Original Research

Journal of Advanced Veterinary Research
(2025) Volume 15, Issue 3, 312-315

Prolactin gene in duck based on GenBank data sequences: A preliminary study

Latifah Latifah[1], Aswin R. Khairullah[2], Imam Mustofa[3*], Yusup Sopian[4], Yelsi L. Dewi[1], Ulvi F. Handayani[1], Lili Anggraini[1], Adeyinka O. Akintunde[5], Daniah A. Afnani[6], Arif N.M. Ansori[7,8,9]

[1]*Research Center for Animal Husbandry, National Research and Innovation Agency (BRIN), Jl. Raya Bogor Km. 46 Cibinong, Bogor 16911, West Java, Indonesia.*
[2]*Research Center for Veterinary Science, National Research and Innovation Agency (BRIN), Jl. Raya Bogor Km. 46 Cibinong, Bogor 16911, West Java, Indonesia.*
[3]*Division of Veterinary Reproduction, Faculty of Veterinary Medicine, Universitas Airlangga, Jl. Dr. Ir. H. Soekarno, Kampus C Mulyorejo, Surabaya 60115, East Java, Indonesia.*
[4]*Doctoral Program in Innovative Tropical Agriculture, School of Industrial Education and Technology, King Mongkut's Institute of Technology Ladkrabang, 1 Chalong Krung 1 Alley, Lat Krabang, Bangkok 10520, Thailand.*
[5]*Department of Agriculture and Industrial Technology, Babcock University, Ilishan-Remo 121103, Ogun State, Nigeria.*
[6]*Department of Microbiology and Parasitology, Faculty of Veterinary Medicine, Universitas Pendidikan Mandalika, Jl. Pemuda No. 59A, Dasan Agung Baru, Mataram 83125, West Nusa Tenggara, Indonesia.*
[7]*Postgraduate School, Universitas Airlangga, Kampus B Dharmawangsa, Jl. Airlangga 4-6, Surabaya, East Java, 60286, Indonesia.*
[8]*Uttaranchal Institute of Pharmaceutical Sciences, Uttaranchal University, Chakrata Rd, Dehradun, Uttarakhand, 248007, India.*
[9]*Medical Biotechnology Research Group, Virtual Research Center for Bioinformatics and Biotechnology, Jl. Gunung Anyar Kidul No. 33, Surabaya, East Java, 60493, Indonesia.*

ARTICLE INFO

Recieved: 03 April 2025

Accepted: 22 April 2025

*Correspondence:

Corresponding author: Imam Mustofa
E-mail address: imam.mustofa@fkh.unair.ac.id

Keywords:

Duck, Genbank, Genetic diversity, Prolactin gene, SNP.

ABstRACT

The prolactin gene is a candidate gene for egg weight, egg number, and egg production in ducks. This study aimed to perform genetics analysis, including single nucleotide polymorphisms, amino *Aci*d change, restriction enzyme, and phylogenetic tree. A preliminary study on the prolactin gene was carried out in 8 GenBank sequences of ducks. Polymorphism was screened in 8 prolactin gene sequences of duck (AB158611.1; JQ677091.2; GU984377.1; DQ660983.1; DQ345782.1; LC565023.1; NM001310372; DQ345783.1) using alignment with Bioedit ver. 7.2. The results showed that 87 SNPs were detected: 10 SNPs in the 5'UTR region, 3 SNPs in exon 1, 25 SNPs in intron 1, 4 SNPs in exon 2, 18 SNPs in intron 2, 2 SNPs in exon 3, 8 SNPs in intron 3, 1 SNP in exon 4, 9 SNPs in intron 4; 5 SNPs in exon 5 and 2 SNPs in 3'UTR. There were 9 SNPs in the exon region that changed amino *Aci*ds. The main result of restriction enzyme mapping was discovered seven enzymes (*Alu*I, *Bst*KTI, *Dpn*I, *Mbo*I, *Aci*I, *Fat*I, *Nla*III) which recognized 4 SNPs regions based on restriction mapping using Bioedit. The recommendation for a restriction enzyme for the next step was the *Alu*I enzyme. The results of this study provide further evidence of the role of the Prolactin gene as a candidate gene in duck.

