Genetics and Molecular Biology, 38, 2, 162-172 (2015)

DOI: http://dx.doi.org/10.1590/S1415-4757382220140203

Research Article

The complete mitochondrial genome sequence of the little egret (*Egretta garzetta*)

Yi Zou, Mei-dong Jing, Xiao-xin Bi, Ting Zhang and Ling Huang

College of Life Sciences, Ludong University, Yantai, Shandong, P.R. China.

Abstract

Many phylogenetic questions in the Ciconiiformes remain unresolved and complete mitogenome data are urgently needed for further molecular investigation. In this work, we determined the complete mitogenome sequence of the little egret (*Egretta garzetta*). The genome was 17,361 bp in length and the gene organization was typical of other avian mtDNA. In protein-coding genes (PCGs), a C insertion was found in *ND3*, and *CO* III and *ND4* terminated with incomplete stop codons (T). tRNA-Val and tRNA-Ser (AGY) were unable to fold into canonical cloverleaf secondary structures because they had lost the DHU arms. Long repetitive sequences consisting of five types of tandem repeats were found at the 3' end of Domain III in the control region. A phylogenetic analysis of 11 species of Ciconiiformes was done using complete mitogenome data and 12 PCGs. The tree topologies obtained with these two strategies were identical, which strongly confirmed the monophyly of Ardeidae, Threskiorothidae and Ciconiidae. The phylogenetic analysis also revealed that *Egretta* was more closely related to *Ardea* than to *Nycticorax* in the Ardeidae, and *Platalea* was more closely related to *Threskiornis* than to *Nipponia* in the Threskiornithidae. These findings contribute to our understanding of the phylogenetic relationships of Ciconiiformes based on complete mitogenome data.

Keywords: *Egretta garzetta*, mitochondrial genome, phylogenomics.

Received: July 4, 2014; Accepted: December 2, 2014.

Introduction

With more than 9,000 living species, Aves is the most diverse class of vertebrates. The huge number of species, complex morphological characters and wide range of ecological behaviors make it difficult to solve the phylogenetic relationship of birds in traditional taxonomy (Bock, 1956; Howard and Moore, 1980; Monroe and Sibley, 1993).

The order Ciconiiformes, consisting of more than 110 species of large or medium size waders, has traditionally be classified into five families (Ciconiidae, Threskiornithidae, Ardeidae, Balaenicipitidae and Scopidae) (Howard and Moore, 1980; Austin, 1985; Gill, 1990; Clements, 2000; Zheng, 2002). However, there have been various uncertainties regarding the evolutionary relationships of different taxa in this order: (1) The phylogenetic relationships among the five families have been questioned in morphological studies (Kahl, 1972; Cracraft, 1981), (2) the Family Ardeidae was divided into two subfamilies (Ardeinae and Botaurinae) by Bock (1956) and Zheng (1997), but into four subfamilies (Ardeinae, Nycticoracinae, Botaurinae and Tigrisomatinae) by Payne and Risley (1976), and (3) the phylogenetic status of several species in the traditional classification of the subfamily Ardeinae has been questioned. For example, the great egret was initially placed in an independent genus *Casmerodius* (Peter, 1931), but was put in *Egretta* by Bock (1956) and *Ardea* by Payne and Risley (1976). Similarly, the intermediate egret was initially included in *Egretta*, but then placed in *Mesophoyx* by Sibley and Monroe (1990). The taxonomic position of the cattle egret had also changed many times; in early taxonomic literature this species belonged to *Bubulcus* (Peter, 1931), but was subsequently placed in *Ardeola* by Bock (1956) and in *Egretta* by Payne and Risley (1976).

Genome sequences, which provide direct information on evolutionary history, are perfect markers for phylogenetic studies since the resulting analyses can be used to assess and revise the conclusions of traditional taxonomy. In the last 30 years, molecular investigations have shed new light on the evolutionary history of the Ciconiiformes. Based on DNA hybridization results, Sibley *et al.* (1988) merged Ciconiiformes and four other orders (Gaviiformes, Podicipediformes, Lariformes and Charadriiformes) into a huge new order. However, recent molecular studies have proposed the paraphyly of Ciconiiformes because the herons and ibises in this group showed a close relationship with Pelecaniformes, whereas the storks were closely related to Sphenisciformes (Hedges and Sibley, 1994; Cracraft *et al.*, 2004; Hackett *et al.*, 2008; Pacheco *et al.*,

Send correspondence to Ling Huang. College of Life Sciences, Ludong University, 264025 Yantai, Shandong, China. E-mail: huangdl@126.com.

2011). The North American Classification Committee (NACC) has recommended that the families Ardeidae, Threskiornithidae, Balaenicipitidae and Scopidae be merged into Pelecaniformes, and Ciconiiformes was restricted to include only the Ciconiidae.

Molecular studies of the Ardeidae have indicated that day herons and night herons are closely related, and that Nycticoracinae should be merged into Ardeinae, while the tiger herons and boat-billed heron were basal lineages and should be placed in the Tigrisomatinae and Cochleariinae, respectively (Sheldon, 1987; Sheldon and Kinnarney, 1993; Sheldon *et al.*, 1995, 2000). This four-subfamily classification (Ardeinae, Botaurinae, Tigrisomatinae and Cochleariinae) has been generally accepted. Molecular investigations of the subfamily Ardeinae have shown that the great egret and intermediate egret form a monophyletic lineage that is more closely related to *Ardea* than to *Egretta*, indicating that they should not be placed in *Egretta* (Sheldon, 1987; Sibley and Monroe, 1990; Sheldon and Kinnarney, 1993; Sheldon *et al.*, 1995, 2000; Chang *et al.*, 2003).

In molecular systematics, the topologies of phylogenetic trees vary with the molecular markers used and the number of taxa involved (Zwickl and Hillis, 2002). Consequently, some phylogenetic uncertainties in the Ardeinae (such as the evolutionary status of the cattle egrets *Ardeola* and *Butorides*) have not been resolved (Chang *et al.*, 2003; Zhou XP, 2008, PhD thesis, Xiamen University, China).

