Nigerian J. Anim. Sci. 2016 (1):11 - 17

Predicting the effects of non-synonymous amino acid variants on protein function in prolactin receptor of cattle and chicken using the MEGA-MD algorithm

Yakubu, A.[1*] and Salako, A.E.[2]

[1]Department of Animal Science, Nasarawa State University, Keffi, Shabu-Lafia Campus, P.M.B., 135, Lafia, 950101, Nigeria.
[2]Animal Breeding and Genetics Laboratory, Department of Animal Science, University of Ibadan, Ibadan, Nigeria.

*Corresponding author's address: abdulmojyak@gmail.com

Target audience: academic researchers, geneticists and molecular biologists.

Abstract

Prolactin receptor (PRLR) is a single transmembrane protein belonging to the cytokine receptor super family through which prolactin plays a wide variety of physiological roles such as mammary gland development, milk production, reproduction and immune function in vertebrates. Many non synonymous single nucleotide polymorphisms (nsSNPs) at the PRLR locus in vertebrates are suspected to impact protein function. This study, therefore, aimed at predicting the likelihood of non synonymous (amino acid change) coding SNPs to cause functional impact on protein at the PRLR locus of cattle and chicken using the MEGA MD bioinformatics tool. In cattle, sixteen out of the first twenty non synonymous amino substitutions obtained: V5A, T9V, T12I, N18S, T19V, C20S, L26S, E32K, F34V, R49E, T52A, S61T, R66K, M72I, I79K and H87Y were beneficial, one was likely neutral, two were deleterious while one was likely deleterious. However, in chicken, L131V, E132N, A134T, V135L, I157A and T161S mutations were found harmless, three were likely neutral, eight were deleterious while three were likely deleterious. This was substantiated by the Evod (–10.70 65.32 versus 0.00-80.03), PolyPhen 2 (0.000-0.859 versus 0.000-0.990) and SIFT (0.16-1.00 versus 0.00-1.00) values in cattle and chicken, respectively. Theoretically, the harmful amino acid substitutions would result in altered spatial structure and functions of the PRLR molecules. Further studies, however, are required to determine whether the beneficial amino acid substitutions obtained will affect the milk yield, reproductive and immune functions of Nigerian livestock species.

Keywords: prolactin receptor; non-synonymous substitutions; bioinformatics tool; protein function; livestock.

Description of problem

Prolactin (**PRL**) is secreted by lactotrophs in the anteriorpituitary gland under dopaminergic control from the hypothalamus and exists in numerous molecular forms, primarily because of alternative posttranslational modifications (1). **PRL** was originally

identified in the late 1920s and named for its ability to stimulate mammary growth and lactogenesis in several species (2). Since then, more than 300 roles for PRL have been identified in a wide range of species including mammals, fish, and birds (3). Among its various target organs, the mammary gland is one of the most sensitive to PRL. More importantly, PRL exerts multiple effects on this organ, ranging from the stimulation of growth to the initiation of milk synthesis and the maintenance of lactation. The actions of PRL are mediated by several PRL receptor (PRLR) isoforms, including its long form and various short PRLR variants that are generated by alternative splicing in a species- and tissue-dependent manner (4). PRLR belongs to the superfamily of class I cytokine receptors, which presumably arose as the result of multiple gene duplications and subsequent divergent evolution (5). The single nucleotide polymorphisms (SNPs) are the most frequent type of genetic variation. In the marker assisted selection (MAS) of livestock, the PRLR seems to be promising candidates. Therefore, SNPs occurring within PRLR gene may regulate important physiological functions such as milk production, egg production, reproduction and immune function or at least be effective DNA markers for this subregion of the livestock genome (6, 7, 8). A great proportion of the known disease-related mutations stems from non-synonymous SNPs, manifested in amino acid mutations (9). Recent advances in high-throughput technologies have generated massive amounts of genome sequence and genotype data for a number of species. The method to identify functional SNPs from a pool, containing both functional and neutral SNPs is challenging by experimental protocols. Therefore, computational predictions have become indispensable for evaluating the impact of non-synonymous single-nucleotide variants discovered in exome sequencing (10). A review of existing computational approaches to estimate the deleteriousness of single nucleotide variants has been recently published by Cooper and Shendure (11). The most common approaches to estimate deleteriousness exploit the fact that sequences observed among living organisms are those that have not been removed by natural selection. Hence, homology searches and conservation analysis are 2 main components of majority of such predictive systems (12). There is dearth of information on the use of *in-silico* methods to detect the presence of beneficial or harmful amino acid substitutions in PRLR gene. Therefore, the present study was undertaken to predict the functional effects of non-synonymous substitutions at the PRLR locus of cattle and chicken using MEGA-MD computational tool.

