HAYATI
Journal of Biosciences

Vol. 31 No. 3, May 2024 548-558
DOI:10.4308/hjb.31.3.548-558
ISSN: 1978-3019
EISSN: 2086-4094

Genetic Structure of *Gallus varius* Based on Middle-Lower Section of Control Region mtDNA

Achmad Alfiyan[1], Achmad Farajallah[1*], Maria Ulfah[2], Dyah Perwitasari-Farajallah[1,3], Muladno Muladno[2]

[1]Department of Biology, Faculty of Mathematics and Natural Sciences, IPB University, Kampus IPB Dramaga, Bogor 16680, Indonesia
[2]Department of Animal Production and Technology, Faculty of Animal Science, IPB University, Kampus IPB Dramaga, Bogor 16680, Indonesia
[3]Primate Research Center, IPB University, Bogor 16151, Indonesia

ARTICLE INFO

Article history:
Received October 31, 2023
Received in revised form January 3, 2024
Accepted January 11, 2024

KEYWORDS:
Control region,
Gallus varius,
haplotype diversity,
middle-lower part

ABSTRACT

***Gallus varius*, an indigenous bird species of Indonesia, demonstrates significant genetic diversity. The genetic diversity assessment in *G. varius* encompassed both mitochondrial DNA (mtDNA) and nuclear DNA, with the mtDNA analysis primarily centered on domain I of the control region. This study aimed to describe the genetic diversity and structure of *G. varius* inhabiting Java, Madura, Lombok and Sumbawa based on the middle-lower part of control region mtDNA. Genomic DNA was extracted from the calamus tip of feather, then the middle-lower part of control region was amplified and sequenced using two pairs of primers. In the examined control region, spanning from 944 to 1,008 bp, 13 bp of nucleotide variation was observed, with nucleotide diversity at 0.0021. Between *G. varius* samples and the reference (NC_007238.1), a total of 7 haplotypes were identified, 5 unique and 2 shared haplotypes, with haplotype diversity at 0.7692. The substantial diversity of haplotypes in this study and two previous study suggests that the genetic diversity of *G. varius* has remained stable over the past two decades. Additionally, genetic distance data indicate there is multiple *G. varius* subspecies, and the haplotype network accentuates signs of population differentiation.**

1. Introduction

Gallus varius Shaw 1798, commonly known as the green junglefowl, is a terrestrial avian species, belonging to the Galliformes order, Phasianidae family, and *Gallus* genus, alongside three congeneric species. This particular junglefowl was endemic to Indonesia, with its distribution areas encompassing Java, Madura, Bawean, Kangean, Bali, and the Lesser Sunda Islands (Madge and McGowan 2010; BirdLife International 2018; Clements *et al.* 2019). Within the taxonomy of *Gallus*, *G. varius* presents distinct morphological features that distinguish it from its congeneric species. Among these distinct features is the iridescent green plumage, prominently displayed in the nape, neck hackles, and mantle areas. Its comb is singular in type, featuring a straight-edged configuration and a red coloration with a central bluish base. Furthermore, *G. varius* presents a single, sizeable throat lappet, characterized by a dominant red hue, with adjacent yellow margins around the neck region and at its distal end displaying bluish colouring (Madge and McGowan 2010).

Over the course of the past two decades, *G. varius* has been recognized for its notable genetic diversity. A noteworthy investigation by Zein and Sulandari in 2008, focused on the domain I of the mitochondrial control region, involved the examination of 33 *G. varius* specimens originating from four different locations. This investigation revealed the formation of 25 distinct haplotypes, with each location displaying a haplotype diversity exceeding 0.8. Within the mitochondrial genome, the control region serves as the primary non-coding region. In *Gallus*, this particular region is flanked by the coding sequences of glutamic acid tRNA (tRNA-Glu) and phenylalanine tRNA (tRNA-Phe) (Nishibori *et al.* 2005; Miao *et al.* 2013). Based on the differential of nucleotide frequency and the distribution of nucleotide variation, control region is divided into 3 domains, which is Domain I (near 5' end), Domain II

* Corresponding Author
E-mail Address: achamad@apps.ipb.ac.id

(central) dan Domain III (near 3'terminus) (Saccone *et al.* 1991; Marshall and Baker 1997). Domain I exhibit the highest frequency of CTP nucleotides compared to the other domains, while Domains II and III showcase the highest frequencies of GTP and ATP nucleotides, respectively, compared to the other domains. Additionally, nucleotide variation is notably concentrated in Domains I and III (Saccone *et al.* 1991; Marshall and Baker 1997; Huang and Ke 2014). In 2016, an investigation into whole genome single nucleotide polymorphisms (SNPs) of *G. varius* from Java and Madura revealed that *G. varius* samples from the same location exhibited genetic distance values of ≥0.32. Importantly, this genetic distance value significantly higher than the value observed in red junglefowl and domestic chicken samples, which was <0.17 (Ulfah *et al.* 2016). The higher value of genetic distance in *G. varius* from Java and Madura suggests a greater level of genetic variation within *G. varius* surpassing the genetic variation observed between samples of red junglefowl and domestic chicken. Ulfah *et al.* (2016) further highlighted that, based on complete mitochondrial genome sequence, the phylogenetic tree formed within the *G. varius* clade displayed extended branch lengths for each *G. varius* sample, indicative of high genetic variation among these specimens.

The utilization of whole-genome SNPs and whole-genome mitochondria as molecular markers in genetic diversity assessments could provide a robust level of comprehensiveness, accuracy, and detail, due to the vast amount of data they offer. Using partial gene or region as genetic marker, particularly one with a high degree of polymorphism such as control region of mtDNA, in evaluating the genetic diversity in a species could simplify the data handling and analysis, thus making it time-efficient approach. Nonetheless, the downside lies in the limited information provided by partial genes or regions, which curtails the overall comprehensiveness, accuracy, and level of detail in understanding genetic diversity.