Mitochondrial DNA (mtDNA), with its intrinsic characteristics (small genome size, simple genome structure, exclusively maternal inheritance, lack of extensive recombination and rapid rate of evolution), has been extensively used in taxonomic and phylogenetic studies of vertebrates (Ingman *et al.*, 2000; Sheldon *et al.*, 2000; Gentile *et al.*, 2009; Zhang and Wake, 2009; Pacheco *et al.*, 2011; Suzuki *et al.*, 2013). Compared to individual genes, complete mitogenomes contain more information on an organisms or taxon's evolutionary history, reduce stochastic errors and minimize the effect of homoplasy in phylogenetic studies (Campbell and Lapointe, 2011). Phylogenies based on complete mitogenomes are generally consistent with those derived from nuclear genes if appropriate sampling of taxa and analysis are applied (Arnason *et al.*, 2002; Reyes *et al.*, 2004; Kjer and Honeycut, 2007). Complete mitogenomes have increasingly been used to address the evolution and radiation of birds (Moum *et al.*, 1994; Sato *et al.*, 1999; Pacheco *et al.*, 2011). To date, more than 260 avian mitogenomes have been deposited in GenBank, only four of which involve species belonging to the Ardeidae (*Egretta eulophotes*, *Ardea novaehollandiae*, *Ixobrychus cinnamomeus* and *Nycticorax nycticora*). The lack of complete mitogenome data is an important limitation in solving the evolutionary puzzles of the Ardeidae and Ciconiiformes.

In this report, we describe the complete mitogenome sequence of the little egret (*Egretta garzetta*) and provide a comprehensive analysis of its genome characters. Although the phylogenetic status of this species has been well-defined by morphological and molecular studies (Bock, 1956; Payne and Risley, 1976; McCracken and Sheldon, 1997; Rabosky and Matute, 2013), the availability of its complete mitogenome data will provide useful information for molecular phylogenetic studies and conservation biology of the Ardeidae.

Material and Methods

Sample collection and extraction of genomic DNA

One specimen of *E. garzetta* was collected from Wuyi Mountain, Fujian Province, China. The specimen was identified based on external characteristics, using the system of Sibley and Monroe (1990). Total genomic DNA was extracted from muscle tissue with a Wizard Genomic DNA purification kit (Promega, Madison, WI, USA) according to the manufacturer's instructions. The concentration of extracted DNA was determined using a spectrophotometer and adjusted to 50 ng/μL.

PCR amplification and sequencing

The *E. garzetta* mtDNA was obtained by polymerase chain reactions (PCR) using 28 primer sets reported by Sorenson *et al.* (1999). The PCR products for each set of primers were < 1,500 bp in size and all fragment sequences overlapped each other by at least 200 bp. PCR amplifications were done with a Mycycler Gradient thermocycler (Bio-Rad) in a final volume of 50 μL, including 5 μL of 10x ^{EX}*Taq* buffer (Mg^{2+}-free; Takara Biotech, Dalian, China), 2.5 mM of each dNTP, 75 mM $MgCl_2$, 10 μM of each primer, 1.5 U of ^{EX}*Taq* polymerase (Takara of Biotech, Dalian, China) and approximately 20-50 ng of total genomic DNA. The reaction included an initial denaturation at 94 °C for 3 min, followed by 35 cycles consisting of denaturation at 94 °C for 10 s, annealing at 50-56 °C for 30 s and extension at 72 °C for 2 min, with a final extension at 72 °C for 10 min. There was a negative control in each round of PCR to check for contamination. The products were electrophoresed on 1.5% agarose gels staining with ethidium bromide and visualized by ultraviolet transillumination. The PCR products were purified with a gel extraction kit (Sangon BioMedical, Shanghai, China) and directly sequenced (both directions) with an ABI 3730XL automatic sequencer (Perkin-Elmer) using an ABI PRISM BigDye Terminator Cycle Sequencing Ready Reaction kit (with AmpliTaq DNA polymerase FS, Applied Biosystems).

Sequence assembly, annotation and analysis

Sequence assembly and annotation were done using the DNASTAR software package (Lasergene version 5.0; Madison, WI, USA). The boundaries of protein-coding genes and rRNA genes were determined by aligning our sequences with the complete mtDNA sequences of *A.*

novaehollandiae (NC_008551) and *Gallus gallus* (NC_001323; Galliformes: Phasianidae) in GenBank. The boundaries and the cloverleaf secondary structures of tRNAs were identified by tRNAscan-SE v 1.12 with the default settings. The complete nucleotide sequence was submitted to GenBank under accession no. NC_023981 and the blast sequences are submitted to DRYAD (doi:10.5061/dryad.3g604). The base composition for protein-coding genes (PCGs), the codon usage of 13 PCGs and the pairwise distances among mitogenomes of the species studied were calculated with MEGA version 5 (Tamura *et al.*, 2011).

Phylogenetic inference using mitogenomes

The phylogenetic relationships among *E. garzetta* and four other species in the Ardeidae (*A. novaehollandiae*, *E. eulophotes*, *I. cinnamomeus* and *N. nycticorax)*, four species in the Threskiornithidae (*Platalea leucorodia*, *Platalea minor*, *Threskiornis aethiopicus* and *Nipponia nippon*) and two species in the Ciconiidae (*Ciconia boyciana*, *Ciconia ciconia*) were constructed with complete mtDNA sequences and 12 PCGs (excluding *ND6*). Two species in the family Anatidae, order Anseriformes (*Branta canadensis*, NC_007011; *Anas platyrhynchos*, EU009397) were designated as outgroups. The relevant information for each genome is presented in Table S1.

The program Modeltest version 3.7 (Posada and Crandall, 1998) was used to choose an appropriate substitution model of sequence evolution. The GTR+I+G model was selected as the best fitting model. For the Bayesian procedure, four independent Markov chains were run for 10,000,000 generations by sampling one tree per 1,000 generations and allowing adequate time for convergence. After discarding the first 2,500 trees (25%) as part of a burn-in procedure that was determined by checking for the likelihood of being stationary, we used the remaining 7,500 sampling trees to construct a 50% majority rule consensus tree. Two independent runs were used to provide additional confirmation of the convergence of the Bayesian posterior probabilities (BPP) distribution.

Results and Discussion

Genome organization and base composition

The complete mitogenome of *E. garzetta* is a circular molecule 17,361 bp in length (Figure 1). This size is intermediate to all available ardeid mitogenomes, which range from 17,180 bp (*I. cinnamomeus*; Zhang *et al.*, 2012) to 17,829 bp (*N. nycticorax*, NC_015807). The gene organization is identical to that of typical avian mtDNA (Wolstenholme, 1992; Boore, 1999; Roques *et al.*, 2004; Gibb *et al.*, 2007; Kan *et al.*, 2010; Zhang, *et al.*, 2012; Figure 1). Table 1 shows the various features of this genome. There are six regions in which genes overlapped by 29 bp and 18 intergenic spacer regions comprising a total of 97 bp.