Introduction

The prolactin gene is one of the genes that influence duck reproduction. In local ducks of Indonesia, two SNP are associated with egg numbers laid up to three months (Damayanti *et al*., 2022). In another study, the prolactin gene affected egg production, egg weight, double-yolk percentage, egg number, and egg production in ducks (Li *et al*., 2009; Wang *et al*., 2011; Chuekwon and Boonlum, 2017; Purwantini *et al*., 2020; Damayanti *et al*., 2022). The prolactin gene in ducks is found on chromosome 2 with five exons (NCBI, 2024b). Based on GenBank Acc. No. NC_051773.1, the number of prolactin gene sequences is 6165 bp with the location of exon 1 until 5, 53-80, 1584-1765, 2170- 2277, 3582-3761, 5673-5864, respectively (NCBI, 2024a).

Several species' prolactin genes can be analyzed using bioinformatic tools such as alignment, gene search, genome analysis, and gene pattern search (Raza, 2012). BioEdit (offline) and NCBI (online) are often used in DNA sequence analysis. BioEdit is one of the most commonly used programs in molecular biology studies. It originated from sequence alignments specifically for Windows. This program contains many features for easy-to-use sequence alignments, separate window views, user-defined colors, and automatic integration with other programs such as ClustalW and BLAST (Hall *et al*., 2011). Many studies on candidate genes in ducks use bioinformatic tools as a genetic analysis that includes DNA sequence alignment, SNP detection, restriction enzyme analysis, and amino *Aci*d change analysis (Purwantini *et al*., 2020; Sabry *et al*., 2020; Damayanti *et al*., 2022; Jax *et al*., 2022).

DNA sequence alignment and SNP detection can also use the BLAST feature on NCBI. Not only that, NCBI has a database and software (analysis tool) that is often used for genetic analysis, such as DNA-RNA Tool which includes GenBank, BioSystem, Database of Expressed Sequence Tags (dbEST), Database of Genome Survey Sequences (dbGSS), and BLAST (Basic Local Alignment Search Tool) (Jenuth, 2000; Maglott *et al*., 2005; Pruitt *et al*., 2007; Hall *et al*., 2011; Barrett *et al*., 2012). In the previous study, genetic analysis was performed on the MC4R gene of the Bligon goat and other species (Latifah *et al*., 2017). Then, its result used an association study with growth traits and feed intake in Bligon goats (Latifah *et al*., 2018; Latifah *et al*., 2020). The study aimed to identify SNPs, amino *Aci*d changes, restriction enzyme mapping, genetic distance, and phylogenetic trees in the prolactin gene of ducks based on Genbank sequences. This preliminary study will be used for further research, namely the association study of Prolactin gene SNPs and reproductive traits in ducks.

Materials and methods

Data collection

This study used 8 Genbank sequences from NCBI. The Acc. No. of Genbank were AB158611.1, JQ677091.2, GU984377.1, DQ660983.1, DQ345782.1, LC565023.1, NM001310372 and DQ345783.1.

Restriction enzyme mapping has been performed on 5 PCR targets (Chuekwon and Boonlum, 2017).

Data Analysis

The 8 Genbank were used for SNP identification, amino *Aci*d change analysis, and restriction enzyme mapping using Bioedit ver.7.2 software. Seven Genbank (except DQ345783.1) were used to perform genetic distance and phylogenetic tree. The genetic distance and phylogenetic tree contracted using Mega 11 software.

ISSN: 2090-6277/2090-6269/

Results

SNP identification and amino Acid change

Sequences from NCBI were aligned using Bioedit ver. 7.2. Based on alignment analysis, 87 SNPs were detected. Ten SNPs were identified in the 5'UTR region, 3 SNPs in exon 1, 25 SNPs in intron 1, 4 SNPs in exon 2, 18 SNPs in intron 2, 2 SNPs in exon 3, 8 SNPs in intron 3, 1 SNP in exon 4, 9 SNPs in intron 4; 5 SNPs in exon 5 and 2 SNPs in 3'UTR. The 11 SNPs in the coding region amino were changed amino *Acid*s (Table 1).