The endemicity of *G. varius* is remarkable due to its distribution across multiple regions (islands). This species further reveals diverse genetic diversity levels across its distribution areas, accompanied by unique genetic features exclusive to specific regions (Zein and Sulandari 2008). At present, within each of its distribution region, this species faces the pressing issue of genetic diversity loss, attributed to over-exploitation, habitat degradation, fragmentation, and other contributory factors. Therefore, there is a pressing need for a thorough examination of the genetic diversity of *G. varius*. This research holds significant importance, as the genetic insights gathered will serve as a fundamental framework for the development of future conservation strategies. Accordingly, the objective of this study was to describe the genetic diversity and structure of *G. varius* inhabiting Java, Madura, Lombok, and Sumbawa based on the middle-lower section of control region. Within this investigation, the concentrated analysis is directed at the middle-lower section of the mitochondrial control region, which includes both domain II and domain III. Of particular importance, domain III represents as one of central region of nucleotide variation in the control region, and the information regarding this area still limited, particularly for *G. varius*. The primary motive behind this approach is to offer supplementary insights into the extent of polymorphism within the mtDNA control region of *G. varius.*

2. Materials and Methods

2.1. Specimen Collection

The *G. varius* samples were collected from breeders (2), a pet seller in the local market (1), and a hunter (1). Preceding the sample collection process, a brief interview was carried out to ascertain the precise geographical origin of each specimen. Chest feather were collected from 12 samples of *G. varius* originating from four locations, namely Madura (1), Java (1), Sumbawa (2), and Lombok (8) (Table 1). Chest feather collection was done based on protocol DNA Diagnostic Center (DDC) (c2019). Every individual's feather was placed within a clean paper container and appropriately labelled (DDC c2019) then stored in a -20°C freezer when in the laboratory. In order to enhance the robustness and representation of our dataset, we included genetic data from 9 *G. varius* individuals in our analysis, encompassing 7 samples from Madura and 2 from Java (Table 1), collection of Dr. Maria Ulfah (Ulfah *et al.* 2016). These genetic records were obtained from GenBank's SRA database.

2.2. Genomic DNA Extraction

The genomic DNA was extracted from the base of the feather, specifically the tip of the calamus section, as it contained tissue from beneath the skin

Table 1. Control region mtDNA sequence list of *Gallus* genera that was used in this study

Species	Origin of location	N	DNA sequence abbreviation	Source	Accession number	Reference
ingroup						
G. varius	Indonesia (Lombok)	8	Gv_L2-6	Pet seller	-	*
			Gv_L7, 8, 13	Breeder	-	
	Indonesia (Sumbawa)	2	Gv_Sw1	Hunter	-	*
			Gv_Sw2	Breeder	-	
	Indonesia (Madura)	8	Gv_Md1	Breeder	-	*
			Gv_Md2-8	GenBank	DRX083687-93	d
	Indonesia (Gunung Ijen/Java)	1	Gv_Ij1	Breeder	-	*
	Indonesia (Java)	2	Gv_Jv1-2	Genbank	DRX083685-86	d
	Indonesia (Bali)	1	Gv_Bli	Genbank	NC_007238.1	a
outgroup						
G. gallus gallus	-	1	Gg_gal	Genbank	X52392.1	b
G. gallus gallus	Philippines	1	Gg_gal_Fil	Genbank	AP003322.1	a
G. gallus bankiva	Indonesia (Bali)	1	Gg_ban_Bli	Genbank	AP003323.1	a
G. gallus spadiceus	Myanmar	1	Gg_spa_Myn	Genbank	NC_040902.1	c
G. gallus jabouillei	China	1	Gg_jab_Chn	Genbank	GU261696.1	c
G. gallus murghi	India	1	Gg_mur_Ind	Genbank	GU261707.1	c
G. lafayetii	Japan	1	Glaf_Jpn	Genbank	NC_007239.1	a
G. sonneratii	India	1	Gson_Ind	Genbank	NC_007240.1	a

-: data unavailable, *: this study, a: Nishibori *et al.* 2005, b: Desjardins and Morais 1990, c: Miao *et al.* 2013, d: Ulfah *et al.* 2016, N: Number of DNA sequence

that contained DNA (DDC c2019). Using a sterile surgical scissor, the tip of the calamus was removed and then placed within a 1.5 ml tube. Following this step, the addition of GT Buffer solution, one of the buffer components of the DNA extraction kit, was carried out, and then the calamus tip was fragmented into smaller portions using the same sterilized surgical scissors. The subsequent procedure was conducted in accordance with the protocol outlined in the GENEAID Genomic DNA extraction kit (Tissue) (Geneaid, Canada).

2.3. DNA Amplification and Sequencing

The amplification of the middle-lower segment of the mtDNA control region from *G. varius* DNA samples was carried out using the Polymerase Chain Reaction (PCR) method, utilizing the Biometra Thermo Cycle tool. The amplification process involved the use of two primer pairs, which were designed based on the *G. varius* sequence (NC_007238.1), namely the primer pair AF623 (forward) 5'-TCG TCC TAC CCC ATA TCC AAC-3' and AF624 (reverse) 5'-CAA GAT GAT TCC CCA TAC ACG-3' and the primer pair AF625 (forward) 5'-TTT TCA CGA AGT CAT CTG TGG-3' and AF626 (reverse) 5'-GAT ACC TGC TCC TTC TGC TG-3' (Figure 1). The amplification process utilized the GoTaq® Green Mastermix reagent. PCR conditions involved an initial pre-denaturation step at 94°C for 3 minutes, followed by three sequential stages: denaturation at 94°C for 1 minute, annealing at 60°C for 1 minute, and elongation at 72°C for 1 minute, which were repeated through 35 cycles. Following this, a post-elongation phase was performed at 72°C for 2 minutes, concluding with a finalization step at 15°C for 10 minutes. The amplicon obtained was subsequently underwent electrophoresis on a 1% agarose gel which had been added with Florosafe staining dye (1st Base, MY) and electrophoresed at 80 V for 55 minutes. Following this step, the amplicon within the agarose gel was visualized using UV light on the Geldock instrument. Amplicons displaying high-quality DNA bands were sent to a sequencing service company, 1st Base, for sequencing process. Sequencing was carried out in accordance with the Sanger method, utilizing 2 pairs of primers (AF623-624 and AF625-626), Applied Biosystem Big Dye® terminator kit V reagents 3.1, and ABI7700 sequencer tools.