The base composition of the *E. garzetta* mitogenome revealed a slight bias towards A+T (31.5% A, 23.2% T, 31.8% C and 13.5% G). The A+T content for the whole H-strand, different genes and control regions was estimated for 11 mitogenomes in Ciconiiformes (Table 2). This analysis showed that, except for the first codon of PCGs, other portions of these mitogenomes showed varying degrees of preference for A/T. The equations AT-SKEW= (A-T)/(A+T) and GC-SKEW= (G-C)/(G+C) can be used to calculate the skew for a given strand to investigate nucleotide bias (Perna and Kocher, 1995). The positive AT-skew (0.138) and negative GC-skew (-0.399) for the *E. garzetta* mitogenome suggested the occurrence of more A and C than T and G, which is consistent with other avian mitogenomes (Haring *et al.*, 2001; Kan *et al.*, 2010; Yang *et al.*, 2010; Zhang *et al.*, 2012).

Protein-coding genes and codon usage

The total length of 13 PCGs in the *E. garzetta* mitogenome was 11,225 bp, and most of the PCGs were separated by one or more tRNAs (Figure 1). The gene sizes and structures were not significantly different from those of other avian species (Yamamoto *et al.*, 2000; Haring *et al.*, 2001; Yang *et al.*, 2010; Kan *et al.*, 2010; Zhang *et al.*, 2012). There is a C insertion at position 174 in *ND3*, and this insertion was also found in some species of Palaeognathae, *e.g.*, NC_002784, NC_002778 and NC_002782 (Härlid *et al.*, 1998) and Neognathae, *e.g.*, NC_011307 and NC_010962 (Zhang *et al.*, 2012). Other analyses have proposed that the insertion is not C at position 174 but A at position 175, as reported in the mitogenomes of *Otis tarda* (Gruiformes: Otididae, NC_014046) (Yang *et al.*, 2010) and *Trachemys scripta* (Testudoformes: Emydidae) (Russell and Beckenbach, 2008). The function of this extra C or A in *ND3* and its phylogenetic implications are not well known (Russell and Beckenbach, 2008), but the effect of this insertion on gene expression can be removed by RNA alternative splicing or a frameshift (Mindell *et al.*, 1998).

The average A+T value of 13 PCGs in *E. garzetta* is 53.10% (Table 3). Except for *ND1*, the other PCGs had positive AT-skew (0.016 ~ 0.563) and negative GC-skew (-0.295 ~ -0.733), indicating the occurrence of more A and C than T and G (Table 3). The nucleotide compositions of three codons in PCGs were estimated for 11 species (Table 4). The results showed that the smallest and greatest variations occurred in the second (A 0.5%, G 0.3%, C 0.6%, T 0.5%) and third (A 4.4%, G 3.0%, C 5.5%, T 3.7%) codons, respectively. The second codon is generally considered to have undergone maximum selective pressure, followed by the first and third codons and other non-coding regions. Different selective pressures result in different nucleotide variability (Zhong *et al.*, 2002). Table 4 also shows that the G content of the third codon (only 4.1%) was the

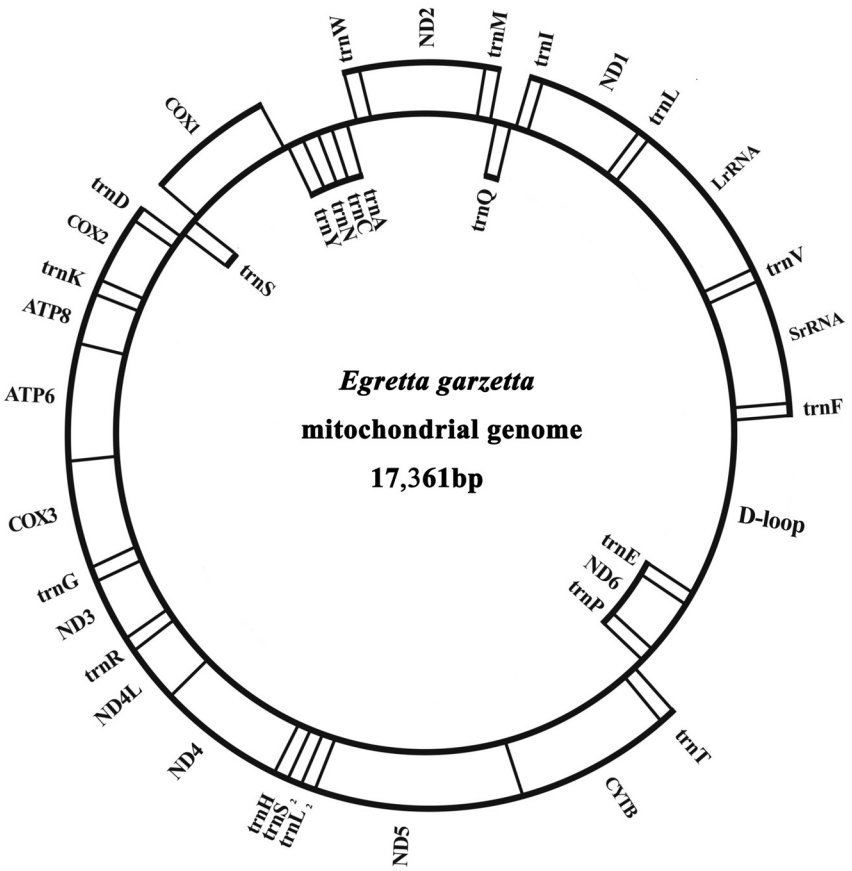

Figure 1 - Gene organization of the *E. garzetta* mitogenome. *ND*1-6 refers to NADH dehydrogenase subunits 1-6, *CO*I-III refer to cytochrome c oxidase subunits 1-3, *ATP6* and *ATP8* refer to ATPase subunits 6 and 8, and *Cyt b* refers to cytochrome b. Twenty-two tRNA genes are designated by single-letter amino acid codes.

smallest of the three codons. A similar phenomenon has also been found in mammalian mitogenomes (Reyes *et al.*, 2004; Gibson *et al.*, 2005).

The start and stop codons for the PCGs of the *E. garzetta* mitogenome are shown in Table 1. *COIII* and *DN4* terminated with an incomplete stop codon (T). The use of an incomplete stop codon (T) is common in avian (Härlid *et al.*, 1998; Haring *et al.*, 2001; Yang *et al.*, 2010; Zhang *et al.*, 2012) and mammalian (Wolstenholme, 1992; Arnason *et al.*, 2002; Gibson *et al.*, 2005; Bi *et al.*, 2012; Chen *et al.*, 2012; Song *et al.*, 2012) mitogenomes, and can form a complete UAA terminal signal by posttranscriptional polyadenylation (Ojala *et al.*, 1981; Boore, 2004).