Materials and Methods

The Gene Search tab of the Mutation Explorer window of The Molecular Evolutionary Genetics Analysis software with mutational diagnosis (MEGA-MD) was used to search for prolactin receptor (PRLR) amino acid variants of cattle and chicken. The amino acid sequence of human prolactin receptor isoform 1 precursor was used as the reference sequence (Peptide ID: NP_000940). MEGA-MD is a suite of

tools developed to forecast the deleteriousness of nsSNVs using multiple methods (EvoD, PolyPhen-2 and bjSIFT) and to explore nsSNVs in the context of the variability permitted in the long-term evolution of the affected position (13). In its graphical interface for use on desktops, it enables interactive computational diagnosis and evolutionary exploration of nsSNVs. As a web service, MEGA-MD is suitable for diagnosing variants on an exome scale. The MEGA-MD suite intends to serve the needs for conducting low- and high-throughput analysis of nsSNVs in diverse applications. MEGA-MD automatically retrieves a 46-species protein sequence alignment that comes from the UCSC resource (14), which has been cached in the MD-DB for quick access. For the selected position, there exists the option to request diagnosis for a specific variant or all possible variants (13).

Results and Discussion

Although there were numerous amino acid substitutions in the prolactin gene, only the first twenty amino acid substitutions for each of cattle and chicken are shown in Tables 1 and 2. Among the non-synonymous substitutions in cattle, sixteen substitutions: V5A, T9V, T12I, N18S, T19V, C20S, L26S, E32K, F34V, R49E, T52A, S61T, R66K, M72I, I79K and H87Y were beneficial, one was likely neutral, two were deleterious while one was likely deleterious. However, in chicken, L131V, E132N, A134T, V135L, I157A and T161S mutations were found harmless, three were likely neutral, eight were deleterious while three were likely deleterious. This was substantiated by the Evod (–10.70-65.32 versus 0.00-80.03), PolyPhen-2 (0.000-0.859 versus 0.000-0.990) and SIFT (0.16-1.00 versus 0.00-1.00) values in cattle and chicken, respectively. The different patterns of amino acid substitutions observed in cattle and chicken reflect adaptive changes (15) Molecular genetic markers are widely used for the characterization of milk production traits, detection of genetically inherited diseases and the determination of the desired breeds; thus they can be utilized to improve livestock production (16, 17). SNP is the most abundant form of genetic variation and a resource for useful genetic traits (18, 19). Identification of sequence variations across the targeted region of candidate gene ultimately leads to genotyping of the animals based on the identified genetic variants that affect the trait of interest (**8**). Therefore, the beneficial amino acid of prolactin receptor may be exploited in improving milk production, reproductive ability and immune functions of Nigerian livestock species through well-structured selection and breeding programmes. The harmful amino acid substitutions obtained in this study may perturb the structure of the α-helix in protein molecules (20), thereby changing the physical and chemical properties of PRLR molecules (19) and overall functions of proteins (21, 22, 23). The obtained results from MEGA-MD may be reliable, since the tool is sensitive to the evolutionary conservation of the positions harbouring the amino acid change which, according to **Gray *et al*. (24),** tends to give better accuracy to predictions.

Table 1. Functional analysis of amino acid mutations in prolactin gene of cattle

Position (AA)	Reference (AA)	Mutant (AA)	Consensus	Evod	PolyPhen-2	SIFT
5	V	A	Neutral	43.29	0.01	0.38
8	A	R	Deleterious	57.41	NA	NA
9	T	V	Neutral	25.55	NA	NA
12	T	I	Neutral	20.99	0.001	1.00
18	N	S	Neutral	42.62	0.009	0.46
19	T	V	Neutral	37.20	NA	NA
20	C	S	Neutral	16.73	0.000	0.88
26	L	S	Neutral	28.18	0.000	1.00
29	G	E	Likely deleterious	55.82	0.859	0.16
32	E	K	Neutral	33.85	0.01	0.98
33	I	L	Likely Neutral	48.07	0.178	1.00
34	F	V	Neutral	51.15	0.008	0.55
40	N	G	Deleterious	65.32	NA	NA
49	R	E	Neutral	38.50	NA	NA
52	T	A	Neutral	42.47	0.000	0.37
61	S	T	Neutral	13.08	0.005	1.00
66	R	K	Neutral	– 12.48	0.001	1.00
72	M	I	Neutral	27.14	0.000	0.60
79	I	K	Neutral	14.50	0.000	1.00
87	H	Y	Neutral	–10.70	0.002	1.00