2.4. Assembling mtDNA Control Region (Middle-Lower) of *G. varius* from SRA Database

The mtDNA control region sequences of 9 *G. varius* specimens, part of Dr. Maria Ulfah's collection, were mapped using BLAST-n (Altschul *et al.* 1990). The reference sequence chosen for this process was the mtDNA control region sequence of *G. varius* from Bali

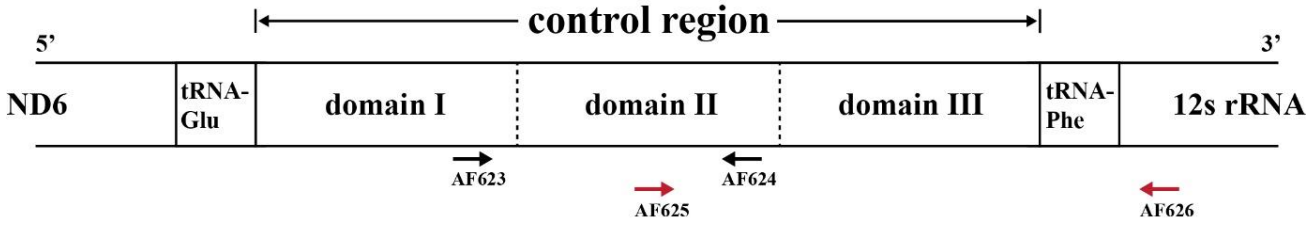

Figure 1. Schematic position of primers that was used in amplification process of middle-lower section of mtDNA control region, the figure was contructed based on genes organization of *G. varius* (NC_007238.1) (Nishibori *et al.* 2005)

(NC_007238.1). Subsequently, the assembly of the fragmented mtDNA control region sequences, which had been mapped, was conducted using MEGA7 (Kumar *et al.* 2016).

2.5. Molecular Analysis

Sequencing products, in the form of chromatograms generated by two primer pairs, were edited and combined using MEGA7. Following this, the combined sequences were compared to the sequences data stored in GenBank using BLAST-N. Upon the completion of a comparative assessment between the sample data with the database, the control region data of *G. varius* was aligned with ingroup and outgroup sequences (Table 1) through the utilization of the ClustalW tool incorporated within the MEGA7. Subsequently, a series of analyses were initiated, including assessments of nucleotide composition, nucleotide variation, genetic distance measurement, allelic differentiation, phylogenetic tree construction, haplotype construction and construction of haplotype network.

The nucleotide composition assessment, genetic distance measurement and the construction of a phylogenetic tree were carried out within the MEGA7. The genetic distance was calculated using the Kimura 2 Parameter method. The establishment of the phylogeny tree was carried out employing the Maximum Likelihood statistical method, with the incorporation of the Hasegawa-Kishino-Yano substitution model and the addition of a Gamma-distribution featuring 5 categories to address the variable evolutionary rates observed across nucleotide sites. The decision to adopt this construction method was guided by the outcomes of the "Find Best DNA/Protein Models (ML)" analysis conducted within MEGA7, where it demonstrated the lowest Bayesian Information Criterion (BIC) and Akaike Information Criterion (AIC) values. Furthermore, the analysis was enhanced by the incorporation of the bootstrap method, generating 1,000 replications to assess phylogeny.

For the analysis of nucleotide variation and the construction of haplotypes within the ingroup (*G. varius*), we utilized the DnaSP6 software (Rozas *et al.* 2017). Allelic differentiation among different *G. varius* locations was assessed via pairwise Fst values, and their statistical significance was determined using the Arlequin 3.5.2 software (Excoffier and Lischer 2010). The establishment of the *G. varius* haplotype network was carried out using Network ver. 10.2 software (Fluxus Technology Ltd c2021) and the median joining calculation method (Bandelt *et al.* 1999).

3. Results

3.1. Comparison and Composition of Control Region DNA Sequence of Directly Collected *G. varius* Samples

A total of 12 control region sequences, with lengths spanning from 944 to 1,008 bp, were effectively acquired from 12 *G. varius* samples which was collected directly. During the sequence database comparison with BLAST-N, these sequences showed a high degree of similarity level with the mitochondrial control region sequence of *G. varius* in the database (Table 2). These results were confirmed by the substantial values of the Identity and Query Cover parameters, as described in Table 2. The nucleotide composition of the 12 control region sequences obtained from *G. varius* samples exhibited a coherent distribution pattern of nucleotides, akin to the control region sequence of *G. varius* as observed in Bali (NC_007238.1) within the database. The mean nucleotide distribution in the 12 control region sequences of *G. varius* samples exhibits a composition of 33.87% TTP, 26.38% CTP, 25.74% ATP, and 14.01% GTP. In parallel, the nucleotide makeup of *G. varius* from Bali (NC_007238.1) comprises 33.70% TTP, 26.38% CTP, 25.90% ATP, and 14.00% GTP. The

Table 2. Comparison analysis result of 12 control region sequence of directly collected *G. varius* samples using BLAST-N

Species	List of sequence accession number	Value range				
		Max score	Total score	Query cover (%)	E-value	Identity (%)
G. varius	AP003324.1	2052-2143	2052-2192	99-100	0	99.15-99.49
G. gallus	GU261676.1					
	KY039423.1	1840-1875	1840-1995	99-100	0	95.08-96.04
	KY039419.1					

Table 3. List of haplotypes, nucleotide variation position and its mutation type of *G. varius* based on middle-lower section of mtDNA control region

Haplotypes	Nucleotide position												
	26	44	74	148	277	381	486	530	531	532	533	535	674
H1	A	C	C	A	G	C	-	T	A	T	C	A	C
H2	G	.	.	.	T	.	-	A	T	C	A	T	.
H3	G	.	T	.	T	.	C	A	T	C	A	T	.
H4	G	.	T	.	T	.	-	A	T	C	A	T	.
H5	G	.	.	.	T	.	C	A	T	C	A	T	A
H6	G	.	T	G	T	T	-	A	T	C	A	T	.
H7	G	A	.	.	T	T	-	A	T	C	A	T	.
Mutation type	Ts	Tv	Ts	Ts	Tv	Ts	indel	Tv	Tv	Ts	Tv	Tv	Tv

.: same nucleotide type, - : gap, Indel: insertion-deletion, Ts: transition, Tv: transversion

findings from BLAST-N analysis and examination of nucleotide distribution pattern collectively validate the mtDNA control region identity of our sequence.

