INNSPUB

International Journal of Biosciences | IJB |
ISSN: 2220-6655 (Print) 2222-5234 (Online)
http://www.innspub.net
Vol. 24, No. 1, p. 152-158, 2024

RESEARCH PAPER OPEN ACCESS

Phylogenetic analysis of Gray-crowned Crocias (*Laniellus langbianis*) in Lam Dong accessed by mitochondrial DNA

Nguyen Thi Phuong Mai[1], Phan Phuoc Thinh[2], Ha Thanh Tung[1], Dinh Thanh Son[3], Nguyen Hong Hanh[3], Ho Nguyen Quynh Chi[3], Le Thanh Long[*3]

[1]*Tay Nguyen institute for Scientific Research, Academy of Science and Technology, Da Lat City, Vietnam*

[2]*Ton Duc Thang University, Ho Chi Minh, Vietnam*

[3]*Institute of Tropical Biology, Vietnam Academy of Science and Technology, Ho Chi Minh, Vietnam*

Key words: *Laniellus langbianis*, mtDNA, Phylogeny

http://dx.doi.org/10.12692/ijb/24.1.152-158 Article published on January 08, 2024

Abstract

This study aimed to evaluate the genetic diversity of *Laniellus langbianis* by using mtDNA analysis. The results showed that mtDNA coverage percentage was as follows: twenty-two tRNA genes (8.64%), thirteen PCGs (63,10%), and two rRNA genes (14,47%). *L. langbianis's* whole mitogenome had a regular A+T content of 54.1%. 13 PCGs in the *L. langbianis* mitogenome were estimated to be 11,262 bp long overall, making up 63,10% of the total mtDNA. The coverage range of the 22 t-RNA genes varied from 66 bp (trnS1, trnC) to 75 bp (trnL2). The tRNAs had an average base composition of A: 27.6%, T: 21.0%, G: 30.6%, and C: 20.8%. Phylogeny analysis revealed that *L. langbianis* is belonged to the family Leiothrichidae.

* **Corresponding Author:** Le Thanh Long ✉ lelong2510@gmail.com

Introduction

The family Leiothrichidae represents the largest clade of babblers in terms of species diversity (Cibois *et al.*, 2018), includes *Laniellus langbianis*. According to Delacour (Delacour, 1946), the *L. langbianis*, which was previously classified as a member of the family Timaliidae (Old World), but is now classified as a member of the family Leiothrichidae (New World), comprises over 400 species (Dickinson and Christidis, 2014; Gill and Donsker, 2017). Based on the melodious whistles of some species, which sometimes duet or include mimicry notes, these Old-World insectivores inhabit Africa, Eurasia, and Australasia, with the majority of their diversity found in temperate and subtropical habitats of the Sino-Himalayan montane region. The family also includes species known as "laughingthrushes" due to their loud calls when in large social groups. Although discovered quite early, it was discovered again 56 years later, on January 29th, 1994, in Chu-Yang-Sin-forest (Dak Lak). Therefore, *L. langbianis* has been in Vietnam Red List of Endangered Species since 1992 with E (Endangered) level (Vietnam Red List of Endangered Species, 2003), IUCN with EN (Endangered) level (IUCN, IUCN Species Survival Commission, 2001). The name of this species is a combination of "*Laniellus*", which is the order this species belong to under the rules of The International Code of Zoological Nomenclature, and "*Langbianis*" come from where it was discovered, which belongs to Langbiang plateau, Lam Dong, Vietnam.

The Gray-crowned Crocias is endemic to Vietnam, mostly found in the Lam Dong plateau at the height from 910 to 1450 meters above the sea level in tropical montane evergreen forest area with such an extremely small number. Nowadays they recently found in Mang Den, NgocLinh mountain (Kon Tum) and rare resident in Dalat plateau, seen often at Ta Nung valley and Tuyen Lam as well as a small part forest at Dam Rong (Vietnam Red List of Endangered Species, 2003). The adult bird has a distinct appearance with its slaty-grey crown, blackish mask, and boldly blackish-streaked flanks. Its dull rufous upperparts are adorned with blackish- brown streaks and faint, pale shaft streaks on the crown, nape, and mask. The rest of its underparts are white. The tail is mostly slaty-grey with white tips, while the greater coverts and secondaries are mostly grey. The juvenile has a browner crown with broader buffish streaks, duller head-sides, smaller flank-streaks, browner greater coverts and secondary feathers, and narrower, white tail feather tips (Vietnam Red List of Endangered Species). The objectives of this thesis include is to sequencing the mitochondrial genome of *Laniellus langbianis* collected in Langbiang plateau, Lam Dong, Vietnam and stored at the Tay Nguyen Institute of Science. Moreover, we also assessing phylogenetic tree based on the sequence of *Laniellus langbianis*' mitochondrial genome to evaluate the genetic diversity.