The *ND6* gene was located in the L-strand and its base composition was very different from the other 12 PCGs (Table 3) so it was excluded from the codon usage analysis. Twelve *E. garzetta* PCGs consisted of 3,626 codons, excluding termination codons (Table S2). The usage frequencies of 21 amino acids ranged from 0.69% (Cys) to 17.9% (Leu). Except for Leu, the most frequently used amino acids were Ile (11.47%), Thr (9.93%) and Ala (7.73%), which was similar with those of other ardeid species (Zhang *et al.*, 2012).

Ribosomal and transfer RNA genes

Animal mitogenomes contain small (*srRNA*) and large (*lrRNA*) subunits of rRNA (Wu *et al.*, 2003; Gibson *et al.*, 2005; Kan *et al.*, 2010; Krajewski *et al.*, 2010; Bi *et al.*, 2012; Chen *et al.*, 2012; Zhang *et al.*, 2012; Gao *et al.*, 2013), and *E. garzetta* was no exception (Figure 1). The A+T content for *srRNA* and *lrRNA* was 50.8% and 54.7%, respectively, and these values were relatively small among the 11 mitogenomes (Table 2).

Based on the respective anticodons and secondary structures, 22 tRNA genes were identified and their sizes ranged from 67 bp ($tRNA^{Cys}$) to 74 bp ($tRNA^{Leu}$ UUR, $tRNA^{Asn}$, $tRNA^{Ser}$ UCN, $tRNA^{Glu}$). Twenty tRNAs can fold into canonical cloverleaf secondary structures, while tRNA-Val and tRNA-Ser (AGY) lost the DHU (dihydrouracil) arms. The cloverleaf structures of tRNA-Val and tRNA-Ser (AGY) were identified by comparing them with counterparts in the *E. eulophotes* mitogenome (NC_009736). In vertebrate mitogenomes, tRNA-Ser (AGY) generally cannot fold into the canonical cloverleaf secondary structure (Härlid *et al.*, 1998; Shi *et al.*, 2002; Wu *et al.*, 2003; Yang *et al.*, 2010; Gao *et al.*, 2013). Al-

though the gene sizes and anticodon nucleotides agreed with those described for other vertebrates, there were some atypical pairings in the stem regions, such as A-A, A-C, U-C and U-U wobbles. Generally, the tRNA cloverleaf structure contained 7 bp in the aminoacyl stem, 5 bp in the TΨC and anticodon stems, and 4 bp in the D-stem. However, some tRNAs, *e.g.*, tRNA-Phe, tRNA-Leu (CUN) and tRNA-Ile, lacked one or two bp in the T-stem, anticodon stem or D-stem.

Non-coding regions

The non-coding region (the control region, mtCR) of the *E. garzetta* mitogenome was determined as1,812 bp in

Table 1 - Organization of the *E. garzetta* mitochondrial genome.

Gene	Position[a]		Size (bp)	Spacer (+)/ Overlap (-)	Strand[b]	Codon	
	From	To				Start[c]	Stop[c]
tRNA-Phe	1	69	69	0	H		
12s-rRNA	70	1040	971	0	H		
tRNA-Val	1041	1111	71	0	H		
16s-rRNA	1112	2718	1607	0	H		
tRNA-Leu (UUR)	2719	2792	74	8	H		
ND1	2801	3778	978	7	H	ATG	AGA
tRNA-Ile	3786	3856	71	11	H		
tRNA-Gln	3868	3937	70	0	L		
tRNA-Met	3938	4005	68	0	H		
ND2	4006	5044	1039	0	H	ATG	TAG
tRNA-Trp	5045	5116	72	2	H		
tRNA-Ala	5119	5186	68	10	L		
tRNA-Asn	5197	5270	74	3	L		
tRNA-Cys	5274	5340	67	-1	L		
tRNA-Tyr	5340	5411	72	13	L		
CO I	5425	6975	1551	-9	H	GTG	AGG
tRNA-Ser (UCN)	6967	7040	74	2	L		
tRNA-Asp	7043	7111	69	1	H		
CO II	7113	7796	684	1	H	ATG	TAA
tRNA-Lys	7798	7867	70	1	H		
ATP8	7869	8036	168	-10	H	ATG	TAA
ATP6	8027	8710	684	-1	H	ATG	TAA
CO III	8710	9493	784	0	H	ATG	T[d]
tRNA-Gly	9494	9562	69	0	H		
ND3	9563	9914	352	2	H	ATT	TAA
tRNA-Arg	9917	9985	69	1	H		
ND4L	9987	10283	297	-7	H	ATG	TAA
ND4	10277	11654	1378	0	H	ATG	T[d]
tRNA-His	11655	11724	70	0	H		
tRNA-Ser (AGY)	11725	11792	68	-1	H		
tRNA-Leu (CUN)	11792	11863	72	0	H		
ND5	11864	13678	1815	10	H	ATG	AGA
Cyt b	13689	14831	1143	3	H	ATG	TAA
tRNA-Thr	14835	14904	70	11	H		
tRNA-Pro	14916	14987	72	8	L		
ND6	14996	15472	477	3	L	ATG	AGA
tRNA-Glu	15476	15549	74	0	L		
Control region	15550	17361	1812	0	H		

[a]Position numbering starts with the 5' position of the Control region; [b]Genes transcribed from the L or H strand; [c]Start and stop codons of protein-coding genes; [d]Protein-coding genes overlapping with tRNA genes end with an incomplete stop codon.

Table 2 - Genomic characteristics of 11 avian mtDNAs.