3.2. Nucleotide Variation and Haplotypes of *G. varius*

Within the middle-lower section of the control region sequence of *G. varius* (ingroup), a total of 13 nucleotide variations were identified. The nucleotide diversity value for this segment was determined to be 0.00210. Among these variations, one nucleotide alteration resulted from an insertion-deletion mutation, while the remaining 12 base pairs underwent substitution mutations, including 4 transitions and 8 transversions (Table 3). The determined haplotype diversity value underscores that the likelihood of encountering distinct alleles within the middle-lower segment of the mtDNA control region between two randomly sampled *G. varius* specimens from a specified location is greater than 76%. Notably, 5 out of the 7 haplotypes were exclusive to specific *G. varius* locations, while the remaining two were shared among two or three distinct *G. varius* locales (Table 4). The observed counts of nucleotide variations, haplotypes, nucleotide diversity, and haplotype diversity values in this study were comparatively lower than those reported in a prior investigation focused on the hypervariable region 1 (HVR1) or domain I of the mtDNA control region, as conducted by Zein and Sulandari in 2008. Nevertheless, when considering the value of haplotype diversity, it becomes evident that the genetic diversity within *G. varius*, as elucidated through the middle-lower section of the mtDNA control region, remains quite substantial.

Table 4. Haplotype composition of *G. varius* based on middle-lower section of mtDNA control region

Haplotypes	N	Sequence composition
H1	1	Gv_Bli_NC_007238.1
H2	9	Gv_Ij1, Gv_Md1, Gv_Lb13, Gv_Md7_DRX083692, Gv_Md4_DRX083689, Gv_Md8_DRX083693, Gv_Md6_DRX083691, Gv_Jv2_DRX083686, Gv_Jv1_DRX083685
H3	6	Gv_Lb2, Gv_Lb3, Gv_Lb4, Gv_Lb6, Gv_Sw1, Gv_Sw2
H4	1	Gv_Lb5
H5	1	Gv_Lb7
H6	1	Gv_Lb8
H7	3	Gv_Md5_DRX083690, Gv_Md3_DRX083688, Gv_Md2_DRX083687

3.3. Genetic Distance and Allelic Differentiation

The genetic distance range between the reference *G. varius* (*G. varius* from Bali) and the study's sampled specimens exhibited a significant variance compared to the genetic distances observed among the *G. varius* samples, spanning from 0.008 to 0.011 and 0.000 to 0.005, respectively (Table 5). Moreover, the mean genetic distance among *G. varius* within this investigation (0.002) surpassed that of the

Table 5. Genetic distance between *G. varius* species based on middle-lower section of mtDNA control region

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
1																						
2	0.008																					
3	0.009	0.001																				
4	0.009	0.001	0.000																			
5	0.009	0.001	0.000	0.000																		
6	0.009	0.001	0.000	0.000	0.000																	
7	0.009	0.001	0.000	0.000	0.000	0.000																
8	0.009	0.001	0.002	0.002	0.002	0.002	0.002															
9	0.011	0.002	0.001	0.001	0.001	0.001	0.001	0.004														
10	0.008	0.000	0.001	0.001	0.001	0.001	0.001	0.001	0.002													
11	0.008	0.000	0.001	0.001	0.001	0.001	0.001	0.001	0.002	0.000												
12	0.009	0.001	0.000	0.000	0.000	0.000	0.000	0.002	0.001	0.001	0.001											
13	0.009	0.001	0.000	0.000	0.000	0.000	0.000	0.002	0.001	0.001	0.001	0.000										
14	0.008	0.000	0.001	0.001	0.001	0.001	0.001	0.001	0.002	0.000	0.000	0.001	0.001									
15	0.011	0.002	0.004	0.004	0.004	0.004	0.004	0.004	0.005	0.002	0.002	0.004	0.004	0.002								
16	0.008	0.000	0.001	0.001	0.001	0.001	0.001	0.001	0.002	0.000	0.000	0.001	0.001	0.000	0.002							
17	0.011	0.002	0.004	0.004	0.004	0.004	0.004	0.004	0.005	0.002	0.002	0.004	0.004	0.002	0.000	0.002						
18	0.011	0.002	0.004	0.004	0.004	0.004	0.004	0.004	0.005	0.002	0.002	0.004	0.004	0.002	0.000	0.002	0.000					
19	0.008	0.000	0.001	0.001	0.001	0.001	0.001	0.001	0.002	0.000	0.000	0.001	0.001	0.000	0.002	0.000	0.002	0.002				
20	0.008	0.000	0.001	0.001	0.001	0.001	0.001	0.001	0.002	0.000	0.000	0.001	0.001	0.000	0.002	0.000	0.002	0.002	0.000			
21	0.008	0.000	0.001	0.001	0.001	0.001	0.001	0.001	0.002	0.000	0.000	0.001	0.001	0.000	0.002	0.000	0.002	0.002	0.000	0.000		
22	0.008	0.000	0.001	0.001	0.001	0.001	0.001	0.001	0.002	0.000	0.000	0.001	0.001	0.000	0.002	0.000	0.002	0.002	0.000	0.000	0.000	

1: Gv_Bli_NC_007238.1; 2: Gv_Jj1; 3: Gv_Lb2; 4: Gv_Lb3; 5: Gv_Lb4; 6: Gv_Lb5; 7: Gv_Lb6; 8: Gv_Lb7; 9: Gv_Lb8; 10: Gv_Md1; 11: Gv_Lb13; 12: Gv_Sw1; 13: Gv_Sw2; 14: Gv_Md7_DRX083692; 15: Gv_Md5_DRX083690; 16: Gv_Md4_DRX083689; 17: Gv_Md3_DRX083688; 18: Gv_Md2_DRX083687; 19: Gv_Md8_DRX083693; 20: Gv_Md6_DRX083691; 21: Gv_Jv2_DRX083686; 22: Gv_Jv1_DRX083685

G. gallus dataset, encompassing all documented subspecies (0.001). It is worth noting that, to date, *G. varius* lacks any formally recognized subspecies. These findings suggest the potential existence of subspecies within *G. varius*. Moderate evidence of allelic differentiation among *G. varius* populations from distinct locations is apparent, as evidenced by the global pairwise Fst value (global Fst: 0.52896, P value = 0.00000±0.00000) and the pairwise Fst values between locations (Table 6). These findings suggest the occurrence of a certain magnitude of genetic exchange among the studied locations.