AA=amino acid, I=isoleucine, L=Leucine, V=valine, F=phenylalanine, M=methionine, C=cysteine, A=alanine, G=glycine, T=threonine, S=serine, Y=tyrosine, N=asparagine, H=histidine, E=glutamic acid, K=lysine, R=arginine

Table 2. Functional analysis of amino acid mutations i n prolactin gene of chicken

Position (AA)	Reference (AA)	Mutant (AA)	Consensus	Evod	PolyPhen-2	SIFT
128	D	G	Deleterious	70.05	0.073	0.07
129	P	S	Deleterious	59.78	0.776	0.04
131	L	V	Neutral	13.70	0.001	1.00
132	E	N	Neutral	0.00	NA	NA
134	A	T	Neutral	–0.44	0.000	1.00
135	V	L	Neutral	–8.00	0.000	1.00
137	V	T	Deleterious	66.25	NA	NA
139	Q	R	Likely deleterious	55.97	0.003	0.12
140	P	S	Likely neutral	55.75	0.003	0.93
141	E	A	Likely deleterious	71.80	0.011	0.15
142	D	N	Likely Neutral	57.15	0.014	0.48
143	R	I	Likely Neutral	69.51	0.001	0.22
149	I	A	Deleterious	58.03	NA	NA
155	T	L	Deleterious	69.33	NA	NA
157	I	A	Neutral	54.17	NA	NA
159	L	A	Deleterious	69.30	NA	NA
160	K	N	Likely deleterious	64.19	0.013	0.01
161	T	S	Neutral	–23.11	0.002	1.00
163	W	S	Deleterious	80.03	0.990	0.00
167	L	H	Deleterious	67.12	NA	NA

AA=amino acid, I=isoleucine, L=Leucine, V=valine, C=cysteine, A=alanine, G=glycine, P=proline, T=threonine, S=serine, W=tryptophan, Q=glutamine, N=asparagine, H=histidine, E=glutamic acid, D=aspartic acid, K=lysine, R=arginine

Conclusions and applications

1. Considering the fact that *PRLR* gene is involved in multiple biological processes in mammals, the beneficial amino acid substitutions obtained in this study may guide subsequent wet and dry laboratory experiments with the ultimate aim of improving genetically the milk yield, reproductive and immune functions of Nigerian livestock species.
2. However, future studies should involve a comparative study using MEGA-D and some other known bioinformatics tools for detecting the deleteriousness of amino acid substitutions in genes of interest.

References

1. Freeman, M. E., Kanyicska, B., Lerant, A. and Nagy, G. (2000). Prolactin: Structure, function, and regulation of secretion. *Physiol. Rev*., 80: 1523–1631.
2. Trott, J. F., Vonderhaar, B.K. and Hovey, R.C. (2008). Historical perspectives of prolactin and growth hormone as mammogens, lactogens and galactagogues—Agog for the future! *J. Mammary Gland Biol. Neoplasia*, 13: 3–11.
3. Bole-Feysot, C., Goffin, V., Edery, M., Binart, N. and Kelly, P.A. (1998). Prolactin (PRL) and its receptor: Actions, signal transduction pathways and phenotypes observed in PRL receptor knockout mice. *Endocr. Rev*., 19:225–268.
4. Trott, J.F., Schennink, A., Petrie, W.K., Manjarin, R., VanKlompenberg, M.K. and Hovey, R.C. (2012). Prolactin: The multifaceted potentiator of mammary growth and function. *J. Anim. Sci*., 90: 1674–1686.
5. Iso-Touru, T., Kantanen, J., Li, M.H., Gizejewski, Z. and Vilkki, J. (2009). Divergent evolution in the cytoplasmic domains of PRLR and GHR genes in Artiodactyla. *BMC Evolutionary Biology*, 9:172 doi:10.1186/1471-2148-9-172.
6. Liang, Y., Cui, J., Yang, G., Leung, F.C. and Zhang, X. (2006). Polymorphisms of 5' flanking region of chicken prolactin gene. Domest. *Anim. Endocrinol*., 30(1): 1-16.
7. Lü, A., Hu, X., Chen, H., Dong, Y. and Pang, Y. (2011). Single nucleotide polymorphisms of the prolactin receptor (PRLR) gene and its association with growth traits in Chinese cattle. *Mol. Biol. Rep*., 38: 261-266.
8. Archana, V. (2013). Nucleotide sequence variability in exon 4 of prolactin locus in Murrah buffaloes (*Bubalus bubalis*). *Research Journal of Biotechnology, 8: 61-65.*
9. Studer, R.A., Dessailly, B.H. and Orengo, C.A. (2013). Residue mutations and their impact on protein structure and function: detecting beneficial and pathogenic changes. *Biochem. J*., 449: 581–594.
10. Liu, L. and Sudhir Kumar, S. (2013). Evolutionary Balancing is Critical for Correctly Forecasting Disease Associated Amino Acid Variants. *Mol. Biol. Evol*., 30 (6):1252-1257.

