somatotropic Function: The somatomedion hypothesis revisited. *Journal of Animal Science* (E. Supply). 8: E239-E214.

Graur, D., and Li, W.H. (2000). *Fundamentals of Molecular Evolution*. 2nd Edition. Sinauer Associates, Sunderland.

Hegarty, R., Macfarlene, S., Banks, R. and Harden, S. (2006) Association of plasma metabolites and hormone with the growth and composition of lambs as affected by nutrition and size genetics. *Austrian Journal of Agricultural Resource* 57: 683-690.

Kang, J.S., Kawakami, Y., Bekku, Y., Ninomiya, Y., Izpisua Belmonte, J.C. and Oohashi, T. (2008). Molecular cloning and developmental expression of a hyaluronan and proteoglycan link protein gen, crtl1/hapln1, in Zebra fish. *Journal of Zoological Science* 25(9): 912-918.

Kemp, S.F. (2007). *Mecasermim rinfabate*. *Drugs Today* (Barc) 43: 149-155.

Mahmoud, A.I., Porrello, E.R., Kimura, W., Olson, E.N. and Sadek, H.A. (2014). Surgical model for cardiac regeneration in neonatal mice. *Nature Protocols* 9:305-311.

Masatoshi, N., Paul, A.F. and Ranajit, C. (1976). Testing the neutral mutation hypothesis by distribution of single locus heterozygosity. *Nature* 262: 491-493.

National DNA Day (2008). *Genetic Variation*. National Human Genome Research Institute, with the National Institutes of Health, in partnership with BSCS and Videodiscovery.

Obetoh, O.J., Akinfenwa, M.O., Peter, S.O., Adenaike, A.S. and Ikeobi, C.O.N. (2011). Comparative analysis and structural modeling of sequence of insulin-like growth factor (*IGF1*) in four poultry species. In: Aduku, A.O., Oluwagbemi, T., Aribido, S.O., Daikwo, S.I and Saliu, O.J. (Eds.) *Research and Value Addition: Key to Transformation of the Nigerian Livestock Industry.*

Proceedings of the 16th Annual conference of the Animal Science Association of Nigeria (ASAN), September 12th-15th, 2011 KSU, Anyigba, Nigeria.

Sakai, K., Busby, W.H. Jr., Clarke, J.B. and Clemmons, D.R. (2001). Tissue transglutaminase facilitates the polymerization of insulin-like growth factor binding protein-1 (IGFBP-1) and leads to loss of IGFBP-1's ability to inhibit insulin-like growth factor-I-stimulated protein synthesis. *Journal of Biological Chemistry* 276: 8740–8745.

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: 4673–4680

Toro, M. and Maki-Tanila, A. (2007). Genomics reveals domestication history and facilitates breed development. In: Oldenbroek, K. (Ed.). *Utilization and Conservation of Farm Animal Genetics Resources*. Wageningen Academic Publishers, Wageningen. pp. 75–102.

Xiong, J. (2006). *Essential Bioinformatics*. Cambridge University Press, New York.

Yilmaz, A., Daus, M.E. and Simmen R.C.M. (2011). Reproductive performance of bulls divergently selected on the basis of blood serum LAF-1 concentration. *Animal Science Journal* 77: 835-839.

Zimin, A.V., Delcher, A.L., Florea, L., Kelly, D.R., Schatz, M.C., Puiu, D., Hanrahan, F., Pertea, G., Van Tazzel, C.P., Sonstegrad, T.S., Marcais, G., Roberts, M., Subramanian, P., Yorke, J.A. and Salzberg, S.L. (2009). A whole-femome assembly of the domestic cow, *Bos Taurus*. *Genomic Biology* 10(4): Article R42.