Heliyon 10 (2024) e30513

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Research article

Molecular identification and phylogenetic analysis of the mitogenome in endangered giant nuthatch *Sitta magna* (*Passeriformes*, *Sittidae*)

Ruixin Mo [a,b,1], Dong Zhu [a,b,1], Jing Sun [b,1], Qingmiao Yuan [a,b], Feng Guo [c], Yubao Duan [a,b,*]

[a] *Key Laboratory for Conserving Wildlife with Small Populations in Yunnan, Southwest Forestry University, Kunming, 650224, China*
[b] *College of Forestry, Southwest Forestry University, Kunming, Yunnan, 650224, China*
[c] *Administration of Zixi Mountain Provincial Nature Reserve, Chuxiong, 675008, China*

ARTICLE INFO

Keywords:
Sitta magna
Mitochondrial genome
Sequence analyses
Phylogeny

ABSTRACT

The Giant Nuthatch *Sitta magna* (family Sittidae) is a passerine bird, the quantification of the number of habitats and species on a global scale remains low. Most species are restricted to low elevations in southwest China, eastern Myanmar, and northern Thailand. To characterize the mitochondrial genome sequence of *S. magna* and its phylogenetic relationships with other members within the genus *Sitta*, the mitochondrial genome of *S. magna* was sequenced using the whole genome shotgun method. The sequencing results showed that the mitochondrial genome was 16,829 bp long and consisted of 13 protein-coding genes (PCGs), 22 transfer RNA genes (tRNAs), 2 ribosomal RNA genes (rRNAs), and one control region (D-loop). All tRNAs were predicted to form a typical clover secondary structure. Among the 13 PCGs, only the start codon in COI was ATC, the start codon by the remaining 12 PCGs was ATG, and the stop codons were TAG, TAA, AGG, AGA, and TA. Bayesian inference and maximum likelihood phylogenetic analysis of the sequences of 17 species generated consistent well-supported phylogenies. The family Polioptilidae and the family Troglodytidae were closely related, and the family Sittidae was confined to a single branch. The genus *Sitta* in the family Sittidae was mainly clustered into three branches. Our findings provide new mitochondrial genomic data that could be used for phylogenetic and taxonomic studies; our results also certificate into the phylogenetic relationships within the genus *Sitta* ((*S. himalayensi*+(*S. nagaensis* + *S. europaea*))+(*S. villosa* + *S. yunnanensis*))+(*S. carolinensis* + *S. magna*).

1. Introduction

The mitochondrial genome is relatively small and maternally inherited, it is also characterized by a rapid evolutionary rate and a low or non-existent recombination rate, this, coupled with the high conservation in gene number among mitochondrial genomes, has made mitochondrial data useful for phylogenetic and population studies of bird taxa [1,2]. In addition, mitochondrial genes are more

* Corresponding author. Key Laboratory for Conserving Wildlife with Small Populations in Yunnan, Southwest Forestry University, Kunming, 650224, China.
E-mail address: boyciana@163.com (Y. Duan).
[1] These authors contributed to the worke qually and should be regarded as co-first authors.

https://doi.org/10.1016/j.heliyon.2024.e30513
Received 20 April 2024; Accepted 29 April 2024
Available online 30 April 2024

conserved than nuclear genes, especially in birds [3]. The mitochondrial genome has become one of the most widely used sources of molecular data in taxonomic, population genetic, evolutionary, and phylogenetic studies [4].

The mitochondrial genome of birds is a closed circular structure ranging in length from 16,003 bp – 23,500 bp [5]. In most vertebrates, it comprises a heavy stand (H strand) and a light strand (L strand) that together to form double-stranded circular DNA, with a total of 37 genes, including 13 PCGs, 22 tRNAs, 2 rRNAs, and one or two non-coding control regions (D-loops) [6,7]. Of the 37 genes, the genes encoded on the L strand are NAD6 and 8 tRNA genes ($tRNA^{Ala}$, $tRNA^{Cys}$, $tRNA^{Glu}$, $tRNA^{Asn}$, $tRNA^{Pro}$, $tRNA^{Gln}$, $tRNA^{Ser2(UCN)}$, and $tRNA^{Tyr}$), the rest of the genes (14 tRNAs, 13 PCGs, 2 rRNAs, and control region) are located in the H strand [8,9].

Nuthatches (*Sitta*) are the most speciose genus in the family Sittidae. There are 29 species of birds in *Sitta*, including 11 in China, according to HBW-Birdlife [10]. Nuthatches in the genus *Sitta* are forest-dwelling passerines, with the exception of the rock nuthatches (*S. neumayer* and *S. tephronota*); their diversity hotspots are located in the southwestern China-Himalayas and the adjacent Indo-Burma Mountains, which features many endemic species with narrow distributions [11]. Due to limited sampling, most studies of Sittidae birds have focused on the phylogeography of individual Asian species. However, the phylogenetic relationships among Sittidae species remain unclear. For example, *Sitta* has been suggested to be a member of Sylvioidea [12]. However, *Sitta* has also been suggested to be classified in Certhioidea [13]. Ericson et al. found that Polioptilidae and Troglodytidae are closely related and that *Sitta* is the closest relative of the clade containing Polioptilidae and Troglodytidae [14]. Treplin et al. suggested that Certhiidae and Troglodytidae are sister groups, and *Sitta* is the closest relative to the clade containing these two groups [15]. *Sitta* is closely related to the genus *Campylorhynchus* [16]. Sittidae has also been suggested to be sister to the family Certhiidae [17].

The Giant Nuthatch (*S. magna*) belongs to the order Passeriformes and the family Sittidae. The *S. magna* is a large-sized *Sitta* with grey, black, and white colors. Its rump is chestnut-colored, and its tail is relatively long. Its body markings resemble those of the *S. nagaensis*, but its body is slightly larger. The black eyestripe is much wider on both sides of the head, and the crown is lighter compared to the grey color on the back. In males, there are fine black streaks on the crown. It is mainly distributed in eastern Myanmar and northwest Thailand. Within China, *S. magna* is mainly distributed in Yunnan; they have also been documented in a few locations in Guangxi, Sichuan, and Guizhou, which are adjacent to Yunnan [18]. Their populations are declining due to their narrow distribution and the continuous loss of habitat, the species is thus listed as endangered (EN) on the IUCN Red List [10,19].

Little is known about the complete mitochondrial genome of Sittidae. Here, we present the mitochondrial genome of *S. magna* to aid phylogenetic studies of Sittidae, we also analyzed the nucleotide composition, codon usage, and secondary structure of tRNAs in the *S. magna* mitochondrial genome. The phylogenetic relationships of *Sitta* species were inferred using 13 PCGs.