Species	Heavy-strand		12 Protein-coding genes					LrRNA gene		SrRNA gene		tRNA gene		Control region	
	Length (bp)	AT%	Length (bp)	AT% (all)	AT% (1st)	AT% (2nd)	AT% (3rd)	Length (bp)	AT%	Length (bp)	AT%	Length (bp)	AT%	Length (bp)	AT%
P. leucorodia	15,585	55.3	10,874	54.9	49.4	58.7	56.8	1,599	56.4	974	53.2	1,567	58.7	1,140	56.1
P. minor	15,784	55.4	10,875	55.0	49.4	58.7	56.6	1,599	56.2	975	53.2	1,552	58.9	1,352	56.6
T. aethiopicus	15,826	55.2	10,874	54.6	50.2	57.7	55.4	1,598	55.1	973	52.2	1,567	59.1	1,382	57.8
N. nippon	15,567	54.0	10,874	53.2	49.3	58.7	51.7	1,603	55.5	977	52.2	1,552	57.5	1,160	59.1
A. novaehollandiae	16,354	55.4	10,875	53.4	49.8	57.7	52.0	1,606	55.4	970	51.4	1,555	57.2	1,922	67.1
E. eulophotes	16,421	55.0	10,870	53.3	48.9	57.7	53.0	1,605	54.8	971	51.6	1,550	56.3	1,997	64.7
E. garzetta	16,245	55.0	10,873	53.4	49.8	57.8	53.0	1,607	54.7	971	50.8	1,553	56.2	1,812	65.3
I. cinnamomeus	16,027	57.1	10,873	56.1	51.0	58.8	58.5	1,591	55.9	971	53.9	1,555	58.9	1,609	65.3
N. nycticorax	16,665	56.1	10,873	54.8	50.1	58.9	55.8	1,596	56.0	973	52.1	1,551	57.0	2,244	62.8
C. boyciana	16,487	53.8	10,871	52.4	48.3	57.5	50.4	1,612	53.6	968	51.6	1,550	57.3	2,053	60.5
C. ciconia	16,212	53.8	10,874	52.6	48.3	57.5	51.5	1,608	54.0	968	51.7	1,550	57.2	1,779	59.5
Average	16,107	55.1	10,873	54.0	49.5	58.2	54.1	1,602	55.2	972	52.2	1,555	57.7	1,677	61.3

length and located between $tRNA^{Glu}$ and $tRNA^{Phe}$ (Table 1, Figure 1). The mtCR controls the replication and transcription of animal mitogenomes (Shadel and Clayton, 1997; Taanman, 1999). Based on the nucleotide composition, the mtCR region of *E. garzetta* contains three domains: a 5'-peripheral domain (Domain I), a central conserved domain (Domain II) and a 3'-peripheral domain (Domain III), an organization that was similar to that of other birds (Southern *et al.*, 1988; Saccone *et al.*, 1991; Randi *et al.*, 2000; Roques *et al.*, 2004; Wang *et al.*, 2008; Yang *et al.*, 2010; Zhang *et al.*, 2012; Figure 2).

In Domain I (nt 1-328), two putative extended termination-associated sequence blocks (ETAS1 and ETAS2) were recognized and two putative termination-associated sequences (TAS, conserved palindromic motifs 5'-TACAT-3' and 5'-TATAT-3') that act as a signal to terminate synthesis of the control region (Saccone *et al.*, 1991; Randi and Lucchini, 1998; Yamamoto *et al.*, 2000; Haring *et al.*, 2001; Roques *et al.*, 2004) were found in ETAS1. In some birds and mammals, there is a C structure located close to the 5'-peripheral domain of Domain I that can potentially form a stable goose hairpin structure (Quinn and Wilson, 1993; Douzery and Randi, 1997; Sbisà *et al.*, 1997; Randi and Lucchini, 1998); this structure consists of a stem with seven complementary 'C's/'G's and a loop containing a TCCC motif (Dufresne *et al.*, 1996; Yang *et al.*, 2010). This structure is speculated to be related to H-strand termination (Dufresne *et al.*, 1996). The hairpin structure cannot be formed in any of the available ardeid mitogenomes because the interrupted poly-C sequences in Domain I of four species (*A. novaehollandiae* NC_008551, *E. eulophotes* NC_009736, *N. nycticora* NC_015807 and *E. garzetta* NC_023981) are not followed by a G stretch and Domain I of *I. cinnamomeus* has no poly-C sequence (Zhang *et al.*, 2012). A sequence block similar to the conserved sequence block (CSB1) was found in Domain I (Figure 2) and similar structures have been observed in other avian mitogenomes (Desjardins and Morais, 1990; Quinn and Wilson, 1993; Randi and Lucchini, 1998; Kan *et al.*, 2010; Zhang *et al.*, 2012).

In Domain II (nt 329-794), four conserved sequence boxes (F, E, D and C) were detected (Figure 2) after aligning with reported counterparts in birds and mammals (Walberg and Clayton, 1981; Southern *et al.*, 1988; Desjardins and Morais, 1990; Quinn and Wilson, 1993; Randi and Lucchini, 1998; Roques *et al.*, 2004; Kan *et al.*, 2010; Yang *et al.*, 2010; Zhang *et al.*, 2012).

Domain III (nt 795-1812) comprised a conserved sequence block (CSB-1) that regulates mtDNA replication (Figure 2). A poly(C) sequence located upstream of the CSB1 was assumed to represent the origin of H-strand replication (O_H) (Walberg and Clayton, 1981; Figure 2). A poly (T) sequence located downstream of the CSB1 was also observed in the mtCR of other birds (NC_008551, NC_009736; NC_015807; Kan *et al.*, 2010; Zhang *et al.*,

Table 3 - Base composition for protein-coding genes found in mtDNA of *E. garzetta*.

Gene	Length (bp)	Proportion of nucleotides (%)					AT Skew	GC Skew
		A	C	G	T	A+T		
ND1	978	26.38	34.46	12.68	26.48	52.86	-0.002	-0.462
ND2	1039	32.63	33.59	10.11	23.68	56.31	0.159	-0.537
COX1	1551	28.24	30.11	16.38	25.27	53.51	0.056	-0.295
COX2	684	31.43	31.43	14.18	22.95	54.38	0.156	-0.378
ATP8	168	32.14	38.69	5.95	23.21	55.35	0.161	-0.733
ATP6	684	30.12	36.84	9.94	23.10	53.22	0.132	-0.575
COX3	784	28.57	31.76	15.43	24.23	52.80	0.082	-0.346
ND3	352	26.70	36.08	11.36	25.85	52.55	0.016	-0.521
ND4L	297	29.97	35.35	11.45	23.23	53.20	0.127	-0.511
ND4	1378	31.49	36.21	9.65	22.64	54.13	0.163	-0.579
ND5	1815	31.90	35.43	10.85	21.82	53.72	0.188	-0.531
CYTB	1143	27.47	37.10	12.60	22.83	50.30	0.092	-0.493
ND6	477	37.53	41.93	10.06	10.48	48.01	0.563	-0.613
Average		30.35	35.31	11.59	22.75	53.10	0.146	-0.506

Table 4 - Nucleotide compositon of the 13 protein-coding genes.