3.4. Phylogenetic Tree and Haplotype Network

The phylogenetic tree's structure revealed that, based on the middle-lower segment of the mtDNA control region, all *G. varius* specimens formed a single clade, exhibiting a closer genetic affinity with *G. sonneratii* when compared to other *Gallus* species (Figure 2). Furthermore, the haplotype network analysis displayed the separation of *G. varius* from Bali with *G. varius* from other locations (Figure 3). This observation suggests the presence of population differentiation within *G. varius*.

4. Discussion

This study has unambiguously identified the DNA sequence in question as the control region of mitochondrial DNA from *G. varius*. Notably, this study has revealed a limited number of nucleotide variations, amounting to only 13 bp, in stark contrast to the earlier investigation by Zein and Sulandari in 2008, which reported 28 bp of variation within domain I of the mitochondrial DNA control region. The variances observed can be attributed to differences in mutation rates specific to the studied domains. Numerous factors influence the magnitude of mutation rates, including the

Table 6. Fixation index between different location of *G. variu*

	1	2	3	4	5
1					
2	0.38844*				
3	0.32681*	0.09774			
4	-0.14886	1.00000	0.46706*		
5	0.83023	1.00000	0.80220	1.00000	

*: significant value with p = <0.05. 1: *G. varius* from Lombok. 2: *G. varius* from Java. 3: *G. varius* from Madura. 4. *G. varius* from Sumbawa. 5. *G. varius* from Bali

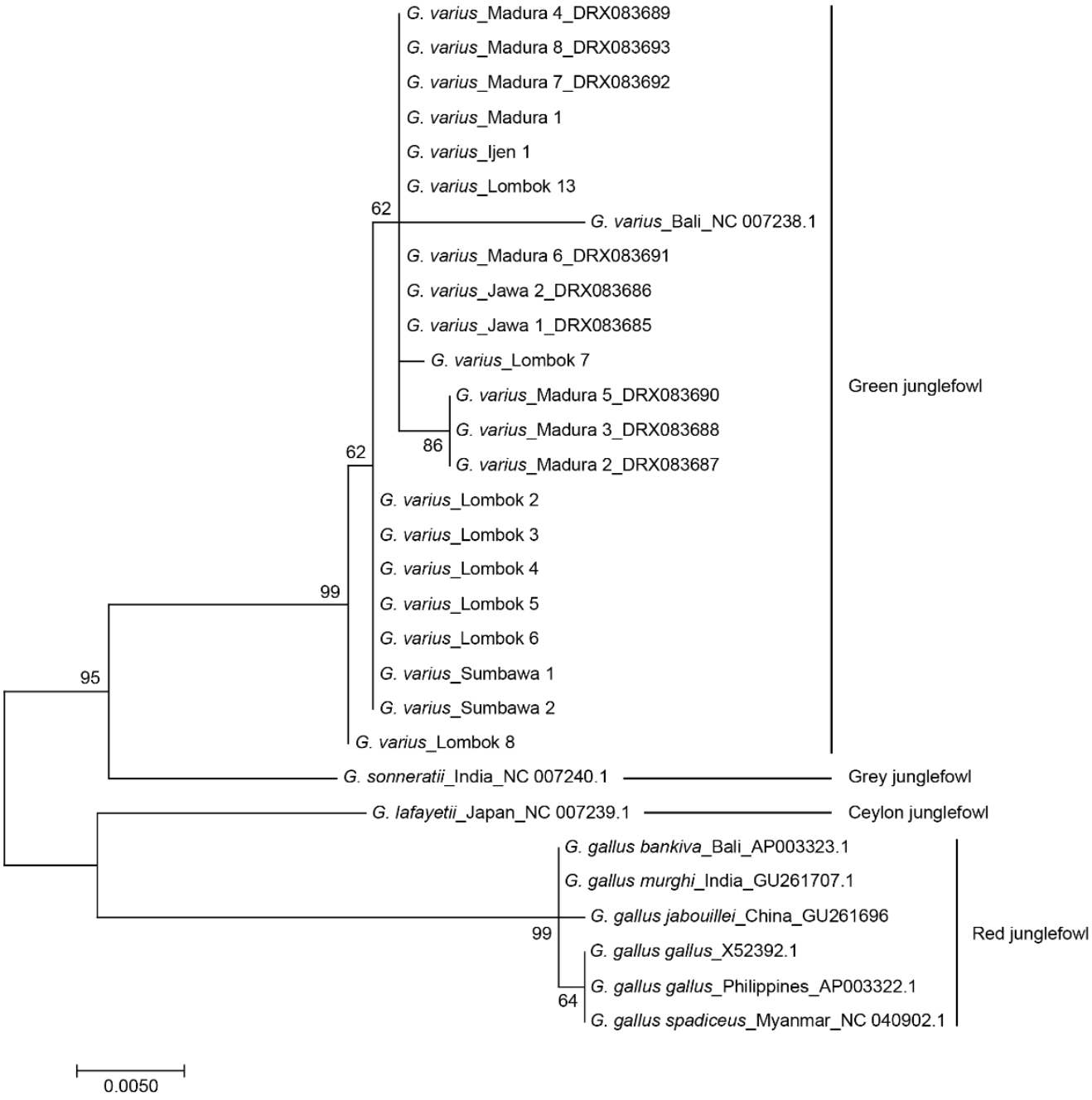

Figure 2. Maximum Likelihood phylogenetic tree of *Gallus* species based on middle-lower section of mtDNA control region, bootstrap value on each branch based on HKY model