2. Materials and methods

2.1. Samples and DNA extraction

In this study, blood samples collected from Zixi Mountain Provincial Nature Reserve, Yunnan Province, China, were preserved using absolute ethanol and stored in a deep freezer at −20 °C at the College of Biodiversity Conservation, Southwest Forestry University, Kunming. The TIANamp Genomic DNA Kit (Tiangen, Beijing, China) was used to extract total genomic DNA from blood samples according to the manufacturer's protocol. The DNA purity was checked using a NanoDrop 2000 spectrophotometer (NanoDrop Technologies, Wilmington, DE, USA), while DNA integrity was measured using 1 % agarose gel electrophoresis.

2.2. Genome sequencing, assembly and annotation

The DNA library of *S. magna* was constructed by Nanjing Paisenno Gene Technology Co., Ltd. (Nanjing, China) using the WGS strategy to build the library and next-generation sequencing technology with the Illumina NovaSeq sequencing platform; these libraries were sequenced using double-terminal paired-end sequencing. After quality control, we obtained a total of 20,169,748 high-quality reads, accounting for 80.94 % of the raw reads. A5-miseqv20150522 [20] and SPAdesv3.9.0 [21] were used for assembly, and sequences were extracted according to the sequencing depth of the spliced sequence; sequences with high sequencing depth were compared with the nt library on NCBI via BLASTN searches (BLAST v2.2.31+), and the mitochondrial sequences of each splicing result were retained. The collinearity of the mitochondrial splicing results was analyzed using mummer v3.1 [22] software to determine the relative positions of the contigs and fill the gaps between contigs. The final mitochondrial genome sequence was obtained after applying corrections using Pilon v1.18 [23], and the mitogenome or mitochondrial genome sequence of *S. magna* was submitted in the GenBank (Accession number: MZ888773). The mitochondrial genome sequence was annotated using MITOS webserver (http://mitos.bioinf.uni-leipzig.de/) for prediction of the putative secondary structure tRNA genes [24]. Mitochondrial whole genome circle mapping was conducted using cgview visualization software [25]. Base composition, codon usage, and relative synonymous codon usage (RSCU) were calculated using MEGA [7], and the base bias of mitochondrial genes was calculated using the following formulas: AT-skew = (A − T)/(A + T) and GC-skew = (G − C)/(G + C) [26].

2.3. Phylogenetic analyses

Currently, only a small number of complete mitochondrial genomes for Sittidae species are available from the NCBI database. To determine the phylogenetic relationships of *S. magna* with other *Sitta* species, we downloaded 16 mitochondrial genome sequences from NCBI, including 16 species in six orders. A total of 13 PCGs identified in *S. magna* were analyzed using the PCGs of passerine birds obtained from the NCBI database, and *Aethopyga gouldiaein* in the Aectariniidae family was used as an outgroup. Table 1 shows the

sequence information.

The nucleotide sequences of 13 PCGs in the 17 mitochondrial genomes were aligned using Clustal X in MEGA v. 7.0; tandem and multiple seuqence alignment was performed using default parameters [34].The sequence alignment was performed in MEGA v.7.0, and saved in both Fasta and Mega formats. Sequence saturation analysis was conducted using DAMBE 5.2.63 software [35]. If Iss < Iss. c and the relationship between the two significantly differed, the nucleotide sequences were not saturated, which indicates that these sequences be used to construct a phylogenetic tree is meaningful; nucleotide sequences were unsaturated. Phylogenetic analysis was performed using Bayesian inference (BI) and maximum likelihood (ML) [36]. According to the Akaike Information Criterion [37], we used jModelTest 0.1.1 to calculate parameters such as the number of nucleotide surrogate types (nst) and the rate distribution between sites (rates); the GTR + I + G model was the optimal model. ML analysis was performed using RaxML GUI 1.3.1, and BI was conducted using MrBayes 3.2.1 with four independent chains running for two million generations. Trees were sampled every 100 generations; the first 2500 trees were removed as burn-in, and the remaining trees were used to calculate Bayesian posterior probabilities. Phylogenetic trees were viewed and edited in Figtree 1.4.0 [38].

3. Results

3.1. Genome organization and nucleotide composition

The mitochondrial genome of *S. magna* is a double-stranded closed circular molecule with a length of 16,829 bp consisting of 37 coding genes and 1 non-coding control region. The 37 coding genes were 13 PCGs, 2 rRNAs, and 22 tRNAs, and the structure of the mitochondrial genome is shown in Fig. 1. Of the 37 genes, 28 were located in the H stand (14 tRNAs, 2 rRNAs, and 12 PCGs), and the remaining 9 genes were located in the L stand (8 tRNAs and 1 PCG) (Table 2). The mitochondrial genome of the *S. magna* contains 19 intergenic spacers, ranging from 1 to 12 bp in length, with a total length of 110 bp. Among them, two intergenic spacers exceed 10 bp, located between atp6 and cox3 (12 bp), and between nad5 and cob (11 bp). There are a total of 10 gene overlaps, ranging from 1 to 10 bp in length, with a total length of 44 bp. Among them, the relatively longer overlaps occur between atp8 and atp6 (10 bp), between trnY and cox1 (8 bp), and between cox1 and trnS2 (9 bp). The remaining 8 gene pairs have no intergenic spacers or overlaps (Table 2).

The nucleotide composition of the complete mitochondrial genome is shown in Table 3: A = 5384 (30.29 %), T = 4109 (23.12 %), G = 2595 (14.60 %), and C = 5684 (31.98 %). The complete mitochondrial DNA sequence of the *S. magna* is similar to those in other *Sitta* mitochondrial genomes. In the mitochondrial genome of the *S. magna*, the A + T content is 54.11 %, showing a pronounced A + T bias. The A-T skew values for the complete mitochondrial sequence, PCGs, tRNAs, and rRNAs are positive (0.06–0.24), indicating that A is more common than T. The G-C skew values are negative (−0.38 to −0.01), indicating that C is more common than G. In the control region, not only is the G-C skew value negative (−0.24), but the A-T skew value is also negative (−0.16), indicating that in the control region, C is more common than G, and T is more common than A.

3.2. Protein-coding genes and codon usage

The total length of the 13 PCGs of *S. magna* was 11,411 bp, accounting for 67.81 % of the entire mitochondrial genome of *S. magna*. The longest PCG was ND5 (1818 bp), and the shortest was ATP8 (168 bp). Of the 13 PCGs, only the start codon usaged by COI was ATC, and the start codon by the remaining 12 was ATG. The stop codons were TAA, TAG, AGG, and AGA, and TAA was the most common. The stop codon in ND2, COII, ATP8, ATP6, ND3, ND4L, and Cytb was TAA, the stop codon in ND1 and ND6 was TAG, the stop codon in COI and ND4 was AGG, the stop codon in ND5 was AGA, and the incomplete termination (TA-) codon was present in COIII. In addition, there was gene overlap or spacing in 13 PCGs; specifically, ATP8 and ATP6 shared 10 nucleotides, ND4L and ND4 shared 7 nucleotides,

Table 1
List of 17 species used for phylogenetic analyses in this study.

