Sequencing and phylogenetic analysis of the *Pyrgilauda ruficollis* (Aves, Passeridae) complete mitochondrial genome

Yong-Gui MA [1,2], Yuan HUANG [1,*] , Fu-Min LEI [3]

1. *School of Life Sciences*, *Shaanxi Normal University*, *Xi'an* 710062, *China*
2. *Ministry of Education Key Laboratory of Qinghai-Tibet Plateau Resources and Environment*, *Qinghai Normal University*, *Xining* 810008, *China*
3. *Key Laboratory of the Zoological Systematics and Evolution*, *Institute of Zoology*, *Chinese Academy of Sciences*, *Beijing* 100101, *China*

Abstract: In this study, both long PCR and conserved primers walking sequencing methods were used to determine the complete sequence of the of *Pyrgilauda ruficollis* mitochondrial genome (KC836121). The results showed that the complete mitochondrial genome of *P. ruficollis* is 16909 bp in length with 55.0% A+T content, harboring the typical 37 genes. The mitogenome had the same gene order with that of *Podoces hendersoni*. All protein coding genes started with ATG codon, except *ND3* with GTG. For the stop codon usage, most genes terminate with codons TAA or TAG, but *ND5* terminated with AGA, while *ND1* and *COI* genes with AGG, and both the genes *COIII* and *ND4* have an incomplete termination codon (T). The secondary structures of 22 tRNA genes were also predicted, showing that all tRNAs can form typical clover-leaf secondary structures, except for the $tRNA^{Ser}$ (AGN) which loses the DHU arm, while $tRNA^{Phe}$ harbor an extra nucleotide inserted in the TψC arm. The predicted secondary structures of 12*S* rRNA and 16*S* rRNA exhibit 47 helices in 4 domains and 60 helices in 6 domains respectively. The control region of *P. ruficollis* with the length of 1305 bp was located between $tRNA^{Glu}$ and $tRNA^{Phe}$, and typical domains of which could be found as other bird groups. Using the data from 13 mitochondrial protein-coding genes, results of a final phylogenetic analysis strongly supports the traditional view that *P. ruficollis* is closely related with Passeridae and Fringillidae.

Keywords: *Pyrgilauda ruficollis*; Complete mitochondrial genome; Phylogeny

To date, there are seven known snow finches (including genus/*Montifringilla* and *Geospiza*), six of them distributed in the Qinghai-Tibet Plateau and adjacent areas, and four of these are endemic to the area. There are different views and arguments on the classification of snow finches—one is that all seven snow finches exist as an independent genus (Howard & Moore, 1980). Ivanitskii (1992) and Gebauer & Kaiser (1994) however proposed dividing them into *Montifringilla* and *Pyrgilauda,* according to their ecological habits. *Montifringilla* would then include *Montifringilla nivalis* and *Montifringilla adamsi*, while *Pyrgilauda* would contain *Pyrgilauda taczanowskii*, *Pyrgilauda ruficollis*, *Pyrgilauda blanfordi*, *Pyrgilauda davidiana* and *Pyrgilauda theresae*. A later clustering analysis by Lei et al (2000) reframed the issue somewhat, with *P. taczanowskii*, *P. ruficollis*, *P. blanfordi* and *P. davidiana* tentatively classified into subgenus *Pyrgilauda* or *Pyrgilauda* species group. A cladistic analysis by Lei et al (2001) also showed that *P. davidiana*, *P. blanfordi* and *P. theresae* have a close relationship and constitute a monophyletic group that forms a clade together with *P. ruficollis*, yielding the monophyletic group of *Pyrgilauda*. The two most recent studies by Zheng (2012) and Gill & Donsker (2013) likewise classified *P. ruficollis* as *Pyrgilauda*, and their analysis is generally accepted by most researchers in the fields.

The variety of findings from the studies on the snow

Received: 18 April 2013; Accepted: 23 December 2013
Foundation items: The National Science Foundation for Outstanding Youth (30925008); Open topic of the Chinese Academy of Sciences Key Laboratory of the Animal Evolution and Systematic(O529YX5105)
*Corresponding author, E-mail: yuanh@snnu.cn

finches and the evolution in the utilized study methodologies underscores both the continued debate on taxonomy and phylogeny. Recently, it has become increasingly popular to usewhole mitochondrial genome sequencing to help settle some of the ongoing disputes and discrepancies in classifying animals and in developing a more complete picture of their evolution. Genetic studies on the mitochondrial gene can be used to resolve the phylogenetic relationships of high-level category (Simon et al, 2006), and also has a better resolution on the genetic relationship of subfamily, genus, inter-species and even the genetic structure of populations (Zhou et al, 2010). Rather surprisingly, as useful as these studies are, to date there has been no research regarding the mitochondrial genome sequencing and phylogenetic studies of *Pyrgilauda ruficollis*—a bird living mainly in Palaearctic areas, belonging to Passeridae in Passeriformes—and instead most phylogenetic studies on this species are mainly focused on morphological evidence, while a small amount reports on RAPD-PCR classification (Qu, 2006).

In this study, we sought to remedy this deficiency by sequencing the mitochondrial genome of *P. ruficollis*, describing in detail the starting position, composition content and structural characteristics of each gene. We also attempted to predict and analyze the secondary structures of 22 tRNAs and two rRNA genes. To provide some new insights into this specie's systematic evolution and the mitochondrial phylogenomics of Passeriformes, we analyzed the phylogenetic position of *P. ruficollis* based on protein-coding genes (PCGs) datasets of 32 Passeriformes mitochondrial genomes published in NCBI using a molecular systematics method.

MATERIALS AND METHODS

Materials

Samples of *P. ruficollis* were collected in Hot Springs Village, Xinghai County, Hainan Prefecture, Qinghai Province on August 12, 2003, at an elevation of 3960 m. Samples were immersed in absolute ethanol at –20 ℃, with replacement of alcohol several times. The voucher specimens (PJ917) were deposited in the Bird Herbarium at the Key Laboratory of Animal Systematic and Evolutionary, Institution of Zoology, Beijing, Chinese Academy of Sciences.

Methods DNA extraction

The total genomic DNA of *P. ruficollis* was extracted and purified using traditional phenol-chloroform-isoamyl alcohol method (Yang et al, 2010) and stored in refrigerator prior to await further analysis.