Materials and methods

Samples collection

Gray-crowned crocias tissue collected in Langbiang plateau, Lam Dong, Vietnam, stored at Tay Nguyen Institute of Scientific Research. The sample was got permission from the local authorities and relatives. The samples were kept in −20°C and transported to the laboratory. The remainders of sequences were derived from GenBank.

DNA extraction

Total DNA were extracted from Gray-crowned crocias samples using GeneJET Genomic DNA Purification Kit (K0721, Thermo scientific). Total DNA was resuspended with TE buffer and preserved at −20°C.

Sequence alignment

Annotation of mitochondrial DNA was performed with Mitos WebServer and MitoFish (Bernt et al, 2013; Iwasaki *et al.*, 2013). A gene map of M. indica's whole mitogenome was produced using MitoAnnotator. The Mitos WebServer and tRNAscan-SE software were used to validate the transfer RNA (t-RNA) predictions and their secondary structures (Lowe and Eddy, 1997). By comparing with other artiodactyl mitogenomes, the r-RNAs, PCGs, and control region were found.

MEGA 6 was used to estimate and compare the amino acid and nucleotide (A + T content) compositions for

Gray-crowned Crocias as well as other representative species. The established method was followed to determine the AT and GC skew values in order to estimate the bias in nucleotide composition among the genes of the complete mitogenome of Gray-crowned Crocias. AT-skew = (A – T)/(A + T) and GC-skew = (G – T)/(G + T) (Perna and Kocher, 1995).

Phylogenetic analysis

The sequence of Gray-crowned crocias will be compared with the other groups derived from Genbank using MEGA6 program. The cytochrome b sequences will be aligned using CLUSTAL W. Tamura & Nei model which used as genetic distance model. Neighbor-joining method was used for phylogenetic tree construction. Bootstrap analyses (1000 replications) are applied to estimate the confidence in branching order.

Results and Discussion

Genome structure, organization and composition

The first recorded mitochondrial genome of *Laniellus langbianis* was showed in this paper. This mtDNA consist of 17,849 bp (Fig. 1), which is distinctly larger than *Corvus moneduloides* (16,846 bp), *Phainopepla nitens* (16,892 bp). This mitochondrion contains 13 PCGs (cox1-3, nad1-6, nad4L, Cyt-B, atp6 and atp8), 22 tRNAs (one for each amino acid, two for Leucine and Serine), 2 rRNAs (rrnS and rrnL) and one A+T rich region (Fig. 1, Table 1), which was separated into two group called HSV1 and HSV2 by tRNA-Thr and tRNA- Pro, tRNA – Glu and tRNA – Phe, respectively.

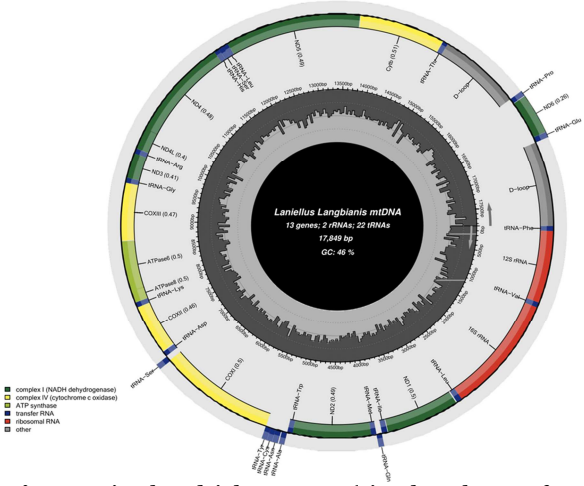

Fig 1. Mitochondrial genome (circular shape of *L. langbianis*)