Family	Species	GenBank No.	References
Sittidae	*Sitta magna*	MZ888773	This study
	Sitta nagaensis	MK343427	Unpublished
	Sitta europaea	MN356255	Unpublished
	Sitta himalayensis	MK343426	Unpublished
	Sitta villosa	MT444149	[27]
	Sitta yunnanensis	MN052793	Unpublished
	Sitta carolinensis	NC_024870	[28]
Troglodytidae	*Campylorhynchus brunneicapillus*	NC_029482	Unpublished
	Campylorhynchus zonatus	NC_022840	[29]
	Henicorhina leucosticta	NC_024673	[30]
Polioptilidae	*Polioptila caerulea*	NC_051031	Unpublished
Sturnidae	*Acridotheres cristatellus*	NC_015613	Unpublished
	Sturnus sericeus	NC_014455	[31]
	Gracupica nigricollis	NC_020423	Unpublished
	Sturnus vulgaris	NC_029360	[32]
Mimidae	*Toxostoma redivivum*	MN356247	Unpublished
Nectariniidae	*Aethopyga gouldiae*	NC_027241	[33]

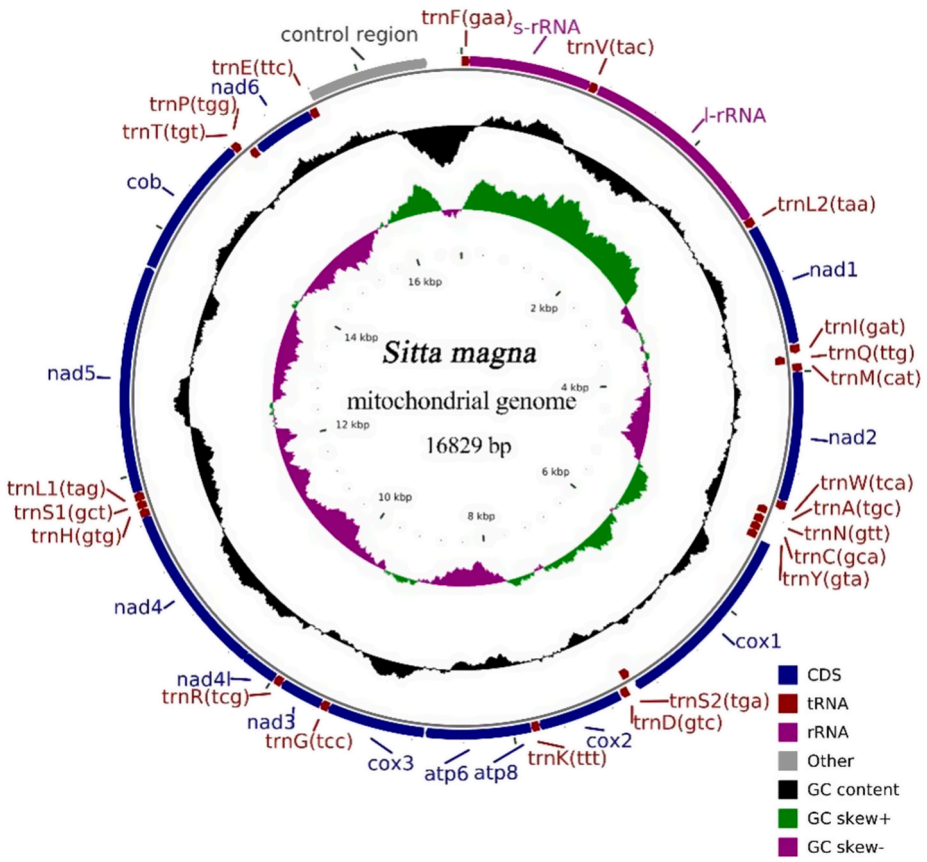

Fig. 1. The circular map of the *S. magna* mitochondrial genome.

ATP6 and COIII shared 1 nucleotide, and ND5 and Cytb shared 11 nucleotides.

RSCU values for the *S. magna* mitochondrial genome for the third position are shown in Table 4. CUA had a maximum RSCU value of 2.83, and AGG had a minimum RSCU value of 0.13. The total number of codons for the 13 PCGs was 3798 the codon encoding Trp was the rarest, and the codon encoding Leu1, Ile, and Thr was the most common (Fig. 2). In the ND6 gene, A was more common than T (−0.51) and C was more common than G (0.56); in the remaining 12 PCGs, T was slightly more common than A (AT-skew = 0.06 to 0.18) and G was more common than C (GC-skew = −0.28 to −0.66) (Table 5).

3.3. Ribosomal and transfer RNA genes

The two rRNA genes were 12S rRNA and 16S rRNA. The 12S rRNA and 16S rRNA were 978 bp and 1591 bp in length, respectively. As in most vertebrates, two rRNAs were isolated by the $tRNA^{Val}$ between $tRNA^{Phe}$ and $tRNA^{Leu2\ (UUR)}$, with A + T levels of 50.82 % and 56.13 %, respectively, and A + T levels were significantly higher than G + C levels (Table 3).

The total length of the 22 tRNAs in the mitochondrial genome of *S. magna* was 1544 bp, and the gene length of a single tRNA ranged from 66 bp ($tRNA^{Ser1\ (AGN))}$) to 75 bp ($tRNA^{Leu2\ (UUR)}$ and $tRNA^{Ser2(UCN)}$). The A + T content of the 22 tRNAs was 58.03 %, which was between 46.38 % ($tRNA^{Met}$) and 68.12 % ($tRNA^{Thr}$). All tRNAs formed a typical clover structure. Among the 22 tRNAs, only 6 tRNAs did not have base mismatches, and the remaining 16 tRNAs had a total of 30 pairs of base mismatches, including G-U, A-A, A-C, U–U, and C–C mismatches, and the most frequent mismatches were G-U mismatches (22 pairs) (Fig. 3).

3.4. Non-coding regions

A putative control region (CR) with a length of 966 bp was observed in the mitochondrial genome of *S. magna*; it was located between $tRNA^{Phe}$ and $tRNA^{Glu}$ and encoded by the H stand. The base composition of CR was as follows: A, 22.88 %; T, 31.78 %; G, 17.18 %; and C, 28.16 %. The content of A + T (54.66 %) was significantly higher than that of G + C, and both AT-skew and GC-skew were observed; specifically, A was more common than T (AT-skew = −0.16), and G was more common than C (GC-skew = −0.24).