Primer design and PCR amplification

According to the mitochondrial genome sequence of *Podoces hendersoni* (NC_014879) alongside the 39 mitochondrial genomes of Passeriformes available from GenBank, we used ClustalX to align and search the highly conserved sequences, and then employed Primer Premier 5.0 to design 29 pairs of PCR amplification primers. The design process based on the principle of primer design, and referred to the primer location identified in Sorenson et al (1999) and Sorenson (2003). Oligo was then used to evaluate and modify each primer. Of these, six pairs of primers were selected for L-PCR, while the rest were used as Sub-PCR primers (primers synthesized by Shanghai Sangon Biological Engineering Technology Co.). The approximate location and approximate length of amplified fragment in the whole genome is listed in Supplementary Tables 1 and 2 (Supporting information of http://www.zoores.ac.cn/).

Using 6 pairs of L-PCR primers, the whole mitochondrial genome of *P. ruficollis* was amplified with six overlapped large fragments (L1-L6), then Sub-PCR was used to amplify small fragments with length of 500−1300 bp. PCR products were detected using 0.8% agarose gel electrophoresis. Good bands were cut and purified using DNA purification kit (U Gene), then the target fragments were sequenced. To assist with Sub-PCR, some samples of the large fragments were stored.

The total volume of the reaction system was 50 μL, with DreamTapTM Green PCR Master Mix (2×) 25 μL, Water, nuclease-free 22 μL, upstream and downstream primers (10 μmol/L) 1 μL for each, DNA template 1 μL. The reaction program of L-PCR was as follows; 93 ℃ denaturation for 2 min; (92 ℃ for 10 s, 58−53 ℃ for 30 s, 68 ℃ 10 min) × 20 cycles; (92 ℃ 10 s, 53 ℃30 s, 68 ℃10 min, and each cycle increases 20 s) × 20 cycles; 68 ℃ 7 min; 4 ℃ incubation. The reaction program of Sub-PCR were as follows: 95 ℃ denaturation 4 min (95 ℃ 45 s, 53−58 ℃ 60 s, 72 ℃60 s) × 30 cycles; 72 ℃ 4 min, 4 ℃ incubation.

Sequencing

Using six pairs of L-PCR amplification primers, the whole mitochondrial genome of *P. ruficollis* was amplified with six large overlapping fragments (L1−L6), the regions of walking sequencing ineffective or poor coverage between long fragments were amplified using

Sub-PCR and then sequenced. Sequencing was completed by the Shanghai Sangon Biological Engineering Technology Company.

Sequence assembly, annotation and analysis

The Staden Package 1.7 was used for the sequencing assembly and annotation, tRNAScan-SE 1.21 (http://lowelab.ucsc.edu/tRNAscan-SE/) was then applied to predict the secondary structure of the tRNA. The mitochondrial DNA sequences of *P. hendersoni* was used as the template during annotation. Using MEGA 5.0 software (Tamura et al, 2011), basic group composition, codon usage frequency of protein genes and other information in mitochondrial genome were calculated.

Phylogenetic analysis

Using the protein-coding genes (PCGs) of 32 Passeriformes species and 4 Anseriformes species published in NCBI (see Supplementary Table 3, Supporting information of http://www.zoores.ac.cn/), as well as those obtained from *P. ruficollis*, we compiled a dataset to construct maximum likelihood (ML), Bayesian inference (BI) and maximum parsimony (MP) trees. The maximum likelihood method was implemented by RAxML 7.0.4 (Stamatakis, 2006). The most suitable model (GTR+G) was selected by Modeltest3.7 (Posada & Crandall, 1998).The process performed rapid bootstrap analysis, and searched for ML tree with best-scoring during a separate program. A random seed number started rapid bootstrap analysis, using model GTRGAMMA. Nodal support was evaluated using 1 000 replications of rapid bootstrapping implemented in RaxML.

MrBayes 3.1 (Ronquist & Huelsenbeck, 2003) was used to reconstruct the BI tree with four Anseriformes species acting as an outgroup. GTR + I + G modelwas selected as the optimal evolutionary model using Modeltest 2.2 (Nylander, 2004). Four Markov chains (one cold chain and three hot chains) runs at one million generations, sampling every 100 generations, were used. The first 1000 samples were discarded as burn-in, and consensus tree was then acquired with the posterior probability (PP) in each branches.

PAUP * 4.0b10 was used to construct MP tree with a 1 000 bootstrap resampling test.

RESULTS

Composition and gene arrangement

The mitochondrial genome of *P. ruficollis* contains 13 protein-coding genes (PCGs), 22 tRNA genes, 2 rRNA genes (lrRNA and srRNA) and a non-coding control region (D-Loop region), with length of 16 909 bp (GenBank accession number KC836121). The position of the 37 other genes (see Table 1) were identical to those other published Passeriformes species. In general, gene arrangement is quite tight, and adjacent genes have little or no non-coding nucleotides except for in the control region. There are a total of 30 bp overlapping sequences between 7 genes with length of 1−10 bp. There are 105 bp interval sequences between 18 pairs of genes with length of 1−22 bp. The largest intergenic region was 22 bp, located between the $tRNA^{leu\ (UUR)}$ and *ND1* gene. Finally, 12 pairs of genes had neither overlap nor interval.

Nucleotide composition

The base content of *P. ruficollis* mtDNA was C > A > T > G, with the content of G + C% is slightly greater than A + T%, but we observed no significant difference. The content of the four bases are very close, which is consistent with mitochondrial genome base content known for the vast majority of reported birds. Base content in protein-coding genes is most similar to the entire mitochondrial genome. One issue in the base content of different codons was that the third codon of *P. ruficollis* exhibited some fluctuations, which is in line with the findings reported by Gao et al (2009) which noted the third base content of different codons in 74 mitochondrial genomes of birds, suggesting that the base changes in the mitochondrial genome were mainly caused by the changes of the third codon.