Table 1. The organization and characterization of the complete mitochondrial genome of *Laniellus langbianis*

Name	Start	Stop	Strand	Length
trnF(ttc)	60	127	+	68
rrnS	127	1110	+	984
trnV(gta)	1110	1179	+	70
rrnL	1180	2778	+	1599
trnL2(tta)	2779	2853	+	75
nad1	2884	3837	+	954
trnI(atc)	3855	3928	+	74
trnQ(caa)	3934	4004	-	71
trnM(atg)	4004	4072	+	69
nad2	4073	5107	+	1035
trnW(tga)	5113	5182	+	70
trnA(gca)	5184	5252	-	69
trnN(aac)	5265	5338	-	74
trnC(tgc)	5339	5404	-	66
trnY(tac)	5405	5474	-	70
cox1	5476	7017	+	1542
trnS2(tca)	7018	7090	-	73
trnD(gac)	7101	7169	+	69
cox2	7180	7851	+	672
trnK(aaa)	7865	7933	+	69
atp8	7935	8096	+	162
atp6	8093	8773	+	681
cox3	8784	9566	+	783
trnG(gga)	9568	9635	+	68
nad3	9636	9983	+	348
trnR(cga)	9986	10055	+	70
nad4l	10057	10350	+	294
nad4	10347	11714	+	1368
trnH(cac)	11725	11794	+	70
trnS1(agc)	11795	11860	+	66
trnL1(cta)	11860	11930	+	71
nad5	11952	13736	+	1785
cob	13769	14890	+	1122
trnT(aca)	14904	14972	+	69
trnP(cca)	16082	16150	-	69
nad6	16160	16675	-	516
trnE(gaa)	16677	16748	-	72

Each gene group's overall mtDNA coverage percentage was as follows: 2 rRNA genes (14,47%), 22 tRNA genes (8.64%), and 13 PCGs (63,10%). In comparison to other species in the family Leiothrichidae, such as *Corvus moneduloides* (66.7% for PCGs, 9.15% for tRNAs, 15.27% for rRNA) and *Phainopepla nitens* (66.6 % for PCGs, 9.15% for tRNAs, 15.3% for rRNA), a similar pattern was recorded. All members of the Leiothrichidae generally had the same gene order and gene orientation seen in the mitochondrial genome of *L. langbianis*. It can be easily seen that while almost all genes of the *L. langbianis* were located in H-strand, others founded in L-strand, which are trnQ, trnA, trnN, trnC, trnY, trnS2, trnP, nad6, trnE.

Table 2. Nucleotide composition indices in various regions of ten representative mitogenomes of the family Leiothrichidae.

Species	Family	Whole		PCGs		Large ribosomal RNA		Small ribosomal RNA	
		Length (%)	AT (%)	Length (%)	AT (%)	Length (%)	AT (%)	Length (%)	AT (%)
L. Langbianis	Leiothrichidae	17849	54.5	11262	54.1	1599	54.7	984	49.8
C. moneduloides	Corvidae	16892	55.5	11282	55.2	1601	56.3	979	51.3
P. nitens	Bombycillidae	16846	54.9	11255	54.7	1601	56.4	984	52.4
P. raggiana	Paradisaeidae	16910	56.4	11268	56.1	1606	56.8	978	53
D. brunneopygia	Petroici dae	16851	54.5	11256	54.4	1606	55.5	984	51
L. ludovicianus	Laniidae	16947	56.1	11274	55.6	1603	57.4	975	51.7
D. chrysoptera	Passeriformes	16925	57.2	11268	56.8	1600	57.4	984	52.7
S. lunatus	Eurylaimidae	16989	52.9	11264	52.3	1583	52.1	973	50.8
C. viridis	Eurylaimidae	14910	53.1	1073	53.1	1593	53.7	982	50.1
A.phoenixceus	Icteridae	16776	54.1	11280	53.8	1598	55.7	976	51.6