Table 2
The Details of the genetic organization of the newly sequenced *S. magna* mitochondrial genome.

Genes	Position	Length (bp)	Start codon	Stop codon	Anticodon	Intergenic nucleotide (bp)	strand
trnF	1–68	68			GAA	−1	H
rrnS	68–1045	978				−1	H
trnV	1045–1113	69			TAC	7	H
rrnL	1121–2711	1591				0	H
trnL2	2712–2786	75			TAA	10	H
nad1	2797–3774	978	ATG	TAG		7	H
trnI	3782–3853	72			GAT	6	H
trnQ	3860–3930	71			TTG	3	L
trnM	3934–4002	69			CAT	0	H
nad2	4003–5043	1041	ATG	TAA		0	H
trnW	5044–5113	70			TCA	1	H
trnA	5115–5183	69			TGC	10	L
trnN	5194–5266	73			GTT	2	L
trnC	5269–5335	67			GCA	−1	L
trnY	5335–5405	71			GTA	−8	L
cox1	5398–6957	1560	ATC	AGG		−9	H
trnS2	6949–7023	75			TGA	5	L
trnD	7029–7097	69			GTC	10	H
cox2	7108–7791	684	ATG	TAA		1	H
trnK	7793–7860	68			TTT	1	H
atp8	7865–8032	168	ATG	TAA		−10	H
atp6	8020–8703	684	ATG	TAA		12	H
cox3	8716–9500	785	ATG	TA(A)		−1	H
trnG	9500–9568	69			TCC	0	H
nad3	9569–9919	351	ATG	TAA		0	H
trnR	9920–9989	70			TCG	1	H
nad4l	9991–10287	297	ATG	TAA		−7	H
nad4	10281–11663	1383	ATG	AGG		−5	H
trnH	11659–11728	70			GTG	0	H
trnS1	11729–11794	66			GCT	−1	H
trnL1	11794–11864	71			TAG	0	H
nad5	11865–13682	1818	ATG	AGA		11	H
cob	13694–14836	1143	ATG	TAA		3	H
trnT	14840–14908	69			TGT	7	H
trnP	14916–14985	70			TGG	10	L
nad6	14996–15514	519	ATG	TAG		0	L
trnE	15515–15587	73			TTC	3	L
CR	15591–16556	966				272	H

Table 3
Composition and skewness values of *S. magna*.

S. magna	Size (bp)	A%	T%	G%	C%	A + T%	G + C%	AT skew	GC skew
mDNA	16829	30.41	23.70	14.49	31.40	54.11	45.89	0.12	−0.37
PCGs	11411	28.26	24.95	14.56	32.23	53.21	46.79	0.06	−0.38
tRNA	1544	30.76	27.27	20.73	21.24	58.03	41.97	0.06	−0.01
12s rRNA	978	30.06	20.76	21.47	27.71	50.82	49.18	0.18	−0.13
16s RNA	1591	34.82	21.31	19.67	24.20	56.13	43.87	0.24	−0.10
CR	966	22.88	31.78	17.18	28.16	54.66	45.34	−0.16	−0.24

3.5. Phylogenetic analyses

To clarify the phylogenetic relationship of *S. magna* with other passerines, the PCGs of 17 species of passerine birds were aligned. BI and ML analysis were carried out; the topologies were all highly supported and consistent (Fig. 4). The 17 species belonged to six families: Trotteridae, Troglodytidae, Polioptilidae, Sturnidae, Mimidae, and Nectariniidae. *Aethopyga gouldiae* was in the Nectariniidae family and was used as an outgroup [39].

Phylogenetic analysis suggested that Sittidae is sister to Troglodytidae and Polioptilidae (posterior probability: 1.00; bootstrap support: 100 %), indicating that Sturnidae and Mimidae are sister groups (posterior probability: 1.00; bootstrap support: 100 %). Seven species of birds in the genus *Sitta* formed a monophyletic clade with high support (posterior probability: 1.00; bootstrap support: 100 %), and the target species *S. magna* and *S. carolinensis* formed a sister group. *S. nagaensis* and *S. europae* were sister to *S. himalayensis*, and *S. villosa* and *S. yunnanensis* comprised a sister group (posterior probability: 1.00; bootstrap support: 100 %).

Table 4
Codon number and relative synonymous codon usage (RSCU) of *S. magna* mitochondrial protein-coding genes (PCGs).

Codon	Count	RSCU	Codon	Count	RSCU	Codon	Count	RSCU	Codon	Count	RSCU
UUU(F)	43	0.47	UCU(S)	50	0.89	UAU(Y)	39	0.64	UGU(C)	8	0.48
UUC(F)	141	1.53	UCC(S)	119	2.13	UAC(Y)	83	1.36	UGC(C)	25	1.52
UUA(L)	51	0.53	UCA(S)	97	1.73	UAA(*)	30	0.61	UGA(*)	86	1.76
UUG(L)	22	0.23	UCG(S)	14	0.25	UAG(*)	31	0.63	UGG(W)	6	1
CUU(L)	62	0.65	CCU(P)	68	0.96	CAU(H)	40	0.66	CGU(R)	8	0.52
CUC(L)	131	1.37	CCC(P)	81	1.14	CAC(H)	82	1.34	CGC(R)	16	1.03
CUA(L)	270	2.83	CCA(P)	119	1.68	CAA(Q)	84	1.46	CGA(R)	45	2.9
CUG(L)	37	0.39	CCG(P)	16	0.23	CAG(Q)	31	0.54	CGG(R)	5	0.32
AUU(I)	74	0.59	ACU(T)	53	0.68	AAU(N)	36	0.5	AGU(S)	11	0.2
AUC(I)	189	1.51	ACC(T)	135	1.73	AAC(N)	109	1.5	AGC(S)	45	0.8
AUA(I)	112	0.9	ACA(T)	114	1.46	AAA(K)	83	1.55	AGA(R)	8	0.52
AUG(M)	38	1	ACG(T)	10	0.13	AAG(K)	24	0.45	AGG(R)	11	0.71
GUU(V)	31	0.73	GCU(A)	49	0.69	GAU(D)	19	0.52	GGU(G)	21	0.43
GUC(V)	48	1.13	GCC(A)	141	1.98	GAC(D)	54	1.48	GGC(G)	55	1.12
GUA(V)	70	1.65	GCA(A)	88	1.24	GAA(E)	53	1.29	GGA(G)	78	1.59
GUG(V)	21	0.49	GCG(A)	7	0.1	GAG(E)	29	0.71	GGG(G)	42	0.86