Protein-coding genes and codon usage

The length of 13 protein-coding genes in *P. ruficollis* were equal to those of the other published Passeriformes species. All protein-coding genes used ATG as an initiation codon, except the *COI* gene which was GTG. TAA is also the most commonly termination codon, but *ND1* and *COI* genes used AGG as termination codons, while *ND5* gene used AGA, and *COIII* and *ND4* genes had an incomplete termination codon T. The remaining genes used typical TAA or TAG. Relative synonymous codon usage in *P. ruficollis* was also similar to that found in other birds. Leu, Thr, Ala, Ser and Ile had the highest frequency in 3787 encoded amino acids, accounting for 49.41% of the total observed content.

Secondary structure of tRNA

Totally, there are 22 tRNAs in the *P. ruficollis* mitochondrial genome, including 14 located on the heavy

Table 1 Composition of the *Pyrgilauda ruficollis* mitochondrial genome

Gene	Coding strand	Start position	End position	Length	Intergenic length	Initiation codons	Termination condones
tRNA^{Phe}	H	1	68	68			
SrRNA	H	69	1046	978	0		
tRNA^{Val}	H	1047	1116	70	0		
LrRNA	H	1117	2713	1597	0		
tRNA^{Leu}(UUR)	H	2714	2788	75	0		
ND1	H	2811	3788	978	22	ATG	AGG
tRNA^{Ile}	H	3794	3865	72	5		
tRNA^{Gln}	L	3872	3942	71	6		
tRNA^{Met}	H	3942	4010	69	–1		
ND2	H	4011	5051	1041	0	ATG	TAA
tRNA^{Trp}	H	5051	5120	70	–1		
tRNA^{Ala}	L	5122	5190	69	1		
tRNA^{Asn}	L	5202	5276	75	11		
tRNA^{Cys}	L	5277	5343	67	0		
tRNA^{Tyr}	L	5343	5413	71	–1		
COI	H	5415	6965	1551	1	ATG	AGG
tRNA^{Ser}(UCN)	L	6957	7029	73	–9		
tRNA^{Asp}	H	7033	7101	69	3		
COII	H	7108	7791	684	6	ATG	TAA
tRNA^{Lys}	H	7793	7862	70	1		
ATP8	H	7864	8031	168	1	ATG	TAA
ATP6	H	8022	8705	684	–10	ATG	TAA
COIII	H	8713	9496	784	7	ATG	T
tRNA^{Gly}	H	9497	9565	69	0		
ND3	H	9566	9916	351	0	GTG	TAG
tRNA^{Arg}	H	9918	9987	70	1		
ND4L	H	9989	10285	297	1	ATG	TAA
ND4	H	10279	11656	1378	–7	ATG	T
tRNA^{His}	H	11657	11726	70	0		
tRNA^{Ser}(AGN)	H	11727	11792	66	0		
tRNA^{Leu}(CUN)	H	11792	11862	71	–1		
ND5	H	11863	13680	1818	0	ATG	AGA
CytB	H	13689	14831	1143	8	ATG	TAG
tRNA^{Thr}	H	14846	14926	81	14		
tRNA^{Pro}	L	14928	14997	70	1		
ND6	L	15013	15531	519	15	ATG	TAG
tRNA^{Glu}	L	15533	15604	72	1		
A+T rich		15605	16909	1305	0		

chain and 8 on light chain. Anticodons used in *P. ruficollis* are the same as those in *Pseudopodoces humilis*. All tRNA secondary structure except for *tRNA$^{Ser\ (AGN)}$* were predicted by tRNAScan-SE 1.21, and all of them form a typical

clover structure (Supplementary Figure 1, Supporting information of http://www.zoores.ac.cn/). There were 36 mismatches in the 20 predicted tRNA secondary structures, though the remaining $tRNA^{Lys}$ and $tRNA^{Arg}$ had no mismatch. Among these, there are 27 pairs of GU mismatches, and the remaining are as follow: 2 pairs AC mismatch occurs in anticodon arm of $tRNA^{Asp}$ and amino acid acceptor arm of $tRNA^{His}$; 2 pairs AA mismatch located in TψC arm of $tRNA^{Asp}$ and the DHU arm of $tRNA^{Trp}$; 2 pairs of UU mismatches in the anticodon arm of $tRNA^{Gly}$ and the TψC arm in $tRNA^{Met}$; 3 pairs of CC mismatches in the anticodon arm of $tRNA^{Gly}$ and amino acid acceptor arm of $tRNA^{Leu\ (UUR)}$ and the TψC arm of $tRNA^{Phe}$. The $tRNA^{Ser\ (AGN)}$ lacks a DHU arm, and a CC mismatch is present in the TψC arm of $tRNA^{Phe}$.

Secondary structure of rRNA

The srRNA gene of *P. ruficollis* is located between $tRNA^{Phe}$ and $tRNA^{Val}$ with a length of 978bp. The predicted srRNA gene secondary structure (Supplementary Figure 2) is roughly the same as other published bird srRNA structures, containing four domains with 47 stem-loop structures. The lrRNA gene is located between $tRNA^{Val}$ and $tRNA^{Leu\ (UUR)}$ with a length of 1 597 bp. The predicted lrRNA gene secondary structure of the *P. ruficollis* lrRNA gene contains six domains, containing 60 stem-loop structures (Supplementary Figure 3).

Control region

The control region of *P. ruficollis* is located between $tRNA^{Glu}$ and $tRNA^{Phe}$ with a length of 1305 bp. Similar to the structure in *P. hendersoni* and *Larus brunnicephalus* (Yang et al, 2012), the control region of *P. ruficollis* contains an F, E, D, C, B domain, and conserved sequence blocks (CSB) (Figure 1) .