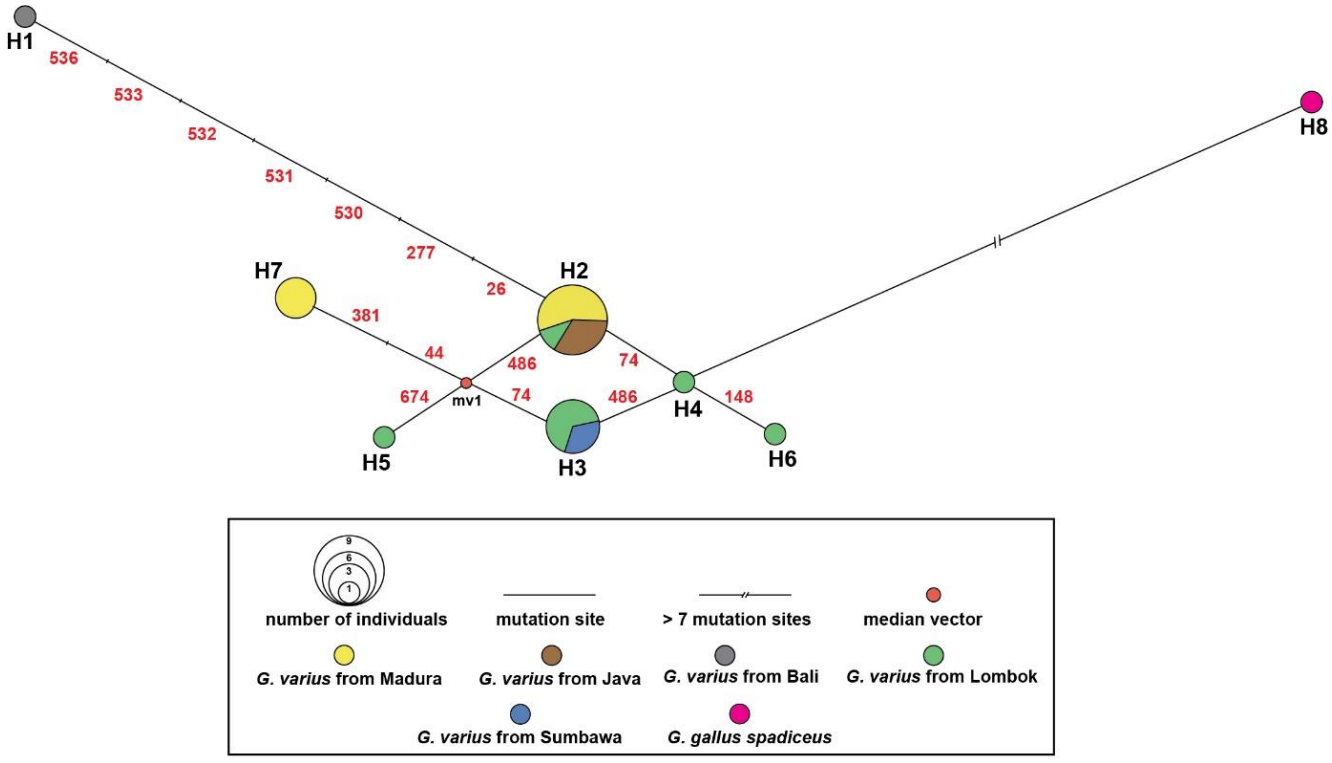

Figure 3. Haplotype network of middle-lower section of mtDNA control region of *G. varius* and *G. gallus spadiceus* (as an outgroup), constructed based on median joining method

composition of elements within the DNA region. It is worth highlighting that, among the three assessed domains, only domain I in the has been reported to contain a tandem repeat element (Figure 4) (Randi and Lucchini 1998). This distinctive characteristic renders it vulnerable to mutational events, implying an elevated mutation rate in this specific domain.

Meyer *et al.* 1999 study, centered on nucleotide substitution patterns and heterogeneity rates within human mitochondrial DNA, indicated that the substitution rate in Hyper Variable Region 1 (HVR1/ domain I) was twice as pronounced as that in Hyper Variable Region 2 (HVR2/domain III). This distinction was attributed to an elevated frequency of transition substitutions involving pyrimidine-based nucleotides in domain I. A similar predominance of pyrimidine substitution was also recognized in the investigation conducted by Zein and Sulandari in 2008. Our study has revealed that pyrimidine substitutions were observed in only 23.1% of the total mutational events. If we combined the previous study based on domain I and our study (comprises domain II and III), the frequency of pyrimidine substitution in *G. varius* was in congruent with Meyer's study (in human). This specific pattern of variation was also found in several genera in Phasianidae, namely *Tetraophasis, Alectoris, Francolinus, Perdix, Arborophila, Bambusicola, Ithaginis, Tragopan, Pucrasia, Lophophorus, Gallus* (only *G. gallus*), *Lophura, Crossoptilon, Syrmaticus, Phasianus, Chrysolophus, Polyplectron*, and *Pavo* (Huang and Ke 2014).

In our investigation, the haplotype diversity value was determined to be lower than that reported in the earlier study, which focused on domain I. Nevertheless, it is noteworthy that this value still maintains a relatively high level at 0.7662. Haplotype diversity values range between 0 and 1, signifying the probability of dissimilarity between two randomly selected alleles. Consequently, a value of 0.7662 indicates a greater than 76% likelihood that two arbitrarily chosen *G. varius* individuals from a specific location will possess different alleles within the middle-lower section of control region mtDNA.

The average genetic distance observed among *G. varius* datasets significantly exceeds that of G. *gallus*, in line with earlier research by Ulfah *et al.* in 2016 that employed single nucleotide polymorphisms (SNPs) from whole genome sequencing as a genetic marker. These findings signify a greater level of genetic differentiation within *G. varius* in comparison to G.