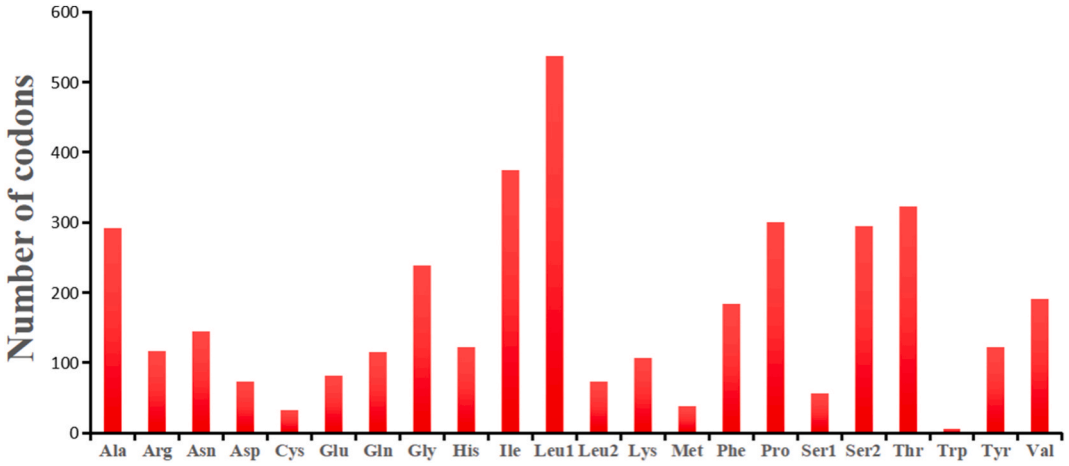

Fig. 2. Relative synonymous Codon Usage (RSCU) in *S. magna*.

Table 5
Base composition for protein-coding genes in the mitochondrial genome of *S. magna*.

Gene	Length (bp)	Proportion of nucleotides (%)					AT skew	GC skew
		A	T	G	C	A + T		
nad1	978	28.43	24.43	13.60	33.54	52.86	0.08	−0.42
nad2	1041	30.45	23.44	11.05	35.06	53.89	0.13	−0.52
cox1	1560	27.88	24.49	17.05	30.58	52.37	0.06	−0.28
cox2	684	30.85	23.24	14.77	31.14	54.09	0.14	−0.36
atp8	168	33.93	23.81	7.14	35.12	57.74	0.18	−0.66
atp6	684	31.58	25.15	10.53	32.74	56.73	0.11	−0.51
cox3	785	27.01	23.56	16.31	33.12	50.57	0.07	−0.34
nad3	351	28.49	24.22	12.82	34.47	52.71	0.08	−0.46
nad4l	297	26.94	22.56	11.78	38.72	49.50	0.09	−0.53
nad4	1383	29.14	23.72	10.99	36.15	52.86	0.10	−0.53
nad5	1818	31.74	23.59	11.72	32.95	55.33	0.15	−0.48
cob	1143	28.34	23.80	13.21	34.65	52.14	0.09	−0.45
nad6	519	12.53	38.34	38.34	10.79	50.87	−0.51	0.56

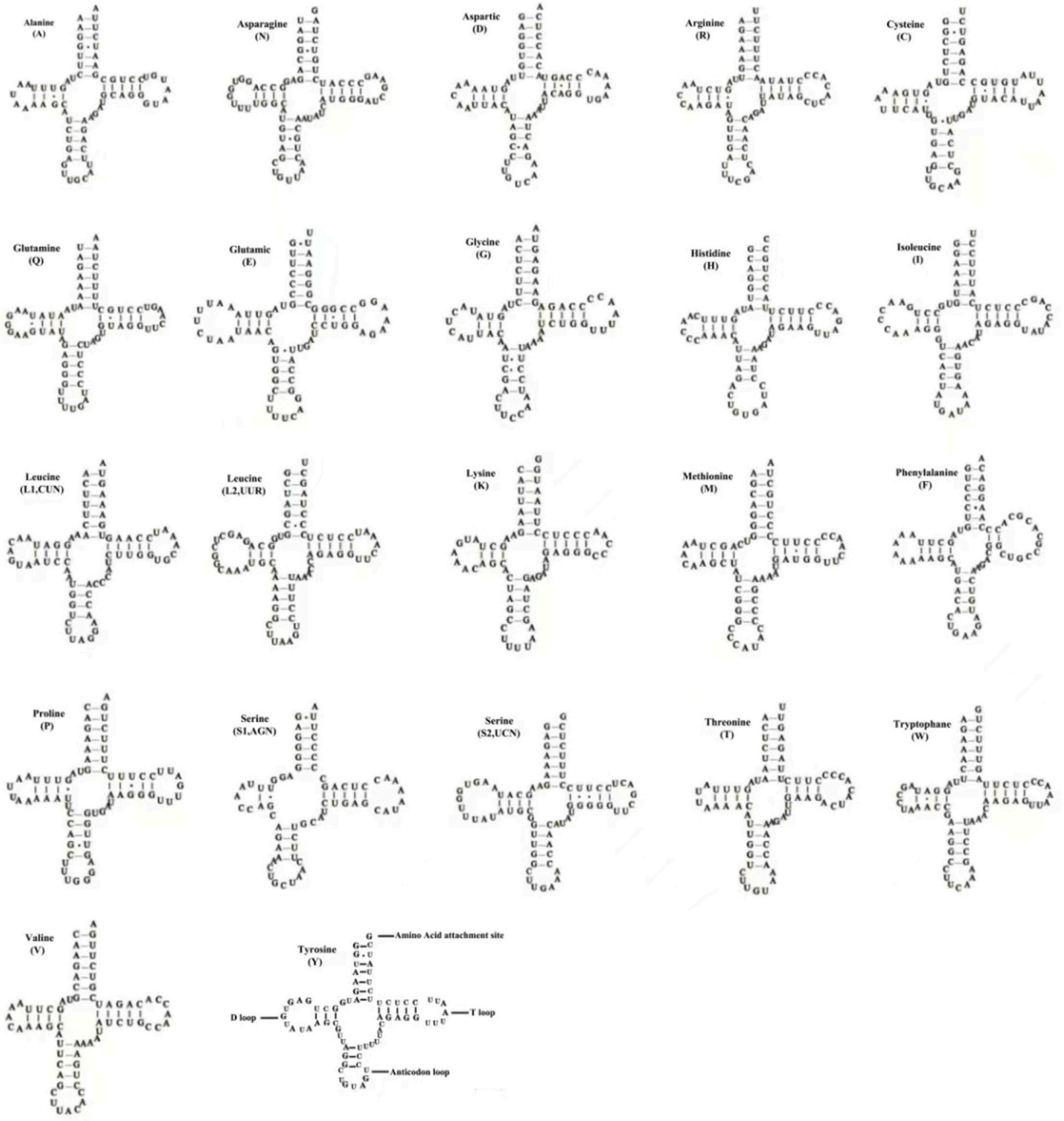

Fig. 3. Secondary structures of the 22 transfer RNA genes of *S.magna*.