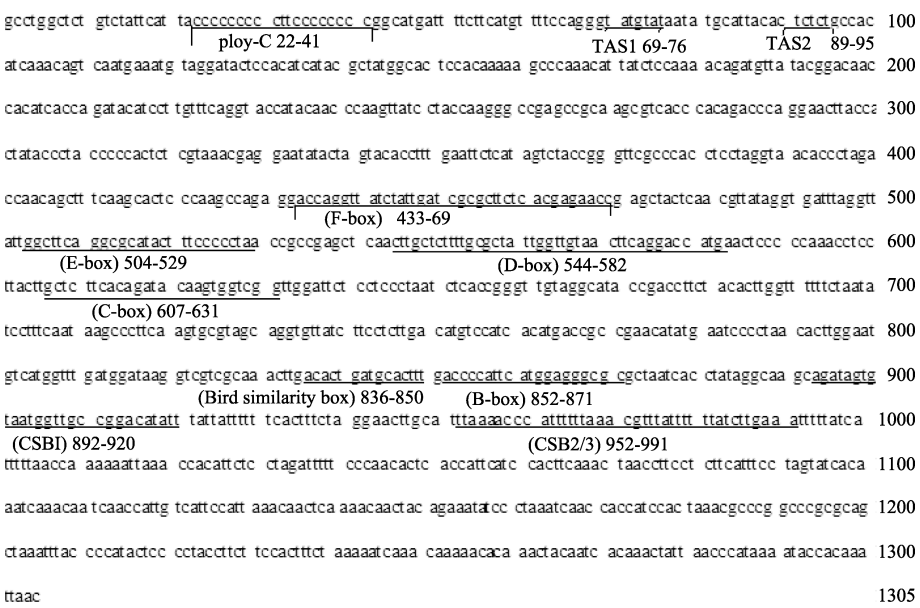

Figure 1 Structure of *Pyrgilauda ruficollis* mitochondrial control region

Phylogenetic analysis

The maximum likelihood (ML), Bayesian inference (BI) and maximum parsimony (MP) trees based on PCG dataset of 33 Passeriformes and 4 Anseriformes species are shown in Figure 2.

The tree topology obtained *via* three different methods are generally similar. The main difference is that ML and MP trees support the partition of Oscines and Suboscines, while the BI tree supports that *Acanthisitta chloris* is the sister group with Oscines and Suboscines. In Passerii, Menuridae is the sister group with Corvida and Passerida, and the monophyly of Passerida and Corvida except for Menuridae is supported. In the 3 superfamily of Passerida, Sylvioidea is located at the base of the trees, and the evolutionary relationship of Passeroidea and Musicicapoidea is closer. Several families of Sylvioidea formed a clade, except *P. humilis* gathered together with Musicicapoidea. Pycnonotidae, Acrocephalidae, Sylviidae and Timaliidae are mutually sister group. Among them, the relationship of Pycnonotidae and Acrocephalidae is closer, while Sylviidae has a close relationship with Timaliidae. InPasseroidea, *P. ruficollis*

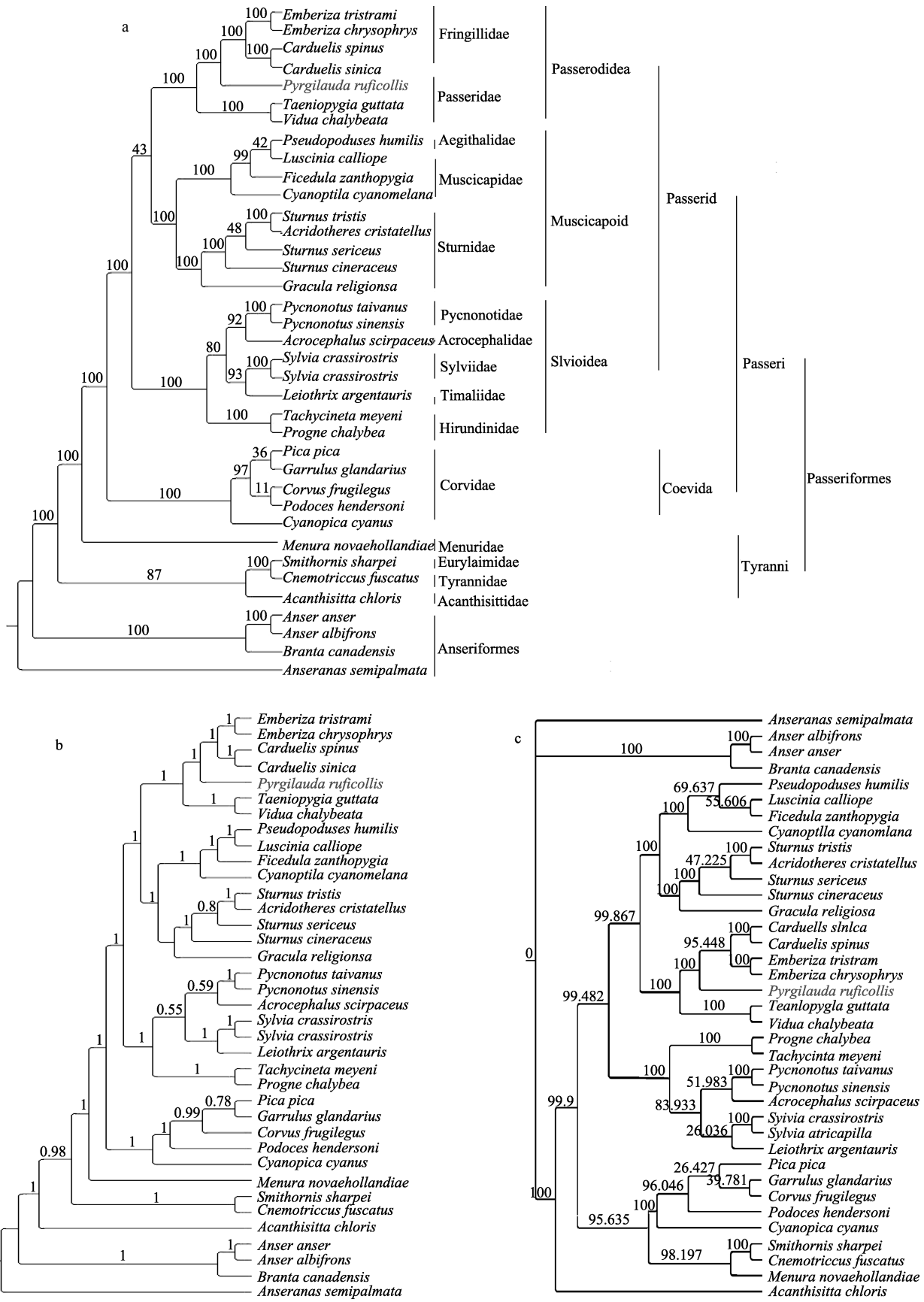

Figure 2 ML tree (a), BI tree (b), and MP tree (c) reconstructed using complete mitochondrial PCG sequence data sets

is close to Fringillidae, but has certain distant with Estrildidae. Musicicapoidea is divided into Sturnidae and Turdidae, and *P. humilis* get together with Turdidae.