11. Cooper, G.M. and Shendure, J. (2011). Needles in stacks of needles: finding disease-causal variants in a wealth of genomic data. *Nat. Rev. Genet.,* 12: 628–640.
12. Zemla, D., Kostova, T., Gorchakov, R., Volkova, E., Beasley, D.W.C., Cardosa, J., Weaver, S.C.,Vasilakis, N. and Naraghi-Arani, P. (2014). Genesv – An approach to help characterize possible variations in genomic and protein sequences. *Bioinformatics and Biology Insights,* 8: 1–16.
13. Stecher, G., Liu, L., Sanderford, M., Peterson, D., Tamura, K. and Kumar, S. (2014). MEGA-MD: molecular evolutionary genetics analysis software with mutational diagnosis of amino acid variation. *Bioinformatics*, 30 (9): 1305-1307.
14. Fujita, P.A. et al. (2011). The UCSC Genome Browser database: update 2011. Nucleic Acids Res., 39: 876–882.
15. Li, Y., Wallis, M. and Zhang, Y. (2005). Episodic evolution of prolactin receptor gene in mammals: coevolution with its ligand. *Journal of Molecular Endocrinology,* 35: 411–419.
16. Brym, P., Kamiński, S. and Wójcik, E. (2005). Polymorphism within the bovine prolactin receptor gene (*PRLR*). *Animal Science Papers and Reports,* 23: 61-66.
17. Khaizaran, Z.A. and Al-Razem, F. (2014). Analysis of selected milk traits in Palestinian Holstein-Friesian cattle in relation to genetic polymorphism. *Journal of Animal and Cell Biology*, 8: 74-85.
18. Yakubu, A., Salako, A.E., De Donato, M., Takeet, M.I., Peters, S.O., Adefenwa, M.A., Okpeku, M., Wheto, M., Agaviezor, B.O., Sanni, T.M., Ajayi, O.O., Onasanya, G.O., Ekundayo, O.J., Ilori, B.M., Amusan, S.A. and Imumorin, I.G. (2013a). Genetic diversity in Exon 2 at the major histocompatibility complex DQB1 locus in Nigerian indigenous goats. *Biochemical Genetics*, 51: 954-966.
19. Abula, R., Zhang, H.L., Chen, Y. and Yao, X.K. (2013). Novel polymorphisms detected in the prolactin receptor gene of Yili horse (*Equus caballus*) by PCR-SSCP. *Journal of Animal and Feed Sciences, 22: 70–76.*
20. Visiers, I., Braunheim, B.B. and Weinstein, H. (2000). Prokink: a protocol for numerical evaluation of helix distortions by proline. *Protein Eng*., 13: 603–606.
21. Mathe, E., Olivier, M., Kato, S., Ishioka, C., Hainaut, P. and Tavtigian, S.V. (2006). Computational approaches for predicting the biological effect of p53 missense mutations: a comparison of three sequence analysis based methods. *Nucleic Acids Research*, 34 (5): 1317–1325.
22. Yakubu, A., Musa-Azara, I.S., Yakubu, B.N.S., Daikwo, S.I., Vincent, S.T., Momoh, O.M. and Dim, N.I. (2013b). Biocomputational genome-wide analysis of microRNA genetic variability in some vertebrates. *Genetika*, 45 (3): 799-810.

23. Bhattacharya, T.K., Chatterjee, R.N., Sharma, R.P., Rajkumar, U. and Niranjan, M. (2011). Genetic polymorphism at 5' flanking region of the prolactin gene and its effect on egg quality traits in naked neck chickens. *Journal of Applied Animal Research*, 39: 72-76.

24. Gray, V.E., Kukurba, K.R. and Kumar, S. (2012). Performance of computational tools in evaluating the functional impact of laboratory-induced amino acid mutations. *Bioinformatics*, 28: 2093–2096.