4. Discussion

4.1. Characteristics of the mitochondrial genome

In this study, the complete mitochondrial genome of *S. magna* was reported for the first time; the complete length of the genome was 16,829 bp. Most bird mitochondrial genomes that have been sequenced range from 16,300 bp to 23,500 bp in length [40]. The mitochondria genomes of all *Sitta* species contained a large D-loop region and 37 coding genes (22 tRNAs, 13 PCGs, and 2 rRNAs); most genes were located in the H stand, and only 8 tRNAs and ND6 were located in the L stand [2,27]. The content of A + T (54.11 %) was significantly higher than that of G + C (45.89 %), and this is consistent with the typical base bias observed in vertebrate genomes [38].

A total of four complete stop codons (TAG, TAA, AGG, AGA) and one incomplete stop codon (TA) were detected in the 13 PCGs; the

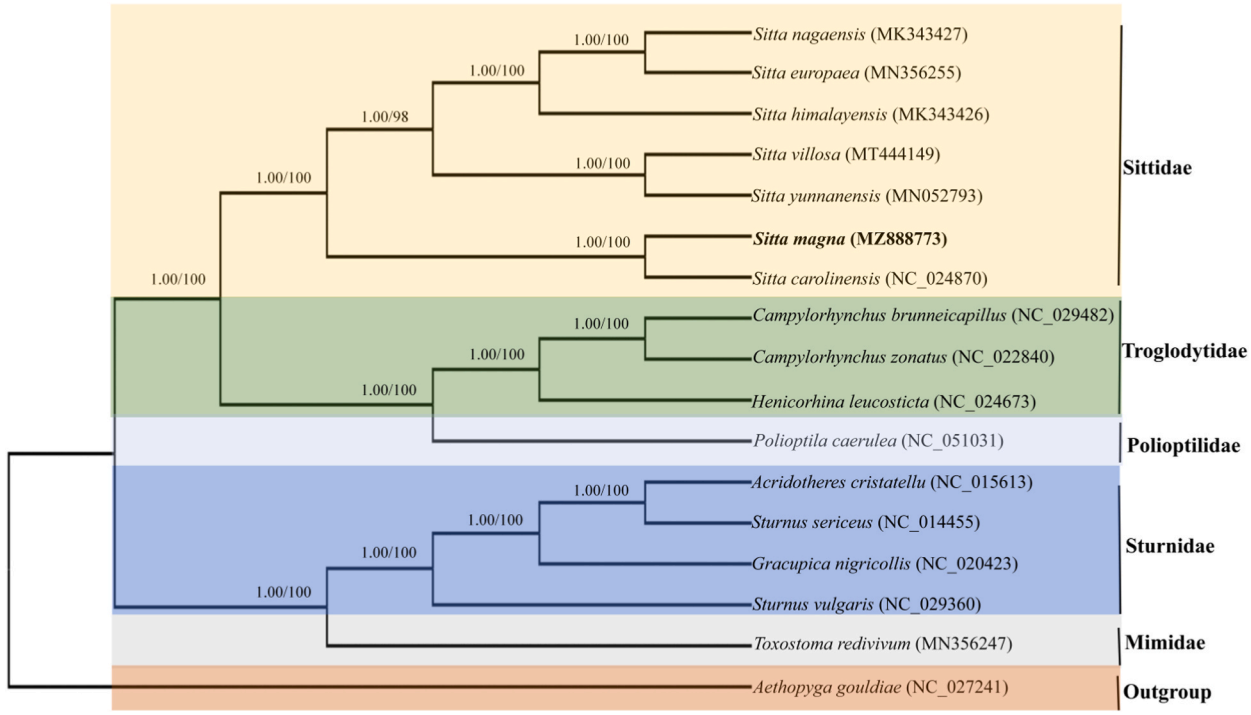

Fig. 4. The phylogenetic tree of *S. magna* based on 13 protein-coding genes; the topology was generated using BI (Numbers on each branch refer to the BI/ML support values)

incomplete stop codon was probably completed by polyadenylation and finally became the stop codon TAA [41]. Codons encoding Leu1 were the most common, and codons encoding Trp were the least common; this might be related to the function of mitochondria [38]. Some genes were separated by 1–12 bp or overlapped by 1–10 bp, whereas others were closely arranged. This stems from the compact nature of the mitochondrial genome [42]. Gene overlap is commonly observed in higher eukaryotes; overlapping genes have coding sequences and regulatory sequences, and gene overlap can conserve bases. In addition, base overlap permits more information to be encoded, which enhances the utilization of the genetic information in DNA and might play a role in the regulation of gene expression [43].

4.2. Phylogenetic analyses

In vertebrates, mitochondrial genomes evolve faster than nuclear genes; gene structure and the recombination rate are also more stable in mitochondrial genomes than in the nuclear genome. Therefore, mitochondrial genome sequences are widely used to infer phylogenetic relationships [1,4]. In this study, the phylogenetic relationships of *S. magna* with 17 other passerine bird species were investigated. The well-supported phylogenetic trees generated by BI and ML were consistent (Fig. 4).

The taxonomic status of the Sittidae family remains unclear. The results provide strong support for the close relationship between the taxa comprising Troglodytidae and Polioptilidae (posterior probability: 1.00; bootstrap support: 100 %) and confirmed that Sittidae was sister to the clade containing Troglodytidae and Polioptilidae (Sittidae + (Troglodytidae + Polioptilidae)). This is consistent with the findings of Ericson et al. [14]. *S. magna* is a globally endangered species. Few studies of *S. magna* have been conducted due to limitations in sampling. The phylogenetic analysis in this study showed that *S. magna* is closely related to *S. carolinensis*. This is consistent with the findings of Pasquet et al. [44] and Päckert et al. [11]. These findings indicate that Sturnidae and Mimidae comprise a highly supported clade, which is consistent with the results of Tobias et al. [39].

Sitta is the most speciose genus in the passerine family Sittidae; according to current systematics studies, this family comprises two additional genera in the subfamily rank: the monotypic wallcreeper (*Tichodroma muraria*; Tichodrominae) and spotted creepers (*Salpornis,* Salpornithinae) [11]. However, there remains much uncertainty regarding the phylogenetic relationships of the Sittidae based on mitochondrial genomes in recent years, and this might stem from limitations in mitochondrial genome data; more mitochondrial genomes need to be sequenced to resolve this uncertainty.