Restriction enzyme mapping

Based on restriction enzyme mapping using Bioedit, 7 enzymes recognized 3 SNPs consisting 1 enzyme in SNP g.1886G/A (*Alu*I), 3 enzymes in SNPs g.2463C/T (*Bst*KTI, *Dpn*I, and *Mbo*I), 1 enzyme in SNP g.5962G/A (*Aci*I), and 2 enzymes (*Fat*I and *Nla*III) in SNP g.5995A/T (Table 2). One enzyme performed large fragments (>100 bp) with a specific character. Among seven enzymes, there was one recommended enzyme cleavage of the PCR product in the SNP g.1886G/A. For the G base, the *Alu*I enzyme recognizes 2 sites and produces 3 fragments (108, 172, and 121), and for the A base, *Alu*I recognizes 1 site and performs 2 fragments (108 and 293). In this study, the *Alu*I enzyme may be used to genotype ducks using the PCR-RFLP method. The illustration recommendation restriction enzyme is shown in Figure 1.

Genetic distance and phylogenetic tree

The genetic distance among the seven studied prolactin gene sequences revealed that AB158611.1 revealed very close (0.01) with JQ677091.2, (0.07) with GU984377.1, and DQ345782.1, with (0,08) DQ660983.1 and (2.66 and 2.70) with NM001310372.1 and DQ345783.1, respectively (Table 3). The construction of a phylogenetic tree can use the partial gene observed. In this study, a phylogenetic tree was constructed based on the prolactin gene of several sequences of ducks. The phylogenetic tree was joined first on AB158611.1, JQ677091.2, GU984377.1, DQ660983.1, and DQ345782.1, followed by the LC565023.1 and then NM001310372 (Figure 2).

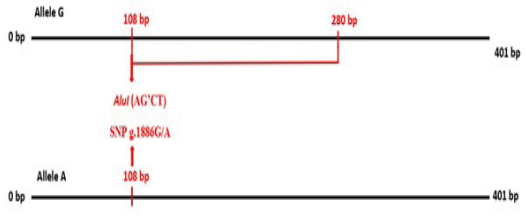

Fig 1. The recommendation of restriction enzyme.

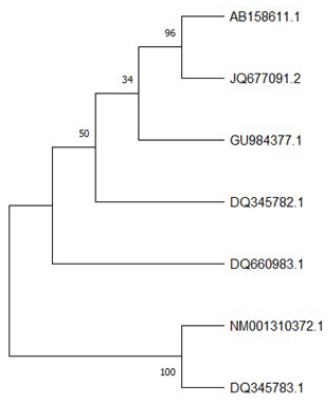

Fig 2. Phylogenetic tree

Table 1. SNP identification and amino *Aci*d change

No	SNP*	Location	Amino *Aci*d change
1	g.86G/A	5'UTR	–
2	g.132AG	5'UTR	–
3	g.135T/C/A	5'UTR	–
4	g.137C/G	5'UTR	–
5	g.178A/G	5'UTR	–
6	g.187A/G	5'UTR	–
7	g.189G/A	5'UTR	–
8	g.213C/T	5'UTR	–
9	g.214A/G	5'UTR	–
10	g.234G/A	5'UTR	–
11	g.259A/C	Exon 1	D/A
12	g.263C/A	Exon 1	S/S
13	g.267G/A	Exon 1	E/K
14	g.295C/G/T	Intron 1	–
15	g.301G/T	Intron 1	–
16	g.314G/T	Intron 1	–
17	g.358T/C	Intron 1	–
18	g.326G/T	Intron 1	–
19	g.343G/A	Intron 1	–
20	g.395C/T	Intron 1	–
21	g.312T/G/C	Intron 1	–
22	g.412A/G	Intron 1	–
23	g.419G/A	Intron 1	–
24	g.434T/C	Intron 1	–
25	g.437A/T	Intron 1	–
26	g.1619T/C	Intron 1	–
27	g.1620A/G	Intron 1	–
28	g.1626G/C	Intron 1	–
29	g.1654C/A	Intron 1	–
30	g.1659G/A	Intron 1	–
31	g.1681T/C	Intron 1	–
32	g.1686A/T	Intron 1	–
33	g.1697A/T	Intron 1	–
34	g.1698C/A	Intron 1	–
35	g.1702T/G	Intron 1	–
36	g.1712C/T	Intron 1	–
37	g.1715C/T	Intron 1	–
38	g.1740T/C	Intron 1	–
39	g.1766C/G	Intron 1	–
40	g.1796A/G	Exon 2	L/L
41	g.1818A/G	Exon 2	K/E
42	g.1886G/A	Exon 2	E/E
43	g.1925T/C	Exon 2	H/H
44	g.2114T/A	Intron 2	–
45	g.2130T/C	Intron 2	–
46	g.2170T/C	Intron 2	–
47	g.2176C/T	Intron 2	–
48	g.2194A/G	Intron 2	–
49	g.2202C/T	Intron 2	–
50	g.2240A/C	Intron 2	–
51	g.2248C/T	Intron 2	–
52	g.2264T/C	Intron 2	–