DISCUSSION

The cytosine insertion phenomenon in ND3 gene

According to Gao et al (2009), cytosine insertion phenomenon in *ND3* gene was found in almost all of the Palaeognathae species while most Neognathae species had no such phenomenon. The *ND3* sequence of *P. ruficollis* and four Passeriformes published in NCBI shows no cytosine insertion phenomenon at position 174 (Figure 3).

Figure 3 Comparison on the 174 site of the mitochondrial ND3 gene sequence among five representative avian species

Secondary structure of tRNA

The TψC arm of $tRNA^{Phe}$ in the *P. ruficollis* mitochondrial genome has no "C" insertion but forms a CC mismatch, which is different from other Passeriformes species (Figure 1, Box 1) . This findings is differs from that in Harrison et al (2004), in which three matching forms of TψC arm in $tRNA^{Phe}$ in the avian mitochondrial genome is normally present (Figure 4a－c). This result also differs from that reported by Yang et al (2012) regarding the *Larus brunnicephalus* mitochondrial genome (Figure 4-d). It is generally accepted that form one (Figure 4-a) is the ancestor form of modern birds, and commonly exist in four Model populations and owls. Form two meanwhile (Figure 4-b) in prevalent in modern birds, and is more primitive than form three, which is itself a derviative of form two that has lost guanine (G). According to statistics, TψC arm of $tRNA^{Phe}$ has a "C" insertion in the majority of published Passeriformes species (Figure 4-b), while the TψC arm of tRNAPhe in *P. ruficollis* formed a CC mismatch (Figure 4-e). The DHU arm is also missing in the $tRNA^{Ser\ (AGN)}$ gene, which is common in many Vertebrate mitochondria including avians (Wolstenholme, 1992a, b). The missing length of the DHU arm of $tRNA^{Ser\ (AGN)}$ gene gets shorter from invertebrates to higher vertebrates. An earlier study about mammal $tRNA^{Ser\ (AGN)}$ showed that they can form potential L-shaped structure to maintain the distance between the anticodon arm and CCA acceptor arm (Hanada et al, 2000). The function of tRNA transferring specific amino acids and tertiary structure of L shape are maintained by the amino acid acceptor arm and anticodon arm, while the TΨC loop and DHU loop as well as the corresponding arms seems to not inhibit the function of tRNA (Dirheimer et al, 1995; Wolstenholme et al, 1994).

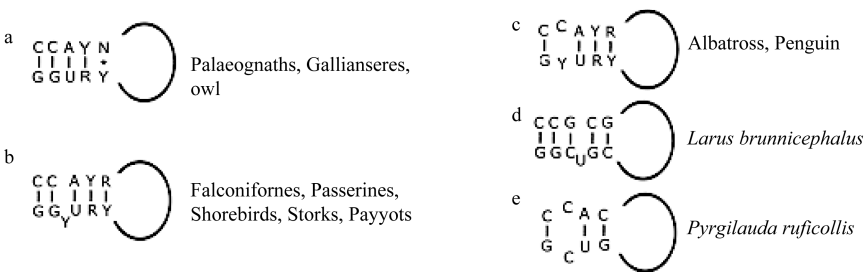

Figure 4 Comparison of the mitochondrial $tRNA^{Phe}$ TψC arm structure among five representative avian species

Secondary structure of rRNA

The single-stranded loop region and double-stranded stem region are alternately arranged in the rRNA gene, being subject to different selection pressures. In general, the loop region has very fast evolution, with high variability between species, while stem region is relatively conserved (Noller, 1984; Woese et al, 1980). However, there are some stem regions containing the same variable length as the loop region, and some loci with a higher variability, while some loop regions do have very conserved sequences (Simon et al, 1994; Vawter & Brown, 1993). Predicting the secondary structure of rRNA can help explain the function and evolutionary information of different sections observed

in the *P. ruficollis* rRNA genes, which can in turn help answer some issues in phylogeny.

For srRNA, the secondary structure of Passeriformes mitochondrial genome is very similar on the whole. The secondary structure of srRNA is the same in *P. Ruficollis* and R*emiz consobrinus* (Gao et al, 2013), and zone one in domain I has a extra base and a cytosine insertion compared to *Gallus gallus* (Eberhard et al, 2001) and *P. hendersoni* (Ke et al, 2010) (Figure 5-a). A similar phenomenon is also appearea in *Pharomachrus pavoninus* (Espinosa De los Monteros A, 2003). There are 5 base pairs and 2 base pairs in zone 24 of Domain II of *P. ruficollis* and *R. consobrinus* respectively, while *G. gallus* and *P. hendersoni* have not formed into paired stems, and 3 base pairs were formed in *Anas platyrhyncho*s and *Coscoroba coscoroba*, which has a prominent cytosine insertion between the first and the second base pairs of the left stem. Moreover, the loop region has more differences among species, indicating that this region has great variability in different species (Figure 5-b). Zone 39 in the stem region of *P. ruficollis* has a significantly reduced stem and relative increased loop regions compared to other species (Figure 5-c). Most Stem variations tend to arise from compensatory mutation, which can guarantee the conservation of secondary structure not interfered by primary sequence variation.