Table 3. Amino acid frequency in PCGs, tRNAs, and rRNAs

	PCGs	tRNAs	Ribosomal RNA
Ala (A)	7.9	5.6	7.2
Arg (R)	1.9	1.5	3.0
Asn (N)	3.6	3.9	5.9
Asp (D)	1.8	1.5	3.4
Cys (C)	0.8	2.4	2.1
Gln (Q)	2.5	4.5	6.3
Glu (E)	2.2	4.5	2.9
Gly (G)	6.0	5.2	6.4
His (H)	2.7	3.5	3.0
Ile (I)	7.9	2.4	3.3
Leu (L)	17.5	14.1	9.9
Lys (K)	2.3	9.7	8.7
Met (M)	4.7	3.2	2.8
Phe (F)	5.6	4.3	1.9
Pro (P)	5.9	4.8	8.6
Ser (S)	7.3	9.5	4.4
Thr (T)	8.9	5.6	8.4
Trp (T)	2.8	1.5	2.6
Tyr (Y)	2.9	2.6	2.9
Val (V)	4.7	9.5	3.1

Base composition and skewness

The parameters AT-skew, GC-skew, and A+T content are frequently used to investigate the pattern of nucleotide composition in mitochondrial genomes. The complete mitogenome of *L. langbianis* had a regular A+T content (54.1%), similar to other species in the family Leiothrichidae (Table 2), with trnT having the highest A+T content (69.6%) while the smallest amout of A+T content was recorded in trnF gene (48.5%).

From Table 3, it can be easily seen that while Leu was the most frequent amino acid in PCGs, tRNAs and rRNA, which is 17.5%, 14.1%, and 9.9% respectively. In the order hand, the lowest percent of each amino acid in PCGs, tRNAs, and rRNAs groups is Cys (0.8%), both Asp (1.5%) and Trp (1.5%), and Phe (1.9%). Amino acids that have a high frequency is Gly (6.0%), Ala (7.9%), Ile (7.9%), Ser (7.3%), Thr (8.9%) in PCGs; Lys (9.7%), Ser (9.5%), and Val (9,5%) in tRNAs; and Ala (7.2%), Gln (6.3%), Gly (6.4%), Lys (8.7%), Pro (8.6%), Thr (8.4%) in rRNAs. This pattern is similar to *Corvus moneduloides* but different from *Phainopepla nitens*, which Pro was the most frequent amino acid.

Table 4. Amino acids frequency of *L. langbianis*, *C. moneduloides*, and *P. nitens*.

	L. Laniellus	*C. Moneduloides*	*P. nitens*
Ala (A)	4.2	5.3	6.2
Arg (R)	3.2	2.8	2.7
Asn (N)	6.5	4.9	4.2
Asp (D)	1.9	2.2	2.4
Cys (C)	1.7	1.9	1.4
Gln (Q)	4.8	4.3	4.3
Glu (E)	2.5	2.4	2.8
Gly (G)	2.6	3.6	3.8
His (H)	6.2	4.5	4.4
Ile (I)	4.5	5.0	4.5
Leu (L)	10.7	11.9	12.7
Lys (K)	4.6	4.5	4.8
Met (M)	3.1	3.4	3.3
Phe (F)	3.3	3.8	4.1
Pro (P)	11.8	9.5	9.5
Ser (S)	10.2	10.4	9.8
Thr (T)	10.2	9.5	9.1
Trp (T)	2.0	2.4	2.1
Tyr (Y)	3.6	4.1	4.0
Val (V)	2.4	3.5	3.6

The most common amino acid of *Laniellus langbianis*, *Corvus moneduloides* and *Phainopepla nitens* was showed in Table 4 and Fig. 2. In detail, Pro, Leu, Lei is the most common amino acid in *Laniellus langbianis*, *Corvus moneduloides* and *Phainopepla nitens* mtDNA respectively while Cys was uncommon in these three species.