Note: The SNP numbering is based on Genbank Acc. No. AB158611.1

Table 1 (Continue). SNP identification and amino *Aci*d change

No	SNP*	Location	Amino *Aci*d change
53	g.2268T/G	Intron 2	–
54	g.2279A/G	Intron 2	–
55	g.2281A/G	Intron 2	–
56	g.2296T/A	Intron 2	–
57	g.2299A/G	Intron 2	–
58	g.2311T/G	Intron 2	–
59	g.2312A/G	Intron 2	–
60	g.2318C/T	Intron 2	–
61	g.2336C/T	Intron 2	–
62	g.2463C/T	Exon 3	S/F
63	g.2466T/C	Exon 3	M/F
64	g.2473T/C	Intron 3	–
65	g.2480T/C	Intron 3	–
66	g.2482G/A	Intron 3	–
67	g.2505C/A	Intron 3	–
68	g.2507G/A	Intron 3	–
69	g.3712C/T	Intron 3	–
70	g.3729A/C	Intron 3	–
71	g.3738C/A	Intron 3	–
72	g.3862T/C	Exon 4	S/P
73	g.3974C/A	Intron 4	–
74	g.3983T/A	Intron 4	–
75	g.3987T/C	Intron 4	–
76	g.4000T/C	Intron 4	–
77	g.4009T/C	Intron 4	–
78	g.4042G/C	Intron 4	–
79	g.4050C/A	Intron 4	–
80	g.5719A/C	Intron 4	–
81	g.5751G/A	Intron 4	–
82	g.5861T/C	Exon 5	I/I
83	g.5871A/T	Exon 5	I/F
84	g.5912C/T	Exon 5	A/A
85	g.5962G/A	Exon 5	R/H
86	g.5995A/T	Exon 5	K/M
87	g.6068T/C	3'UTR	–
88	g.6079G/A	3'UTR	–

Note: The SNP numbering is based on Genbank Acc. No. AB158611.1

Table 2. Restriction enzyme map of SNP in coding region.

SNP	Enzyme	recognition	Frequency	Fragment
g.1886G/A	*Alu*I	AG'CT		
	G		2	108, 172, 121
	A		1	108, 293
g.2463C/T	*BstKT*I	G_AT'C		
	C		3	27, 216, 110, 49
	T		2	27, 216, 159
	*Dpn*I	GA'TC		
	C		3	26, 217, 110, 49
	T		2	26, 217, 159
	*Mbo*I	'GATC_		
	C		3	24, 219, 110, 49
	T		2	24, 219, 159
g.5962G/A	*Aci*I	C'CG_C		
	G		2	303, 48, 97
	A		1	303, 97
g.5995A/T	*Fat*I	'CATG_		
	A		2	312, 55, 33
	T		3	287, 25, 55, 33
	*Nla*III	_CATG'		
	A		2	316, 55, 29
	T		3	291, 25, 55, 29

Table 3. Genetic distance of prolactin gene based on GenBank data

	1	2	3	4	5	6	7
AB158611.1							
JQ677091.2	0.01						
GU984377.1	0.07	0.07					
DQ345782.1	0.07	0.07	0				
DQ660983.1	0.08	0.08	0.02	0.02			
NM001310372.1	2.66	2.63	2.55	2.55	2.56		
DQ345783.1	2.7	2.66	2.52	2.52	2.53	0.02	

Discussion

A total of 87 SNPs were detected in ducks based on Genbank sequences in this study using the Bioedit program. According to research by Chuekwon and Boonlum (2017), 5 SNPs were identified and compared with the Genbank sequence Acc. No. AB158611 (55G/A, 191G/T, 255G/A, 287T/G, and 359C/A). An SNP was identified at position c.164G/A and has changed amino guanine (G) to adenine (A), indicating the nucleotide variation in the region (Purwantini *et al*., 2020). SNPs of the prolactin gene were reported to be associated with reproduction and production traits in ducks (Alsudany *et al*., 2023; Astuti, 2019; Susanti and Yuniastuti, 2020). The association study of the prolactin gene with reproduction traits in the Khaki Campbell population described that the SNP 359A/C in intron influences egg production at 300 days of age (Chuekwon and Boonlum, 2017). An SNP was identified at position c.164G/A) with three different genotypes (AA, GA, and GG) that have significant associations with egg production in Tegal and Magelang (F0) ducks (Purwantini *et al*., 2020). In other studies, the prolactin gene SNP SNP 5796C/A and 5817T/C were found to be associated with egg numbers laid up to three months in Alabio and Mojosari Ducks (Damayanti *et al*., 2022). The different results in the association study prolactin gene SNP and production trait were reported by Alsudany *et al*. (2023) that SNP and haplotypes of the prolactin gene exon 1 did not have a significant effect on the average weights of Iraqi ducks. Based on previous research, many of the prolactin genes have the potential to make a candidate gene for reproduction and production traits in ducks.