Species	1st codon position				2nd codon position				3rd codon potion			
	A%	G%	C%	T%	A%	G%	C%	T%	A%	G%	C%	T%
P. leucorodia	29.7	20.2	30.0	20.1	20.1	12.3	29.3	38.3	41.4	3.8	40.4	14.4
P. minor	29.7	20.2	29.9	20.2	20.0	12.4	29.2	38.4	41.3	4.0	40.0	14.7
T. aethiopicus	29.5	20.4	29.7	20.3	20.0	12.4	29.6	38.0	41.0	4.1	40.8	14.1
N. nippon	29.5	20.4	30.7	19.4	20.0	12.4	29.3	38.3	39.2	5.4	43.3	12.1
A. novaehollandiae	30.1	20.1	30.4	19.4	20.1	12.2	29.7	38.0	40.6	3.9	44.1	11.4
E. eulophotes	30.3	20.0	30.6	19.1	20.0	12.4	29.7	37.9	40.6	4.0	43.9	11.5
E. garzetta	30.0	20.1	30.5	19.4	20.1	12.2	29.7	38.0	40.5	4.0	43.7	11.8
I. cinnamomeus	30.8	19.4	28.8	21.0	20.1	12.3	29.4	38.2	43.2	2.4	39.6	14.8
N. nycticorax	30.5	20.0	30.0	19.5	20.2	12.3	29.3	38.2	40.4	4.6	39.9	15.1
C. boyciana	29.7	20.5	31.1	18.7	19.8	12.4	29.8	38.0	38.8	4.6	45.1	11.5
C. ciconia	29.7	20.5	31.1	18.7	19.7	12.5	29.8	38.0	39.1	4.4	44.5	12.0
Range	1.3	1.1	2.3	2.3	0.5	0.3	0.6	0.5	4.4	3.0	5.5	3.7
Average	30.0	20.2	30.3	19.6	20.0	12.3	29.5	38.1	40.6	4.1	42.3	13.0

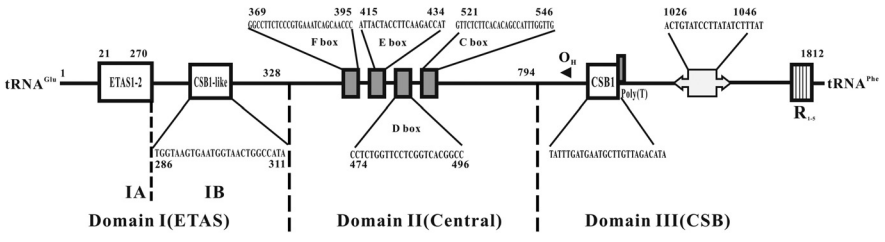

Figure 2 - Schematic representation of the control region in the mitogenome of *E. garzetta*. The first box represents the extended termination-associated sequences (ETAS1 and ETAS2). Boxes F, E, D and C represent the conserved sequence boxes in the central domain. CSB – conserved sequence block, CSB-like – a sequence similar to CSB, LSP and HSP – light-strand and heavy-strand transcription promoters, respectively, and Rs – tandem repeats in the control region.

2012). The bidirectional light- and heavy-strand transcription promoters (LSP/HSP) described in other birds (L'abbé *et al.*, 1991; Randi and Lucchini, 1998; Ritchie and Lambert, 2000; Kan *et al.*, 2010; Zhang *et al.*, 2012) also existed in Domain III of *E. garzetta*. In addition, long tandem repeats were found at the 3' end of Domain III and could be divided into two regions: the first region (nt 977 to 1399) contained three types of tandem repeats: 5'-TACTTTAAAGCACTAAAA-3' (6x18 bp), 5'-TTTCATTAAAAATATACTATACCCTTCATGAAC-3' (5x33 bp), and 5'- TGTATCCTTATATCTTTATGT TACCTTTAC-3' (4x30 bp) while the second region (nt 1406 to 1804) comprised two types of tandem repeats: 5'-TAAACAA -3' (26x7 bp) and 5'- CAAACAA -3' (30x7 bp). The existence of repetitive sequences contributed to the large size of the mtCR and the high content of A. Similar tandem repeats (CAAA or CAAACAA) were found in species of Charadriiformes (NC_003712, NC_003713, NC_007978, NC_018548, NC_017601, NC_024069; Wenink *et al.*, 1994) and Gruiformes (Yang *et al.*, 2010), and in *C. boyciana* in Ciconiiformes (Yamamoto *et al.*, 2000). These repetitive sequences have been speculated to result from the pause of H-strand replication and subsequent slipped mispairing (Fumagalli *et al.*, 1996). The presence of similar conserved repeat sequences in different animal groups (Douzery and Randi, 1997; Nesbø *et al.*, 1998) has led some researchers to propose that these tandem repeats may have an important role in regulating mitogenome replication and transcription (Delarbre *et al.*, 2001; Delport *et al.*, 2002).

Phylogenomic relationships of 11 species in Ciconiiformes

Mitochondrial sequences provide valuable information for tracing the history of gene rearrangements and phylogenetic reconstructions (Härlid *et al.*, 1998; Braband *et al.*, 2010; Oh *et al.*, 2010; Yang *et al.*, 2010; Cerasale *et al.*, 2012). The availability of an increasing number of complete avian mitogenomes has allowed the construction of phylogenetic trees with better resolution, the results of which show better agreement with morphological and nuclear marker studies (Zhang and Wake, 2009; Pacheco *et al.*, 2011). The phylogenetic tree that included *E. garzetta* and ten other species in Ciconiiformes (Table S1) was constructed using complete mitogenome sequences, with *A. platyrhynchos* (EU009397) and *B. canadensis* (NC_007011) as outgroups. Since some investigators have preferred to use PCGs in tree construction (Härlid *et al.*, 1998; Gibson *et al.*, 2005; Shen *et al.*, 2009; Zhang *et al.*, 2012), we also ran an analysis with 13 PCGs to assess the congruence between these two strategies. The results showed that although several regions (tRNAs, CR, rRNAs and *ND6*) presented some problems in the analysis, *e.g.*, difficulties in alignment, numerous gaps, potential saturation and heterogeneous base composition (Gardner *et al.*, 2005; Sullivan and Joyce, 2005; Krajewski *et al.*, 2010; Oh *et al.*, 2010), the topologies of the phylogenetic trees generated by the two strategies were the same (Figure 3).