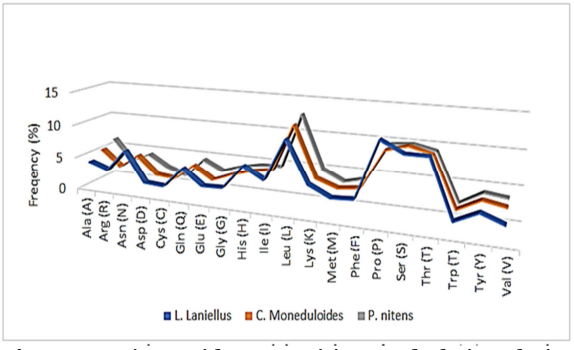

Fig. 2. Amino acid composition and their relative frequency (%) in complete mitogenome of *L. langbianis, C. moneduloides* and P. *nitens*

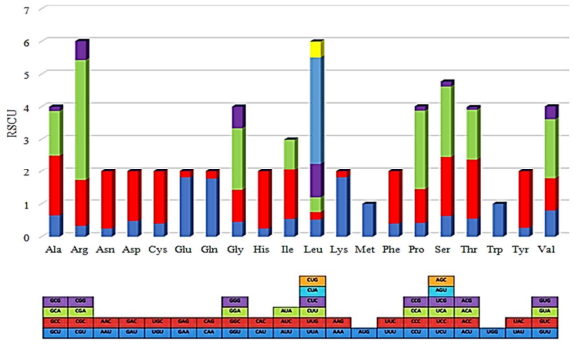

Fig. 3. The Relative Synonymous Codon Usage (RSCU) of the mitochondrial protein- coding genes of L. *Langbianis*. Different codons present in PCGs are plotted on X axis. Codons which are not present in mitogenome are not indicated.

Protein-coding genes and rate of evolution

The total length of 13 PCGs in the mitogenome of *L. langbianis* was estimated to be 11,262 bp, accounted for 63,10% of the total mtDNA. While 12 PCGs were located on the H – strand (majorrity strand), the nad6 gene located on the L – strand (minority strand), which is similar to *C. moneduloides* and *P. nitens*. The A + T content of 13 PCGs in *L. langbianis* was 54.1%, with the lowest and highest A+T content was 50.7% (nad4l) and 56.8% (atp8). Fig. 3 show relative Synonymous Codon Usage (RSCU) of the mitochondrial protein- coding genes of *L. langbianis*.

Base skew was calculated to determine the degree of base bias among all PCGs. Table 5 show the average AT and GC skew values in *L. langbianis*' mtDNA in contrast to other species in the Leiothrichidae family. Positive AT-skew values were identified in the major of PCGs, showing that adenines occur more frequently than thymines, which is similar to other species in the Leiothrichidae family that was showed in Table 5. Negative GC skewness was founded for most of the PCGs of *L. langbianis* with the value between -0.571 and -0.299. The nad6 region showed an AT skew deviation of – 0.620 and a GC skew deviation of 0.628, which was likewise seen in *C. moneduloides* (AT skew= - 0.418, GC skew=0.500) and *P. nitens* (AT skew= -0.535, GC skew=0.535). Fig. 4 depicts the trend of the AT-skew and GC-skew values over all 13 PCGs of *L. langbianis*. 12 out of 13 PCGs began with ATG or ATA (putative codons), ATT for cob gene, differed in nad1 (ATC) and cob (ATG) of *C. moneduloides*.

Table 5. The AT and GC skew in the protein-coding genes of ten representative mitogenomes of the family Leiothrichidae used in this study

Species	Protein Coding Genes (PCGs)					
	T	C	A	G	AT-skew	GC-skew
L. langbiangis	23.8	30.1	30.3	15.8	0.120	-0.312
C. moneduloides	25.3	29.0	29.9	15.9	0.083	-0.292
P. nitens	24.8	29.2	29.9	16.1	0.093	-0.289
P. raggiana	25.7	27.9	30.4	16.0	0.084	-0.271
D. brunneopygia	24.9	29.3	29.5	16.3	0.085	-0.285
L. ludovicianus	25.8	28.2	29.8	16.2	0.072	-0.270
D. chrysoptera	26.3	27.9	30.5	15.2	0.074	-0.295
S. lunatus	24.0	31.8	28.3	15.9	0.082	-0.333
C. viridis	22.4	33.3	30.7	13.6	0.156	-0.420
A. phoeniceus	24.5	30.4	29.3	15.8	0.120	-0.312