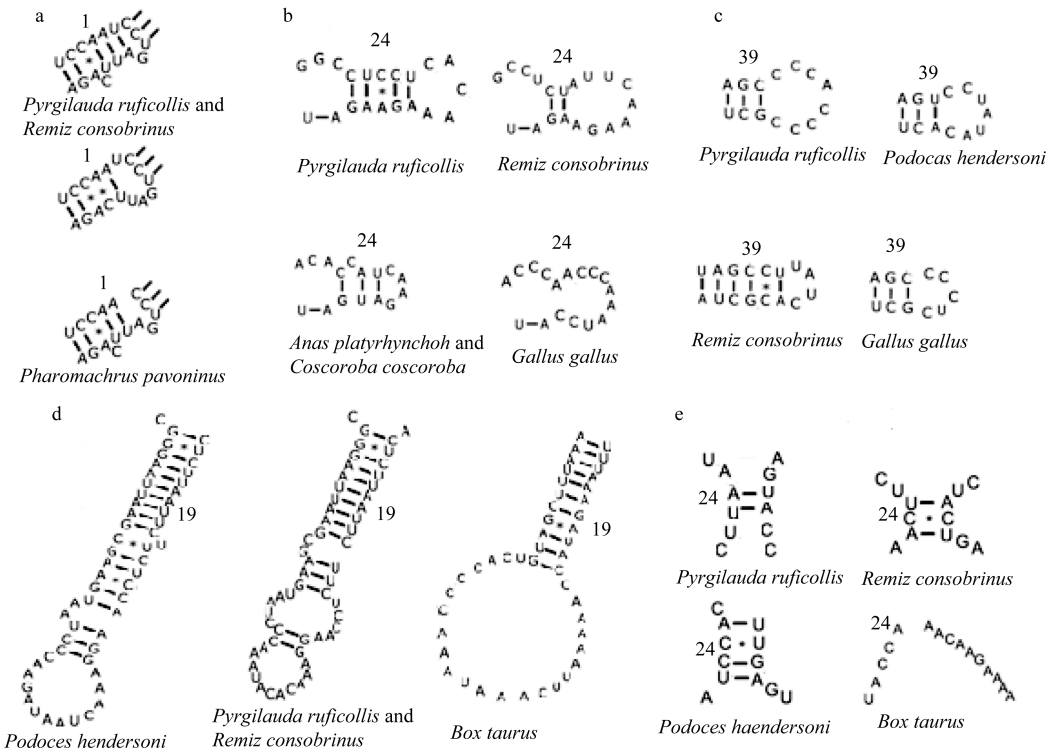

Figure 5 Comparison of variable regions in rRNA secondary structure

For lrRNA, the predicted secondary structure of *P. ruficollis* is identical with *R. consobrinus,* and is basically the same with *P. hendersoni* and *Bos taurus* (Burk et al, 2002). There are some differences: four species have formed a 15 bp base pairing stem in zone 19 of domain II, and formed a hairpin structure stem-loop, while the length of this stem is 19 bp and 11 bp in *P. hendersoni* and *B. taurus* respectively, but the loop region in *B. taurus* is much larger than the other species (Figure 5-d). *P. ruficollis*, *R. consobrinus* and *P. hendersoni* formed a 2−4 bp matching in zone 24 of domain II, while *B. taurus* has matching bases in the same position (Figure 5-e).

Control region

Like vast majority of birds, one control region presented in *P. ruficollis*' mitochondrial genome is located between $tRNA^{Glu}$ and $tRNA^{Phe}$ with a length of 1305 bp. The control region contains regulatory elements of mitochondrial DNA transcription and replication. It suffered high variability according to under less evolutionary pressures, causing it as the fastest evolution region in mitochondrial DNA. However, control region also includes some conservative fragment. According to the

structure information of control region in other species, we make a analysis and prediction to control region in *P. ruficollis* and determines the domains of F, E, D, C, B, CSB and Bird similarity (Figure 1).

The control region in *P. ruficollis* mitochondrial genome has three domains, that is domain I, domain II and domain III, which is similar to other vertebrate. Domain II is more conservative, while domain I and domain III have heterogeneity in base composition and length (Baker & Marshall, 1997).

The domain I in the control region contains an interrupted poly-C structure at 25−42 sites. In Galliformes and Anseriformes, this region and a short sequences below can form hairpin structures Ruokonen & Kvist (2002), while this phenomenon is not found in *P.ruficollis*, and the the causes is not clear. Refers to the termination associated sequence (TAS) of control region in *L.brunnicephalus* (Yang et al, 2012) and *Chrysolophus pictus* (Huang & Liao, 2011), we confirmed the TAS in *P.ruficollis*.

We described the conservative box B, C, D, E and F in domain II of *P. ruficollis* based on *P.hendersoni* and *L.brunnicephalus*, and a more conservative Bird similarity-box was founded before B-box. The specific function of these conserved boxes is not entirely clear, it may related to the formation of D loop and the regulation of H chain synthesis (Figure 1) .

For domain III of control region, CSB1-3 boxes in vertebrates domain III are closely related to the replication initiation of mitochondrial genome (Yokobori & Pääbo, 1995). CSB1 box in *P. ruficollis* is located 892−920 bp of domain III and is extremely conservative compared with other birds. According to CSB2-3 boxes description of *C. pictus* and *A. chukar*, we determined the position of CSB2/3 boxes at 952 − 991 bp, but there is no obvious conservative box features founded in the sequence analysis.

Phylogeny

One of the major reasons for pursuing this study of the mitochondrial genome was to help answer some long-standing questions regarding how several of these avian species fit together. According to the phylogenetic tree we constructed in this study, the position of categories under Passeriformes are basically the same as those found in the other currently available studies. In general though, the mitochondrial genome data presented in this study supports the division of Passeriformes by Sibley & Ahlqulst (1990). In the aspect of Suboscines, the position of *Acanthisitta chloris* is different in different phylogenetic tree. One is that *A. chloris* is clustered with *Cnemotriccus fuscatus* and *Smithornis sharpie*, and another is located in the base of Passeriformes, which is the sister group to Oscines and Suboscines. The former is consistent with the conclusion of traditional morphological classification (Mayr & Greenway, 1956) and DNA hybridization studies (Sibley & Ahlqulst, 1990), while the latter is supported by the study based on nuclear genes and mitochondrial genes (Pacheco et al, 2011; Ericson et al, 2002). From this perspective, the specific location of *A. chloris* needs to be verified using more molecular markers.

For Suboscines, there were many conclusions drawn from difference taxa being different from traditional conclusions reached by taxonomy and DNA hybridization. Menuridae was placed in Suboscines according to traditional taxonomy (Mayr & Greenway, 1956), while it was served as a superfamily under Corvida of Oscines based on DNA hybridization (Sibley & Ahlqulst, 1990). Our results, meanwhile, are consistent with the majority study conclusions based on DNA sequences (Barker et al, 2002; Ericson et al, 2002), supporting the contention that Menuridae is the sister group with Corvida and Passerida. Unfortunately this is largely a preliminary investigation the reliability of this phenomenon has yet to be tested, and the mainly reason is inadequate taxon sampling. Perhaps further studies can shed some light on this issue.