The phylogenetic relationships among species/genera within the three families examined here were consistent with the conclusions of previous investigations (Sheldon *et al.*, 2000; Chang *et al.*, 2003; Zhang *et al.*, 2012). The monophyly of the Ardeidae, Threskiorothidae and Ciconiidae was strongly confirmed (posterior probabilities = 1.00; Figure 3). In the Ardeidae, *I. cinnamomeus* was the basal clade and *Egretta* more closely related to *Ardea* than to *Nycticorax*. In Threskiornithidae, *Platalea* was more closely related to *Threskiornis* than to *Nipponia*. The relationships revealed by the phylogenetic trees were also supported by the pairwise distances among mitogenomes (Table S3).

With regard to the evolutionary relationships among the three families, our results supported a closer relationship between Threskiorothidae and Ciconiidae than between Threskiorothidae and Ardeidae, a conclusion similar to that based on amino acid data from 12 PCGs (Zhang *et al.*, 2012), but different from that of Hackett *et al.* (2008) and Pacheco *et al.* (2011). Since the topologies of molecular phylogenetic trees often vary with the markers and taxa used (Zwickl and Hillis, 2002), divergent evolutionary relationships have often been suggested for the families of

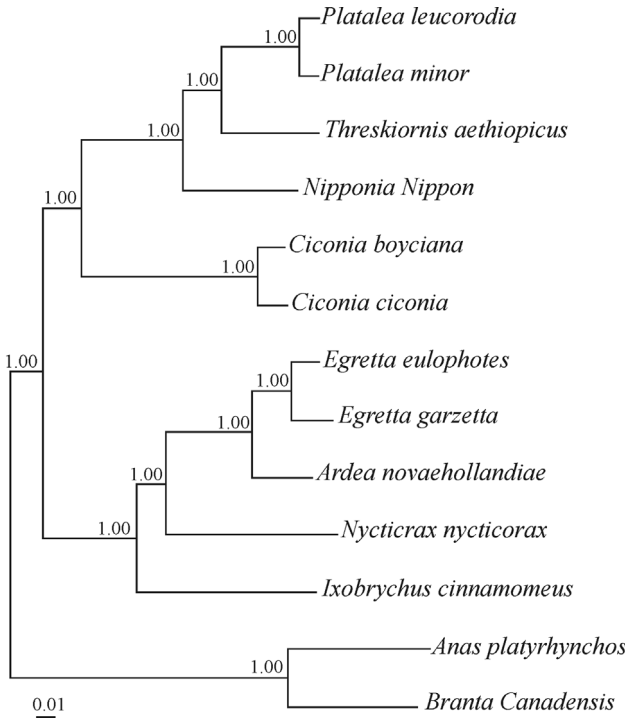

Figure 3 - Bayesian tree based on the complete mitochondrial genome data and 13 PCGs with the GIR+I+G model. The horizontal length of each branch corresponds to the substitution rates estimated with the model. *Anas platyrhynchos* and *Branta canadensis* were used as outgroups. Numbers on the branches are the bootstrap values for Bayesian posterior probability.

Ciconiiformes (Gibb *et al.*, 2007; Hackett *et al.*, 2008; Pacheco *et al.*, 2011; Zhang *et al.*, 2012; this study). More complete mitogenome data for the Ardeidae (and other families in Ciconiiformes) are urgently needed for detailed molecular systematic analyses in this order. The mitogenome sequence data presented here represent a contribution to this long-term goal.

Acknowledgments

This work was supported by the Natural Scientific Foundation of China (grant nos. 31171189 and 31371252).

References

Arnason U, Adegoke JA, Bodin K, Born EW, Esa YB, Gullberg A, Nilsson M, Short RV, Xu X and Janke A (2002) Mammalian mitogenomic relationships and the root of the eutherian tree. Proc Natl Acad Sci USA 99:8151-8156.

Austin OL (1985) Families of Birds. Golden Press, New York, pp 7-37.

Bi XX, Huang L, Jing MD, Zhang L, Feng PY and Wang AY (2012) The complete mitochondrial genome sequence of the black-capped capuchin (*Cebus apella*). Genet Mol Biol 35:545-552.

Bock WJ (1956) A generic review of the family Ardeidae (Aves). Am Mus Novit 1779:1-49.

Boore JL (1999) Animal mitochondrial genomes. Nucleic Acids Res 27:1767-1780.

Boore JL (2004) Complete mitochondrial genome sequence of *Urechis caupo*, a representative of the phylum Echiura. BMC Genomics 5:e67.

Braband A, Cameron SL, Podsiadlowski L, Daniels SR and Mayer G (2010) The mitochondrial genome of the onychophoran *Opisthopatus cinctipes* (Peripatopsidae) reflects the ancestral mitochondrial gene arrangement of Panarthropoda and Ecdysozoa. Mol Phylogenet Evol 57:285-292.

Campbell V and Lapointe FJ (2011) Retrieving a mitogenomic mammal tree using composite taxa. Mol Phylogenet Evol 58:149-156.

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

Chang Q, Zhang BW, Jin H, Zhu LF and Zhou KY (2003) Phylogenetic relationships among 13 species of herons inferred from mitochondrial 12S rRNA gene sequences. Acta Zool Sin 49:205-210 (in Chinese with English abstract).

Chen W, Sun Z, Liu Y, Yue B and Liu S (2012) The complete mitochondrial genome of the large white-bellied rat, *Niviventer excelsior* (Rodentia, Muridae). Mitochondrial DNA 23:363-365.

Clements JF (2000) Birds of the World: A Checklist. 5th edition. Ibis Publishing Company Press, Vista, pp 18-25.

Cracraft J (1981) Toward a phylogenetic classification of the recent birds of the world (Class Aves). Auk 98:681-714.

Cracraft J, Barker FK, Braun MJ, Harshman J, Dyke GJ, Feinstein J, Stanley S, Cibois A, Schikler P, Beresford P, *et al.* (2004) Phylogenetic relationships among modern birds (Neornithes). Toward an avian tree of life. In: Cracraft J and Donoghue MJ (eds) Assembling the Tree of Life. Oxford University Press, New York, pp 468-489.

Delarbre C, Rasmussen AS, Arnason U and Gachelin G (2001) The complete mitochondrial genome of the hagfish *Myxine glutinosa*: Unique features of the control region. J Mol Evol 53:634-641.

Delport W, Ferguson JW and Bloomer P (2002) Characterization and evolution of the mitochondrial DNA control region in hornbills (Bucerotifoormes). J Mol Evol 54:794-806.

Desjardins P and Morais R (1990) Sequence and gene organization of the chicken mitochondrial genome: A novel gene order in higher vertebrates. J Mol Biol 212:599-634.