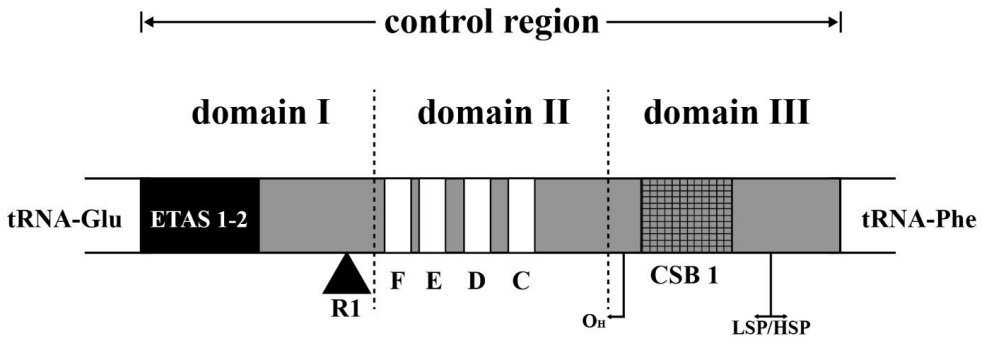

Figure 4. Schematic structure of mtDNA control region in Aves, contructed based on mtDNA control region sequences of *Alectoris* (Randi and Lucchini 1998), ETAS: extended termination-associated sequences; R: tandem repeat; F, E, D and C: conserved block F, E, D and C, respectively; CSB: conserved sequence block; OH: origin of H-strand (heavy strand) replication; LSP: light-strand transcription promoter; HSP: heavy-strand transcription promoter

Gallus. Furthermore, the genetic distance value of *G. varius* from Bali and Madura (Gv_Bli_NC_007238.1, Gv_Md5_DRX083690, Gv_Md3_DRX083688, and Gv_Md2_DRX083687) with the others *G. varius* was considerably high, suggesting the possibility of multiple subspecies contributing to the analysed *G. varius* population. The phylogenetic tree's topology further lends support to this hypothesis. The current hypothesis requires additional validation, specifically through the inclusion of morphological characteristics, morphometrics, vocalizations, and genetic information (Hull *et al.* 2014; Maley *et al.* 2016; Balakirev *et al.* 2017).

The pairwise fixation index (Fst), both on a global scale and between specific geographical locations, highlights the occurrence of gene flow within *G. varius*, with a particular emphasis on *G. varius* populations originating from Java, Madura, Lombok, and Sumbawa islands. Furthermore, a discernible degree of allelic differentiation is observed, despite the fact that many of the Fst values fail to achieve statistical significance. The findings emphasize the importance of increasing the sample size to bolster the statistical validity of the observed Fst values. The impact of anthropogenic factors is likely to underlie the gene flow observed in *G. varius*. The likelihood of natural migration is considered unlikely, given the substantial geographical barrier presented by the wide inter-island strait and the documented constrained flight abilities of *Gallus* species. Meijer *et al.* (2023), in their research investigating the origins of skeletal elements discovered during archaeological excavations at Liang Bua, Flores, Indonesia. Their findings propose that *Gallus* species, specifically the Red junglefowl (G. *gallus*) and potentially *G. varius*, were introduced to the Wallacean region through human-mediated dispersal. The absence of chicken fossils from Island Southeast Asia (ISEA) further supports this hypothesis. The haplotype network exhibits a separation between *G. varius* from Bali and those from other regions, highlighting the existence of population differentiation within *G. varius*.

The genetic diversity of *G. varius*, as evaluated within the middle-lower region of the mtDNA control region, was found to be notably high, as highlighted by the haplotype diversity value. These results concur with prior studies utilizing domain I of the mitochondrial DNA control region, whole-genome single nucleotide polymorphisms, and mitochondrial genome-wide assessments. The present investigation, along with two preceding studies (Zein and Sulandari *et al.* 2008; Ulfah *et al.* 2016), collectively suggests that the genetic diversity of *G. varius* has remained stable over the past two decades. Notably, despite the escalating demand for *G. varius* as a pet, as well as challenges such as habitat fragmentation and loss, common threats faced by numerous wildlife species, discernible genetic diversity loss has yet to manifest (Minh Nguyen *et al.* 2002; Wan *et al.* 2018; Schlaepfer *et al.* 2018; Zhang *et al.* 2018). Ongoing conservation efforts remain imperative for this species, as there is a need to monitor its genetic diversity, which plays a crucial role in enhancing its survivability in its natural habitat. The analysis of 22 *G. varius* DNA sequences have revealed the presence of 7 distinct haplotypes, including two shared haplotypes between location and five unique ones (specific to only 1 location). Nearly in all eukaryotes, mtDNA was inherited maternally, based on this phenomenon the shared haplotype (H2 and H3) between location

underscoring the maternal linkage inherent in the development of each haplotype at respective locations. These findings, specifically based on genetic distance, phylogenetic tree and haplotype network, strongly suggest the potential existence of multiple subspecies within *G. varius* and indicate clear signs of population differentiation. To obtain a more comprehensive and accurate understanding of the genetic structure of *G. varius*, additional biological data must be incorporated.

Acknowledgements

We would like to thank to Ministry of Research and Higher Education for financial support through the grant of PMDSU no. 1/E1/KP.PTNBH/2021.

References

[DDC] DNA Diagnostic Center, c2019. Avian sample collection procedure, feather collection. Available at: https://www.vetdnacenter.com/files/collecting-feather-samples.pdf. [Date accessed: 25 August 2019]

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment search tool. *J. Mol. Biol.* 215, 403-410. https://doi.org/10.1016/S0022-2836(05)80360-2

Balakirev, A.E., Abramov, A.V., Rozhnov, V.V., 2017. The phylogeography of red spiny rats *Maxomys surifer* (Rodentia, Muridae) in Indochina with comments on taxonomy and description of new subspecies. *Zool. Stud.* 56, e6. https://doi.org/10.6620%2FZS.2017.56-06

Bandelt, H.J., Forster, P., Röhl, A., 1999. Median-joining networks for inferring intraspecific phylogenies. *Mol. Biol. Evol.* 16, 37-48. https://doi.org/10.1093/oxfordjournals.molbev.a026036

BirdLife International, 2018. *Gallus varius* the IUCN red list of threatened species 2018: e.T22679213A131905483. Available at: http://dx.doi.org/10.2305/IUCN.UK.2018-2.RLTS.T22679213A131905483.en. [Date accessed: 1 August 2019]

Clements, J.F., Schulenberg, T.S., Iliff, M.J., Billerman, S.M., Fredericks, T.A., Sullivan, B.L., Wood, C.L. 2019. The eBird/Clements Checklist of Birds of the World: v2019. Available at: https://www.birds.cornell.edu/clementschecklist/download/. [Date accessed: 1 August 2019]

Desjardins, P., Morais, R., 1990. Sequence and gene organization of the chicken mitochondrial genome: a novel gene order in higher vertebrates. *J. Mol. Biol.* 212, 599-634. https://doi.org/10.1016/0022-2836(90)90225-b