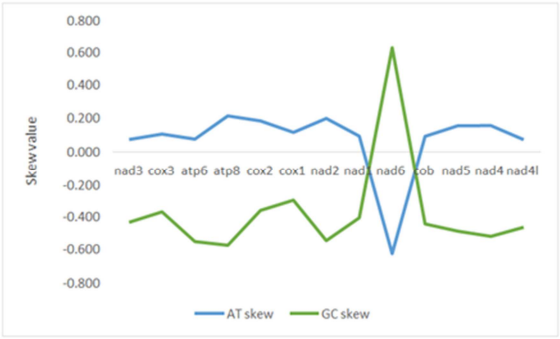

Fig. 4. Graphical representation of AT- and GC-skew in all the 13 protein-coding genes of *L. langbianis* mitogenome

Ribosomal RNA and transfer RNA genes

In the mitogenome of *M. indica*, the rrnS and rrnL genes were found between trnF and trnV, and between trnV and trnL1, respectively. Both rRNAs were separated by trnV, which is common in most vertebrates. The rrnS and rrnL lengths were 984bp and 1599bp, respectively.

The total A+T content of t rrnS and rrnL was 49.8% and 54.7%, which matches with the other two species in Leiothrichidae family, the *C. Moneduloides* with the total A+T content in rrnS and rrnL was 51.3% and 56.3% respectively, and P. nitens with 52.4% (rrnS) and 57.4% (rrnL). The length and A+T concentration of both rRNAs were very similar among all Leiothrichidae representative species.

Table 6. The details of the mismatched t-RNA base pairs from *M. indica*. AA=amino acid acceptor, TΨC=pseudouridine, AC=anticodon

tRNA	MBP	Stem	Frequency
trnF	A-C	AA	1
	A-C	TΨC	1
trnL2	C-C	AA	1
trnM	A-C	AA	1
	UU	TΨC	1
trnW	A-A	D-arm	1
trnN	A-C	AC	1
trnG	U-U	AC	1
	C-C	AC	1

tRNAscan-SE calculated the total number of tRNA genes coding for amino acids in the mitogenome of *L. langbianis*. The anticodons of all the tRNAs discovered in *L. langbianis'* whole mtDNA were identical to those discovered in other Leiothrichidae species. The coverage range of the 22 t-RNA genes varied from 66 bp (trnS1, trnC) to 75 bp (trnL2). The tRNAs had an average base composition of A: 27.6%, T: 21.0%, G: 30.6%, and C: 20.8%, with trnT having the greatest AT content (53.1%) and trnF having the lowest (48.1%). 14 of the 22 tRNA genes were found on the H strand (trnF, trnV, trnL2, trnI, trnM, trnW, trnD, trnK, trnG, trnR, trnH, trnS1, trnL1, trnT) while the rest were found on the L strand (trnA, trnN, trnC, trnY, trnP, trnE). All of the tRNA might form a secondary clover-leaf structure as predicted by Mitos WebServer. Beside the normal secondary base pair structure of tRNA, which is A-U, G-C, nine mismatched base pairs were discovered in six tRNAs of *L. langbianis* mitogenome. All of them had different types of mismatches with 3 on the amino acid acceptor stems, 3 on the anticodon stems, two on the TΨC (pseudouridine) stems, and one on the D-arm stem (Table 6).