This study found that 7 enzymes recognize SNP areas. The study of restriction enzyme mapping was done in the MC4R gene in the previous study (Latifah *et al*., 2017; Perdana and Hartatik, 2022). Restriction enzymes are essential tools in in vivo and diverse functions of biological molecules (Loenen *et al*., 2014). Various studies have used restriction enzymes to identify the SNP of the prolactin gene in ducks. Previous research used restriction enzymes for genotyping ducks based on the prolactin gene (Chuekwon and Boonlum, 2017; Sabry *et al*., 2020; Susanti and

Yuniastuti, 2020). The 566 bp PCR product was digested using the DraI enzyme in Central Javanese local ducks (Susanti and Yuniastuti, 2020). The other study used Xba1 and Pst1 enzymes for genotyping prolactin in intron 1 and exon 5 regions in Shanma, Shaoxing, Jingjiang, Youma, and Jingyun ducks (Mazurowski *et al*., 2016).

The genetic distance and phylogenetic tree in this study were contracted using the Mega 11 program. In a previous study, a phylogenetic study explored egyptian duck breeds prolactin gene exon 1 with the other sequences of avian prolactin gene species in the genbank database (Sabry *et al*., 2020).

Conclusion

This study identified 87 SNPs in ducks, emphasizing the significance of the prolactin gene in reproductive and production traits, with various SNPs showing associations with egg production across different duck breeds, while also utilizing restriction enzymes and phylogenetic analysis to enhance understanding of genetic relationships.

Conflict of interest

The authors have no conflict of interest to declare.

References

Alsudany, S.M., Saad, H.F., Zaqeer, B.F., Owaid, J.M., 2023. Polymorphisms of the prolactin gene in Iraqi ducks and its association with body weight. Tex. J. Agric. Biol. Sci. 14, 79–82.

Astuti, D., 2019. The characters and diversity of prolactin and growth hormone genes in the local Turi ducks of Indonesia. AIP Conf. Proc. 2155, 020045. doi: 10.1063/1.5125549.

Barrett, T., Wilhite, S.E., Ledoux, P., Evangelista, C., Kim, I.F., Tomashevsky, M., Marshall, K. A., Phillippy, K.H., Sherman, P.M., Holko, M., 2012. NCBI GEO: archive for functional genomics data sets—update. Nucleic *Acid*s Res. 41, D991–D995. doi: 10.1093/nar/gks1193.

Chuekwon, K., Boonlum, S., 2017. Association of prolactin gene with egg production in Khaki Campbell ducks. Walailak J. Sci. Technol. 14, 849–853.

Damayanti, I., Sudaryati, S., Maharani, D., 2022. The association of prolactin gene polymorphism with egg production traits in Alabio and Mojosari ducks. J. Indones. Trop. Anim. Agric. 47, 184–191. doi: 10.14710/jitaa.47.3.184-191.

Hall, T., Biosciences, I., Carlsbad, C., 2011. BioEdit: an important software for molecular biology. GERF Bull. Biosci, 2, 60–61.

Jax, E., Franchini, P., Sekar, V., Ottenburghs, J., Parera, D.M., Kellenberger, R.T., Magor, K.E., Müller, I., Wikelski, M., Kraus, R.H.S., 2022. Comparative Genomics of the Waterfowl Innate Immune System. Mol. Biol. Evol. 39, msac160. doi: 10.1093/molbev/msac160.

Jenuth, J.P., 2000. The NCBI. In Bioinformatics Methods and Protocols. Springer pp. 301–312.