Excoffier, L., Lischer, H.E.L., 2010. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Mol. Ecol. Resour.* 10, 564-567. https://doi.org/10.1111/j.1755-0998.2010.02847.x

Fluxus Technology Ltd, c2021. Network version 10.2.0.0. Available at: www.fluxus-engineering.com. [Date accessed: 28 July 2021]

Huang, Z., Ke, D., 2014. Structure and evolution of the Phasianidae mitochondrial DNA control region. *Mitochondrial DNA A DNA Mapp. Seq. Anal.* 27, 350-354. https://doi.org/10.3109/19401736.2014.895987

Hull, J.M., Englis, A., Medley, J.R., Jepsen, E.P., Duncan, J.R., Ernest, H.B., Keane, J.J., 2014. A new subspecies of great gray owl (*Strix nebulosa*) in the Sierra Nevada of California, USA. *J. Raptor Res.* 48, 68–77. https://doi.org/10.3356/JRR-13-35.1

Kumar, S., Stecher, G., Tamura, K., 2016. MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol. Biol. Evol.* 33, 1870-1874. https://doi.org/10.1093/molbev/msw054

Madge, S., McGowan, P., 2010. *Pheasants, Partridges & Grouse: Including buttonquails, sandgrouse and allies*. Bloomsbury Publishing, London.

Maley, J.M., McCormack, J.E., Tsai, W.L.E., Schwab, E.M., Dort, J.V., Juarez, R.C., Carling, M.D., 2016. Fonseca mangrove rail: a new subspecies from Honduras. *West Birds.* 47, 262-273. https://doi.org/10.21199/wb47.4.1

Marshall, H.D., Baker, A.J., 1997. Structural conservation and variation in the mitochondrial control region of fringilline finches (*Fringilla* spp.) and the greenfinch (*Carduelis chloris*). *Mol. Biol. Evol.* 14, 173-84. https://doi.org/10.1093/oxfordjournals.molbev

Meijer, H.J., Walker, S.J., Sutikna, T., Wahyu Saptomo, E., Tocheri, M.W., 2023. Why did the chicken cross the Wallace Line? Archeological evidence suggests human-mediated dispersal of *Gallus* to Flores first occurred at least~ 2.25 ka cal BP. *Int. J. Osteoarchaeol.* 33, 631-641. https://doi.org/10.1002/oa.3192

Meyer, S., Weiss, G., Haeseler, A.V., 1999. Pattern of nucleotide substitution and rate heterogeneity in the hypervariable region I and II of human mtDNA. *Genetics*. 152, 1103-1110. https://doi.org/10.1093/genetics/152.3.1103

Miao, Y.W., Peng, M.S., Wu, G.S., Ouyang, Y.N., Yang, Z.Y., Yu, N., Liang, J.P., Pianchou, G., Beja-Pereira, A., Mitra, B., Palanichamy, M.G., Baig, M., Chaudhuri, T.K., Shen, Y.Y., Kong, Q.P., Murphy, R.W., Yao, Y.G., Zhang, Y.P., 2013. Chicken domestication: an updated perspective based on mitochondrial genomes. *Heredity*. 110, 277-282. https://doi.org/10.1038/hdy.2012.83

Minh Nguyen, D., Lan Phan Nguyen, H., Minh Nguyen, T., 2022. Genetic structure of the endemic *Dipterocarpus condorensis* revealed by microsatellite markers. *AoB Plants.* 14, 1-9. https://doi.org/10.1093/aobpla/plac007

Nishibori, M., Shimogiri, T., Hayashi, T., Yasue, H., 2005. Molecular evidence for hybridization of species in the genus *Gallus* except for *Gallus varius*. *Anim. Genet.* 36, 367-375. https://doi.org/10.1111/j.1365-2052.2005.01318.x

Randi, E., Lucchini, V., 1998. Organization and evolution of the mitochondrial DNA control region in the avian genus *Alectoris*. *J. Mol. Evol.* 47, 449-62. https://doi.org/10.1007/pl00006402

Rozas, J., Ferrer-Mata, A., Sánchez-DelBarrio, J.C., Guirao-Rico, S., Librado, P., Ramos-Onsins, S.E., Sánchez-Gracia, A., 2017. DnaSP 6: DNA sequence polymorphism analysis of large data sets. *Mol. Biol. Evol.* 34, 3299-3302. https://doi.org/10.1093/molbev/msx248

Saccone, C., Pesole, G., Sbisa, E., 1991. The main region of mammalian mitochondrial DNA structure-function model and evolutionary pattern. *J. Mol. Evol.* 33, 83-91. https://doi.org/10.1007/bf02100199

Schlaepfer, D.R., Braschler, B., Rusterholz, H.P., Baur, B., 2018. Genetic effects of anthropogenic habitat fragmentation on remnant animal and plant populations: A meta-analysis. *Ecosphere.* 9, e02488. https://doi.org/10.1002/ecs2.2488

Ulfah, M., Kawahara-Miki, R., Farajallah, A., Muladno, M., Dorhorst, B., Martin, A., Kono, T., 2016. Genetic features of red and green junglefowls and relationship with Indonesian native chickens Sumatera and Kedu Hitam. *BMC Genomic.* 17, 1-9. https://doi.org/10.1186/s12864-016-2652-z

Wan, H.Y., Cushman, S.A., Ganey, J.L., 2018. Habitat fragmentation reduces genetic diversity and connectivity of the mexican spotted owl: a simulation study using empirical resistance models. *Genes.* 9, 403. https://doi.org/10.3390%2Fgenes9080403

Zein, M.S.A., Sulandari, S., 2008. Struktur populasi genetik Ayam Hutan Hijau menggunakan sekuen hypervariable-1 D-loop DNA mitokondria. *Biota.* 13, 182-190. https://doi.org/10.24002/biota.v13i3.2573

Zhang, X., Shi, M.M., Shen, D.W., Chen, X.Y., 2018. Habitat loss other than fragmentation per se decreased nuclear and chloroplast genetic diversity in a monoecious tree. *PLoS One.* 7, e39146. https://doi.org/10.1371/journal.pone.0039146