5. Conclusions

We sequenced the whole mitochondrial genome of *S. magna*. The mitochondrial genome of *S. magna* was similar to that of other Sitta species. It was a double-stranded closed-ring molecule with a length of 16,829 bp. It comprised 37 coding genes and 1 non-coding

control region. The 37 coding genes included 13 PCGs, 2 rRNAs, and 22 tRNAs. Phylogenetic analysis indicated that *S. magna* is a sister group of *S. carolinensis*, which is consistent with the results of previous studies.

Funding

This work was supported by the Science and Technology Project of Yunnan (202301BD070001-082 and 202101AT070040) and the Yunnan Provincial University Student Innovation and Entrepreneurship Training Program Support Project (202010677023).

Institutional review board statement

All samples collected in this study were non-invasive sampling. All animal experiments were approved by the Academic Committee of Southwest Forestry University (IACUC-SWFU L20200206), which includes some regulations on animal ethics, animal welfare, and wildlife conservation. And all methods were performed in accordance with the Guidelines for the ethical review of laboratory animal welfare People's Republic of China National Standard (GB/T 35892-2018) and the study complies with the ARRIVE guidelines (https://arriveguidelines.org).

Informed consent statement

Not applicable.

Data availability statement

The mitogenome sequence of *S. magna* has been submitted to GenBank under accession number MZ888773.

CRediT authorship contribution statement

Ruixin Mo: Methodology, Writing – original draft. **Dong Zhu:** Conceptualization, Investigation, Methodology, Software. **Jing Sun:** Methodology, Writing – original draft. **Qingmiao Yuan:** Software. **Feng Guo:** Investigation. **Yubao Duan:** Conceptualization, Resources, Writing – review & editing.

Declaration of competing interest

The author(s) declared no potential conflicts of interest with respect to the research, author-ship, and/or publication of this article.

Acknowledgments

There are grateful to Zezhao Kang for assistance with the fieldwork. We also thank the reviewers for their valuable comments and suggestions to improve the manuscript.

References

[1] S.C. Chen, X.Q. Wang, P.W. Li, X. Hu, J.J. Wang, P. Peng, The complete mitochondrial genome of aleurocanthus camelliae: insights into gene arrangement and genome organization within the family aleyrodidae, Int. J. Mol. Sci. 17 (11) (2016) 1843.
[2] Q.M. Yuan, Q. Guo, J. Cao, X. Luo, Y.B. Duan, Description of the three complete mitochondrial genomes of *Sitta* (S. Himalayensis, S. Nagaensis, and S. Yunnanensis) and phylogenetic relationship (aves: Sittidae), Genes 14 (3) (2023) 589.
[3] F.E. Rheindt, S.V. Edwards, Genetic introgression: an integral but neglected component of speciation in birds, The Auk 128 (4) (2011) 620–632.
[4] Z.T. Chen, Y.Z. Du, A.S. Bustin, First mitochondrial genome from nemouridae (plecoptera) reveals novel features of the elongated control region and phylogenetic implications, Int. J. Mol. Sci. 18 (5) (2017) 996.
[5] W. Bai, Mitochondrial Genome and Phylogenetic Relationship of Phasianidae Birds, Shanxi University, 2020.
[6] D. Bi, NUMT and Mitochondrial Phylogenetic Genomics in Passeriformes, Anhui Normal University, 2019.
[7] Y.X. Huang, F.J. Ren, R.C. Bartlett, Y.S. Wei, D.Z. Qin, Contribution to the mitogenome diversity in Delphacinae: phylogenetic and ecological implications, Genomics 112 (2) (2020) 1363–1370.
[8] C.H. Sun, H.Y. Liu, X. Min, C.H. Lu, Mitogenome of the little owl *Athene noctua* and phylogenetic analysis of Strigidae, Int. J. Biol. Macromol. 151 (2020) 924–931.
[9] Q.M. Yuan, J.B. Sha, Y.B. Duan, The new mitogenome of erpornis zantholeuca (aves: Passeriformes): sequence, structure, and phylogenetic analyses, Cytogenet. Genome Res. 162 (5) (2022) 250–261.
[10] HBW-BirdLife, Handbook of the Birds of the World and Birdlife, 2022. Version 2022-12, http://datazone.birdlife.org. (Accessed 20 April 2023).
[11] M. Päckert, M. Bader-Blukott, B. Künzelmann, Y.H. Sun, Y.C. Hsu, C. Kehlmaier, F. Albrecht, J.C. Illera, J. Martens, A revised phylogeny of nuthatches (Aves, Passeriformes, *Sitta*) reveals insight in intra- and interspecific diversification patterns in the Palearctic, Vertebrate Zoology 70 (2) (2020) 241–262.
[12] F.H. Sheldon, F.B. Gill, A reconsideration of songbird phylogeny, with emphasis on titmice and their sylvioid relatives, Syst. Biol. 45 (4) (1996) 473–495.
[13] U.S. Johansson, J. Fjeldså, R.C. Bowie, Phylogenetic relationships within Passerida (Aves: Passeriformes): a review and a new molecular phylogeny based on three nuclear intron markers, Mol. Phylogenet. Evol. 48 (3) (2008) 858–876.
[14] G.P. Ericson, S.U. Johansson, Phylogeny of Passerida (Aves: Passeriformes) based on nuclear and mitochondrial sequence data, Mol. Phylogenet. Evol. 29 (1) (2003) 126–138.
[15] S. Treplin, R. Siegert, C. Bleidorn, H.S. Thompson, R. Fotso, R. Tiedemann, Molecular phylogeny of songbirds (Aves: Passeriformes) and the relative utility of common nuclear marker loci, Cladistics 24 (3) (2008) 328–349.
[16] K.F. Barker, Mitogenomic data resolve basal relationships among passeriform and passeridan birds, Mol. Phylogenet. Evol. 79 (2014) 313–324.