For Passerida, the partition of its three general superfamilies are basically consistent with the results of our study. That said, for *P. humilis*, which was moved from Corvidae to Paridae over the last debate, (James et al, 2003), the taxonomic status is still in doubt and its specific location in phylogenetic tree needs more research. In Passeroidea, Fringillidae first clustered into a branch, and then get together with *P. ruficollis* to form a large branch, which constituted a sister group. Finally, they gather together with the clade containing *Taeniopygia guttata* and *Vidua chalybeate*. According to this, we would speculate that *T. guttata* and *V. chalybeate* were first separated from Fringillidae, and *P. ruficollis* is the evolutional intermediate form. Though an interesting possibility, due to a dearth of data on Passeridae, we can make no definitive conclusions on the evolutionary position of *P. ruficollis*. We do however hope that this data can be paired with other datasets that will allow researchers to provide a more satisfactory answer.

References

Baker AJ, Marshall HD. 1997. Mitochondrial control region sequences as tools for understanding evolution. *In*: Mindell DP. Avian Molecular Evolution and Systematics. San Diego: Academic Press. 51-82.

Barker FK, Barrowclough GF, Groth JG. 2002. A phylogenetic hypothesis for passerine birds: taxonomic and biogeographic implications of an analysis of nuclear DNA sequence data. *Proceedings of the Royal Society B: Biological Sciences,* **269**(1488): 295-308.

Burk A, Douzery EJP, Springer MS. 2002. The secondary structure of mammalian mitochondrial 16S rRNA molecules: refinements based on a comparative phylogenetic approach. *Journal of Mammalian Evolution,* **9**(3): 225-252.

Cerasale DJ, Dor R, Winkler DW, Lovette IJ. 2012. Phylogeny of the Tachycineta genus of New World swallows: Insights from complete mitochondrial genomes. *Molecular Phylogenetics and Evolution,* **63**(1): 64-71.

Dirheimer G, Keith G, Dumas F, Westhof E. 1995. Primary, Secondary and tertiary Structures of tRNAs. *In*: Hill WE, Dahlbert A, Garrett RA, Moore PB, Schlessinger D, Warner JR. tRNA: Structure, Biosynthesis and Function. Washington DC: American Society for Microbiology Press, 93-126.

Eberhard JR, Wright TF, Bermingham E. 2001. Duplication and concerted evolution of the mitochondrial control region in the parrot genus Amazona. *Molecular Biology and Evolution,* **18**(7): 1330-1342.

Ericson PGP, Christidis L, Irestedt M, Norman JA. 2002. Systematic affinities of the lyrebirds (Passeriformes: Menura), with a novel classification of the major groups of passerine birds. *Molecular Phylogenetics and Evolution,* **25**(1): 53-62.

Espinosa De los Monteros A. 2003. Models of the primary and secondary structure for the 12S rRNA of birds: A guideline for sequence alignment. *DNA Sequence,* **14**(4): 241-256.

Gebauer A, Kaiser M. 1994. Biology and behavior of General Asiatic snow finches (*Montifringilla*) and mountain-steppe sparrows (*Pyrgilauda*). *Journal of Ornithology,* **135**(1): 55-57.

Gao RR, Huang Y, Lei FM. 2013. Sequencing and analysis of the complete mitochondrial genome of Remiz consobrinus. *Zoological Research,* **34**(3): 228-237.

Gao YK, Miao YW, Su XQ, Chi ZF, Yu B, Jiang F. 2009. A comprehensive analysis on 74 avian mitochondrial genome base compositions. *Journal of Yunnan Agricultural University,* **24**(1): 51-58.

Gill F, Donsker D. 2013. IOC World Bird List (version 3.3). Available at http://www.worldbirdnames.org.

Hanada T, Suzuki T, Watanabe K. 2000. Translation activity of mitochondrial tRNA with unusual secondary structure. *Nucleic Acids Symposium Series,* **44**(1): 249-250.

Harrison GL, McLenachan PA, Phillips MJ, Slack KE, Cooper A, Penny D. 2004. Four new avian mitochondrial genomes help get to basic evolutionary questions in the late cretaceous. *Molecular Biology and Evolution,* **21**(6): 974-983.

Howard R, Moore A. 1980. A Complete Checklist of the Birds of the World. Oxford: Oxford University Press, 1-613.

Huang ZH, Liao XJ. 2011. Structure of the mitochondrial DNA control region and genetic variation of *Chrysolophus pictus*. *Life Science Research,* 2011, **15**(3): 223-228.

Ivanitskii VV. 1992. On the social behavior of david's snow finches (*Pyrgilauda davidiana*) and snow finches (*Montifringilla nivalis*). *Zoologichesky Zhurnal,* **70**(9): 104-117.

James HF, Ericson PGP, Slikas B, Lei FM, Gill FB, Olson SL. 2003. *Pseudopodoces humilis*, a misclassified terrestrial tit (Paridae) of the Tibetan Plateau: evolutionary consequences of shifting adaptive zones. *Ibis,* **145**(2): 185-202.

Kan XZ, Yuan J, Zhang LQ, Li X, Yu L, Chen L, Guo ZC, Yang JK. 2013. Complete mitochondrial genome of the Tristram's Bunting, *Emberiza tristrami* (Aves: Passeriformes): The first representative of the family Emberizidae with six boxes in the central conserved domain II of control region. *Mitochondrial DNA,* **24**(6): 648-650.

Ke Y, Huang Y, Lei FM. 2010. Sequencing and analysis of the complete mitochondrial genome of *Podoces hendersoni* (Ave, Corvidae). *Hereditas,* **32**(9): 951-960.

Lei FM, Qu YH, Yin ZH. 2001. Phylogeny of species of *Montifringilla* based on cladistic (Passeriform: Ploceidae). *Acta Zootaxonomica Sinica,* **26**(1): 1-7.

Lei FM, Qu YH, Feng ZJ, Ran CZ. 2000. A tentative study on the snow finch (*Montifringilla* spp.) in China based on Clustering. *Acta Zootaxonomica Sinica,* **25**(4): 467-473.

Mayr E, Greenway JC. 1956. Sequence of Passerine families (Aves). *Breviora,* **58**: 1-11.