Latifah, L., Kustantinah, K., Maharani, D., Hartatik, T., 2020. Identification of polymorphism in MC4R gene and its association with dry matter and crude protein intake in post-weaned Bligon goats. J. Indones. Trop. Anim. Agric. 45, 173–180. doi: 10.14710/jitaa.45.3.173-180.

Latifah, L., Maharani, D., Kustantinah, A., Hartatik, T., 2018. Association of melanocortin 4 receptor gene polymorphism with growth traits in Bligon goat. J. Indones. Trop. Anim. Agric. 43, 343–351. doi: 10.14710/jitaa.43.4.343-351.

Latifah, Priyadi, D.A., Maharani, D., Kustantinah, Hartatik, T., 2017. Genetic analysis using partial sequencing of melanocortin 4 receptor (MC4R) gene in Bligon goat. J. Anim. Sci. Technol. 40, 71–77. doi: 10.5398/medpet.2017.40.2.71.

Loenen, W.A., Dryden, D.T., Raleigh, E.A., Wilson, G.G., Murray, N.E., 2014. Highlights of the DNA cutters: a short history of the restriction enzymes. Nucleic *Acid*s Res. 42, 3–19. doi: 10.1093/nar/gkt990.

Li, H., Zhu, W., Chen, K. W., Zhang, T., Song, W., 2009. Association of polymorphisms in the intron 1 of duck prolactin with egg performance. Turk. J. Vet. Anim. Sci. 33, 193–197. doi: 10.3906/vet-0709-4.

Maglott, D., Ostell, J., Pruitt, K.D., Tatusova, T., 2005. Entrez Gene: gene-centered information at NCBI. Nucleic *Acid*s Res. 33, D54–D58. doi: 10.1093/nar/gkq1237.

Mazurowski, A., Frieske, A., Wilkanowska, A., Kokoszyński, D., Mroczkowski, S., Bernacki, Z., Maiorano, G., 2016. Polymorphism of prolactin gene and its association with growth and some biometrical traits in ducks. Ital. J. Anim. Sci. 15, 200–206. doi: 10.1080/1828051X.2016.1153405.

NCBI (National Center for Biotechnology Information), 2024a. Anas platyrhynchos isolate Z2 breed Pekin duck chromosome 2, ZJU1.0, whole genome shotgun sequence. NCBI. https://www.ncbi.nlm.nih.gov/nuccore.

NCBI (National Center for Biotechnology Information), 2024b. PRL prolactin [Anas platyrhynchos (mallard)]. NCBI. https://www.ncbi.nlm.nih.gov/gene/101793643.

Perdana, Y.C., Hartatik, T., 2022. Restriction Mapping of MC4R Gene on Bali Cattle (Bos sondaicus) as Genetic Marker for Breeding Program in Compared to Bos taurus and Bos indicus. J. Trop. Biodivers. Biotechnol. 7, jtbb67636. doi: 10.22146/jtbb.67636.

Pruitt, K.D., Tatusova, T., Maglott, D.R., 2007. NCBI reference sequences (RefSeq): a curated non-redundant sequence database of genomes, transcripts and proteins. Nucleic *Acid*s Res. 35, D61–D65. doi: 10.1093/nar/gkl842.

Purwantini, D., Santosa, R.S.S., Santosa, S.A., Susanto, A., Candrasari, D.P., Ismoyowati, I., 2020. Prolactin gene polymorphisms and associations with reproductive traits in Indonesian local ducks. Vet. World 13, 2301–2311. doi: 10.14202/vetworld.2020.2301-2311.

Raza, K., 2012. Application of data mining in bioinformatics. ArXiv Preprint ArXiv:1205.1125.

Sabry, N.M., Mabrouk, D.M., Abdelhafez, M.A., El-Komy, E.M., Mahrous, K.F., 2020. Polymorphism of the Prolactin Gene in Egyptian Duck Breeds. J. World's Poult. Res. 10, 587–598. doi: 10.36380/jwpr.2020.67.

Susanti, R., Yuniastuti, A., 2020. Prolactin and X-collagen genes polymorphism in Central Javanese local ducks, Indonesia. Biodiversitas 21, 605–610. doi: 10.13057/biodiv/d210223.

Wang, C., Liang, Z., Yu, W., Feng, Y., Peng, X., Gong, Y., Li, S., 2011. Polymorphism of the prolactin gene and its association with egg production traits in native Chinese ducks. S. Afr. J. Anim. Sci. 41, 63–69. doi: 10.4314/sajas.v41i1.66044.