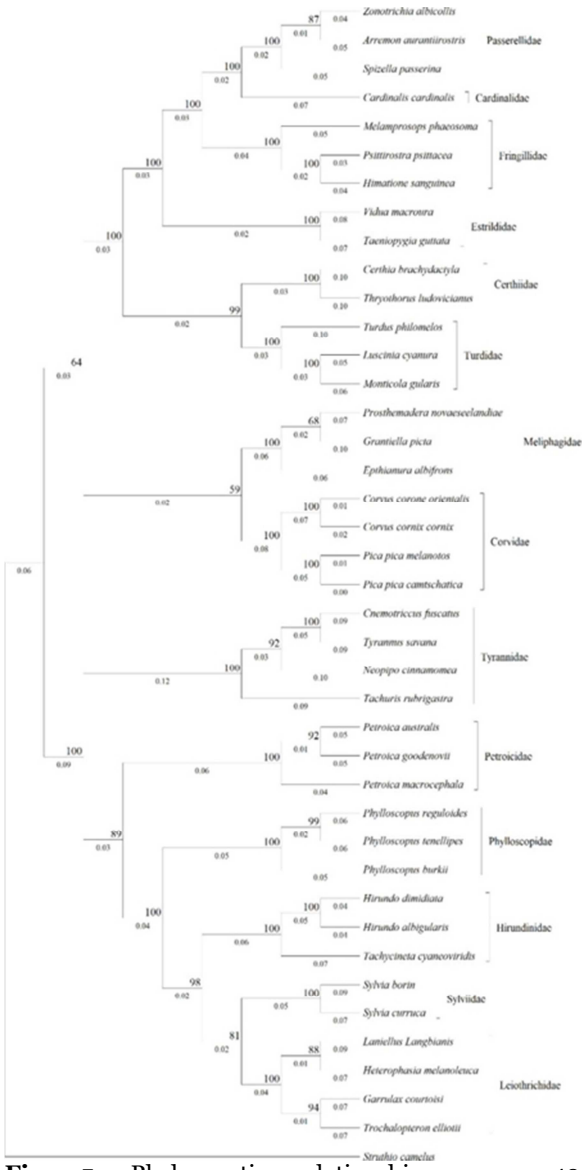

Fig. 5. Phylogenetic relationship among 40 mitogenomes of Passeriformes, constructed from whole mitochondrial genome using Maximum Likelihood (ML) by MEGA 6

Phylogenetic relationship

We provide a fully resolved phylogenetic tree of species in various families of the order Passeriformes. The whole mitochondrial genome was used from 14 families, each containing one or multiple representatives, all together 40 species. All together compare with *L. langbianis*, which is belong to the family Leiothrichidae (includes *Laniellus langbianis*, *Garrulax courtoisi*, *Trochalopteron elliotii*, *Heterophasia melanoleuca*) with the bootstrap value of 100 (Fig. 5). The entire tree was rooted by the *Struthio camelus*, which is a species in the Struthionidae family.

The tree topology was consistent in ML analysis with high bootstrap support (>90). The closest living relatives of Leiothrichidae were Sylviidae, which is similar to other tree from previous study (Cibois *et al.*, 2018).

Conclusion

The results of the study will contribute to the data collection of taxonomy and genetic diversity for the species *Laniellus langbianis*. With the discovering of *Laniellus langbianis*' mitochondrial genome, it will help to improve the biodiversity of Vietnamese birds, in order to contribute to the world's bird phylogenetic tree.

Acknowledgements

This study was supported by Tay Nguyen institute for Scientific Research, VAST.

References

Cibois A, Gelang M, Alström P. 2018. Comprehensive Phylogeny of the Laughingthrushes and Allies (Aves, Leiothrichidae) and a Proposal for a Revised Taxonomy. Zoologica Scripta **47(4)**, 428–440.

Delacour J. 1946. Les timaliinés. L'oiseau **16**, 7–36.

Dickinson EC and Christidis L. 2014. The Howard & Moore Complete Checklist of the birds of the world (4th ed.). Eastbourne, UK: Aves Press.

Gill F and Donsker D. 2017. IOC World Bird List (v 7.3).

Vietnam Red List of Endangered Species. 2023.

IUCN. 2001. IUCN Red List Categories and Criteria: Version 3.1. IUCN Species Survival Commission. IUCN, Gland, Switzerland and Cambridge, UK.

Bernt M, Donath A, Jühling F, Externbrink F, Florentz C, Fritzsch G, Pütz J, Middendorf M, Stadler PF. 2013. MITOS: Improved de novo Metazoan Mitochondrial Genome Annotation. Mol. Phylogenet. Evol. **69**, 313–319.

Iwasaki W, Fukunaga T, Isagozawa R, Yamada K, Maeda Y, Satoh TP, Sado T, Mabuchi K, Takeshima H, Miya M, Nishida M. 2013. MitoFish and MitoAnnotator: A mitochondrial genome database of fish with an accurate and automatic annotation pipeline. Mol. Biol. Evol. **30**, 2531–2540.

Lowe TM and Eddy SR. 1997. tRNAscan-SE: A program for improved detection of transfer RNA genes in genomic sequence. Nucleic Acids Res. **25**, 955–964.

Perna NT and Kocher TD.1995. Patterns of nucleotide composition at fourfold degenerate sites of animal mitochondrial genomes. J. Mol. Evol. **41**, 353–358.