Mindell DP, Sorenson MD, Dimcheff DE, Hasegawa M, Ast JC, Yuri T. 1999. Interordinal relationships of birds and other reptiles based on whole mitochondrial genomes. *Systematic Biology,* **48**(1): 138-152.

Mossman JA, Birkhead TR, Slate J. 2006. The whole mitochondrial genome sequence of the zebra finch (*Taeniopygia guttata*). *Molecular Ecology Notes,* **6**(4): 1222-1227.

Nilsson MA, Gullberg A, Spotorno AE, Arnason U, Janke A. 2003. Radiation of extant marsupials after the K/T boundary: evidence from complete mitochondrial genomes. *Journal of Molecular Evolution,* **57**(Suppl. 1): S3-S12.

Noller HF. 1984. Structure of ribosomal RNA. *Annual Review of Biochemistry,* **53**: 119-162.

Nylander JAA. 2004. MrModeltest v2 [Computer software and manual]. Evolutionary Biology Centre, Uppsala University.

Pacheco MA, Battistuzzi FU, Lentino M, Aguilar RF, Kumar S, Escalante AA. 2011. Evolution of modern birds revealed by Mitogenomics: timing the radiation and origin of major orders. *Molecular Biology and Evolution,* **28**(6): 1927-1942.

Posada D, Crandall KA, 1998. Modeltest: testing the model of DNA substitution. Bioinformatics, 14 (9): 817-818.

Qian C, Ren Q, Kan X, Guo Z, Yang J, Li X, Yuan J, Qian M, Hu Q, Zhang L. 2013. 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): 129-131.

Qu YH. 2006. Studies on systematic of Snow Finches (*Montifringilla* spp.). Master thesis, Shaanxi Normal University, Xi'an.

Ronquist F, Huelsenbeck JP, 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics, 19(12): 1572-1574.

Ruokonen M, Kvist L. 2002. Structure and evolution of the avian mitochondrial control region. *Molecular Phylogenetics and Evolution,* **23**(3): 422-432.

Sibley CG, Ahlquist JA. 1990. Phylogeny and Classification of Birds: A Study in Molecular Evolution. New Haven: Yale University Press.

Simon C, Frati F, Bekenbach A, Crespi B, Liu H, Flook P. 1994. Evolution, weighting, and phylogenetic utility of mitochondrial gene sequences and a complication of conserved polymerase chain reaction primers. *Annuals of the Etomological Society of America,* **87**(6): 651-701.

Simon C, Buckley TR, Frati F, Stewart JB, Beekenbaeh AT. 2006. Incorporating molecular evolution into phylogenetic analysis, and a new compilation of conserved polymerase chain reaction primers for animal mitochondrial DNA. *Annual Review of Ecology, Evolution, and Systematics,* **37**(1): 545-549.

Singh TR, Shneor O, Huchon D. 2008. Bird mitochondrial gene order: insight from 3 warbler mitochondrial genomes. *Molecular Biology and Evolution,* **25**(3): 475-477.

Slack KE, Janke A, Penny D, Arnason U. 2003. Two new avian mitochondrial genomes (penguin and goose) and a summary of bird and reptile mitogenomic features. *Gene,* **302**(1-2): 43-52.

Slack KE, Delsuc F, McLenachan PA, Arnason U, Penny D. 2007. Resolving the root of the avian mitogenomic tree by breaking up long branches. *Molecular Phylogenetics and Evolution,* **42**(1): 1-13.

Sorenson MD. 2003-02-14[2010-01-08]. Avian mtDNA primers. Boston: Boston University, http://people.bu.edu/msoren/Bird.mt.Primers.pdf.

Sorenson MD, Ast JC, Dimcheff DE, Yuri T, Mindell DP. 1999. Primers for a PCR-based approach to mitochondrial genome sequencing in birds and other vertebrates. *Molecular Phylogenetics and Evolution,* **12**(2): 105-114.

Stamatakis A, 2006. RAxML-VI-HPC: Maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics, 22(21): 2688 – 2690.

Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. 2011. MEGA5: Molecular evolutionary genetic analysis using maximum likelihood, evolutionary distance, and maximum parsimory methods. *Molecular Biology and Evolution*, **28**(10): 2731-2739.

Vawter L, Brown WM. 1993. Rates and patterns of base change in the small subunit ribosomal RNA gene. *Genetics,* **134**(2): 597-608.

Woese CR, Magrum LJ, Gupta R, Siegel RB, Stahl DA, Kop J, Crawford N, Brosius J, Gutell R, Hogan JJ, Noller HF. 1980. Secondary structure model for bacterial 16S ribosomal RNA phylogenetic, enzymatic and chemical evidence. *Nucleic Acids Research,* **8**(10): 2275-2293.

Wolstenholme DR. 1992a. Animal mitochondrial DNA: structure and evolution. *International Review of Cytology,* **141**: 173-216.

Wolstenholme DR. 1992b. Genetic novelties in mitochondrial genomes of multicellular animals. *Current Opinion in Genetics & Development,* **2**(6): 918-925.

Wolstenholme DR, Okimoto R, Macfarlane JL. 1994. Nucleotide correlations that suggest tertiary interactions in the TV-replacement loop-containing mitochondrial tRNAs of the nematodes, *Caenorhabditis elegans* and *Ascaris suum*. *Nucleic Acids Research,* **22**(20): 4300-4306.

Yang C, Lei F M, Huang Y. 2010. Sequencing and analysis of the complete mitochondrial genome of *Pseudopodoces humilis* (Aves, Paridae). *Zoological Research,* **31**(4): 333-344.

Yang C, Wang QX, Huang Y, Xiao H. 2012. Analysis of the complete mitochondrial genome sequence of *Larusbrunnicephalus* (Aves, Laridae). *Hereditas,* **34**(11): 1434-1446.

Yokobori S, Pääbo S. 1995. Transfer RNA editing in land snail mitochondria. *Proceedings of the National Academy of Sciences of the United States of America,* **92**(22): 10432-10435.

Zheng GM. 2012. Ornithology. 2nd ed. Beijing: Science Press.

Zhou ZJ, Ye HY, Huang Y, Shi FM. 2010. The phylogeny of Orthoptera inferred from mtDNA and description of *Elimaea cheni* (Tettigoniidae: Phaneropterinae) mitogenome. *Journal of Genetics and Genomics,* **37**(5): 315-324.