[17] V.A. Payevsky, Phylogeny and classification of passerine birds, Passeriformes. Biology Bulletin Reviews 4 (2) (2014) 143–156.
[18] J. Mackinnon, K. Philipps, A Field Guild to the Birds of China, Hunan Education Press, 2000.
[19] BirdLife International, Species factsheet: tetrax tetrax, in: IUCN Red List for Birds, 2022. Downloaded from, http://www.birdlife.org. (Accessed 4 October 2022).
[20] D. Coil, G. Jospin, A.E. Darling, A5-miseq: an updated pipeline to assemble microbial genomes from Illumina MiSeq data, Bioinformatics 31 (4) (2015) 587–589.
[21] A. Bankevich, S. Nurk, D. Antipov, A.A. Gurevich, M. Dvorkin, A.S. Kulikov, V.M. Lesin, S.I. Nikolenko, S. Pham, A.D. Prjibelski, A.V. Pyshkin, A.V. Sirotkin, N. Vyahhi, G. Tesler, M.A. Alekseyev, P.A. Pevzner, SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing, J. Comput. Biol. : a journal of computational molecular cell biology 19 (5) (2012) 455–477.
[22] S. Kurtz, A. Phillippy, A.L. Delcher, M. Smoot, M. Shumway, C. Antonescu, S. Salzberg, Versatile and open software for comparing large genomes, Genome Biol. 5 (2) (2004) R12.
[23] B.J. Walker, T. Abeel, T. Shea, M. Priest, A. Abouelliel, S. Sakthikumar, C.A. Cuomo, Q.D. Zeng, J. Wortman, S.K. Young, A.M. Earl, Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement, PLoS One 9 (11) (2014) e112963.
[24] M. Bernt, A. Donath, F. Juhling, F. Externbrink, C. Florentz, G. Fritzsch, J. Putz, M. Middendorf, Stadler PF: MITOS: improved de novo metazoan mitochondrial genome annotation, Mol. Phylogenet. Evol. 69 (2) (2013) 313–319.
[25] N.F. Alinkhan, N.K. Petty, Z.N. Ben, S.A. Beatson, BLAST Ring Image Generator (BRIG): simple prokaryote genome comparisons, BMC Genom. 12 (1) (2011) 402.
[26] H.Y. Liu, C.H. Sun, Y. Zhu, Q.Z. Zhang, Complete mitogenomic and phylogenetic characteristics of the speckled wood-pigeon (*Columba hodgsonii*), Mol. Biol. Rep. 47 (5) (2020) 3567–3576.
[27] Z. Zhang, S. Mi, Q. Guo, Z. Zhang, P. Yan, Z. Liu, L. Teng, The complete mitochondrial genome of the *Sitta villosa* (Passeriformes: Sittidae) from China, Mitochondrial DNA part B 5 (3) (2020) 2328–2329.
[28] F.K. Barker, Mitogenomic data resolve basal relationships among passeriform and passeridan birds, Mol. Phylogenet. Evol. 79 (2014) 313–324.
[29] F.K. Barker, S. Oyler-McCance, D.F. Tomback, Blood from a turnip: tissue origin of low-coverage shotgun sequencing libraries affects recovery of mitogenome sequences, Mitochondrial DNA 26 (3) (2015) 384–388.
[30] C. Aguilar, L.F.D. Léon, J.R. Loaiza, W.O. McMillan, M.J. Miller, Extreme sequence divergence between mitochondrial genomes of two subspecies of White-breasted Wood-wren (*Henicorhina leucosticte*, Cabanis, 1847) from western and central Panamá, Mitochondrial DNA 27 (2) (2016) 956–957.
[31] C. Qian, Q. Ren, X. Kan, Z. Guo, J. Yang, X. Li, J. Yuan, M. Qian, Q. Hu, L. Zhang, Complete mitochondrial genome of the Red-billed Starling, *Sturnus sericeus* (Aves: Passeriformes): the first representative of the family Sturnidae with a single control region, Mitochondrial DNA 24 (2) (2013) 129–131.
[32] L.A. Rollins, A.P. Woolnough, B.G. Fanson, M.L. Cummins, T.M. Crowley, A.N. Wilton, R. Sinclair, A. Butler, W.B. Sherwin, Selection on mitochondrial variants occurs between and within individuals in an expanding invasion, Mol. Biol. Evol. 33 (4) (2016) 995–1007.
[33] N. Wang, B. Liang, Complete mitochondrial genome of a sunbird, Aethopyga gouldiae (Aves: Passeriformes), the first representative of Nectariniidae, Mitochondrial DNA 27 (4) (2015) 2356–2358.
[34] M.A. Larkin, G. Blackshields, N.P. Brown, R. Chenna, P.A. McGettigan, H. McWilliam, F. Valentin, I.M. Wallace, A. Wilm, R. Lopez, J.D. Thompson, T.J. Gibson, D.G. Higgins, ClustalW and ClustalX version 2.0, Bioinformatics 23 (21) (2007) 2947–2948.
[35] X. Xia, DAMBE5: a comprehensive software package for data analysis in molecular biology and evolution, Mol. Biol. Evol. 30 (7) (2013) 1720–1728.
[36] A. Cibois, M. Gelang, P. Alström, E. Pasquet, J. Fjeldså, P.G.P. Ericson, U. Olsson, Comprehensive phylogeny of the laughingthrushes and allies (Aves, Leiothrichidae) and a proposal for a revised taxonomy, Zool. Scripta 47 (4) (2018) 428–440.
[37] D. Posada, ModelTest: phylogenetic model averaging, Mol. Biol. Evol. 25 (7) (2008) 1253–1256.
[38] C.H. Sun, H.Y. Liu, C.H. Lu, Five new mitogenomes of Phylloscopus (Passeriformes, Phylloscopidae):Sequence, structure, and phylogenetic analyses, Int. J. Biol. Macromol. 146 (2020) 638–647.
[39] J.A. Tobias, C. Sheard, A.l. Pigot, et al., AVONET: morphological, ecological and geographical data for all birds, Ecol. Lett. 25 (3) (2022) 581–597.
[40] Y.K. Gao, Y.W. Miao, X.Q. Su, Z.F. Chi, B. Yu, F. Jiang, A comprehensive analysis on 74 avian mitochondrial genome base compositions, J. Yunnan Agric. Univ. 24 (1) (2009) 54–58.
[41] D. Ojala, J. Montoya, G. Attardi, TRNA punctuation model of RNA processing inhuman mitochondria, Nature 290 (5806) (1981) 470–474.
[42] J.P. Curole, T.D. Kocher, Mitogenomics: digging deeper with complete mitochondrial genomes, Trends Ecol. Evol. 14 (10) (1999) 394–398.
[43] B. Xiao, F. Ma, Y. Sun, Q.W. Li, Comparative analysis of complete mitochondrial DNA control region of four species of strigiformes, Acta Genetica Sin. 33 (11) (2006) 965–974.
[44] E. Pasquet, F.K. Barker, J. Martens, A. Tillier, C. Cruaud, A. Cibois, Evolution within the nuthatches (Sittidae: aves, Passeriformes): molecular phylogeny, biogeography, and ecological perspectives, J. Ornithol. 155 (3) (2014) 755–765.