Douzery E and Randi E (1997) The mitochondrial control region of Cervidae: Evolutionary patterns and phylogenetic content. Mol Biol Evol 14:1154-1166.

Dufresne C, Mignotte F and Gueride M (1996) The presence of tandem repeats and the initiation of replication in rabbit mitochondrial DNA. Eur J Biochem 235:593-600.

Fumagalli L, Taberlet P, Favre L and Hausser J (1996) Origin and evolution of homologous repeated sequences in the mitochondrial DNA control region of shrews. Mol Biol Evol 13:191-199.

Gao RR, Huang Y and Lei FM (2013) Sequencing and analysis of the complete mitochondrial genome of *Remiz consobrinus*. Zool Res 34:228-237 (in Chinese with English abstract).

Gardner PP, Wilm A and Washietl S (2005) A benchmark of multiple sequence alignment programs upon structural RNAs. Nucleic Acids Res 33:2433-2439.

Gentile G, Fabiani A, Marquez C, Snell HL, Snell HM, Tapia W and Sbordoni V (2009) An overlooked pink species of land iguana in the Galápagos. Proc Natl Acad Sci USA 106:507-511.

Gibb GC, Kardailsky O, Kimball RT, Braun EL and Penny D (2007) Mitochondrial genomes and avian phylogeny: Complex characters and resolvability without explosive radiations. Mol Biol Evol 24:269-280.

Gibson A, Gowri-Shankar V, Higgs PG and Rattray M (2005) A comprehensive analysis of mammalian mitochondrial genome base composition and improved phylogenetic methods. Mol Biol Evol 22:251-264.

Gill FB (1990) Ornithology. WH Freeman Company Press, New York, pp 522-524.

Hackett SJ, Kimball RT, Reddy S, Bowie RC, Braun EL, Braun MJ, Chojnowski JL, Cox WA, Han KL, Harshman J, *et al.* (2008) A phylogenomic study of birds reveals their evolutionary history. Science 320:1763-1768.

Haring E, Kruckenhauser L, Gamauf A, Riesing MJ and Pinsker W (2001) The complete sequence of the mitochondrial genome of *Buteo buteo* (Aves, Accipitridae) indicates an early split in the phylogeny of raptors. Mol Biol Evol 18:1892-1904.

Härlid A, Janke A and Arnason U (1998) The complete mitochondrial genome of *Rhea americana* and early avian divergences. J Mol Evol 46:669-679.

Hedges SB and Sibley CG (1994) Molecules *vs.* morphology in avian evolution: The case of the "pelecaniform'' birds. Proc Natl Acad Sci USA 91:9861-9865.

Howard R and Moore A (1980) A Complete Checklist of the Birds of the World. Oxford University Press, Oxford, pp 1-12.

Ingman M, Kaessmann H, Paabo S and Gyllensten U (2000) Mitochondrial genome variation and the origin of modern humans. Nature 408:708-713.

Kahl MP (1972) A revision of the family Ciconiidae (Aves). J Zool 167:451-461.

Kan XZ, Li XF, Zhang LQ, Chen L, Qian CJ, Zhang XW and Wang L (2010) Characterization of the complete mitochondrial genome of the rock pigeon, *Columba livia* (Columbiformes, Columbidae). Genet Mol Res 9:1234-1249.

Kjer KM and Honeycut RL (2007) Site specific rates of mitochondrial genomes and the phylogeny of Eutheria. BMC Evol Biol 7:e8.

Krajewski C, Sipiorski JT and Anderson FE (2010) Complete mitochondrial genome sequences and the phylogeny of cranes (Gruiformes, Gruidae). Auk 127:440-452.

L'abbé D, Duhaime JF, Lang BF and Morais R (1991) The transcription of DNA in chicken mitochondria initiates from one major bidirectional promoter. J Biol Chem 266:10844-10850.

McCracken KG and Sheldon FH (1997) Avian vocalizations and phylogenetic signal. Proc Natl Acad Sci USA 94:3833-3836.

Mindell DP, Sorenson MD and Dimcheff DE (1998) An extra nucleotide is not translated in mitochondrial ND3 of some birds and turtles. Mol Biol Evol 15:1568-1571.

Monroe BL and Sibley CG (1993) A World Checklist of Birds. Yale University Press, New Haven, 400 pp.

Moum T, Johansen S, Erikstad KE and Piatt JF (1994) Phylogeny and evolution of the auks (subfamily Alcinae) based on mitochondrial DNA sequences. Proc Natl Acad Sci USA 91:7912-7916.

Nesbø CL, Arab MO and Jakobsen KS (1998) Heteroplasmy, length and sequence variation in the mtDNA control regions of three percid fish species (*Perca fluviatilis*, *Acerina cernua*, *Stizostedion lucioperca*). Genetics 148:1907-1919.

Oh DJ, Oh BS, Jung MM and Jung YH (2010) Complete mitochondrial genome of three *Branchiostegus* (Perciformes, Malacanthidae) species: Genome description and phylogenetic considerations. Mitochondrial DNA 21:151-159.

Ojala D, Montoya J and Attardi G (1981) tRNA punctuation model of RNA processing in human mitochondria. Nature 290:470-474.

Pacheco MA, Battistuzzi FU, Lentino M, Aguilar RF, Kumar S and Escalante AA (2011) Evolution of modern birds revealed by mitogenomics: Timing the radiation and origin of major orders. Mol Biol Evol 28:1927-1942.

Payne RB and Risley CJ (1976) Systematics and evolutionary relationships among the herons (Ardeidae). Misc Publ Univ Mich Mus Zool 150:1-115.

Perna NT and Kocher TD (1995) Patterns of nucleotide composition at fourfold degenerate sites of animal mitochondrial genomes. J Mol Evol 41:353-358.

Peter JL (1931) Checklist of Birds of the World. Harvard University Press, Cambridge, 110 pp.

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

Quinn TW and Wilson AC (1993) Sequence evolution in and around the mitochondrial control region in birds. J Mol Evol 37:417-425.

Rabosky DL and Matute DR (2013) Macroevolutionary speciation rates are decoupled from the evolution of intrinsic reproductive isolation in *Drosophila* and birds. Proc Natl Acad Sci USA 110:15354-15359.

Randi E and Lucchini V (1998) Organization and evolution of the mitochondrial DNA control region in the avian genus *Alectoris*. J Mol Evol 47:449-462.

Randi E, Lucchini V, Armijo-Prewitt T, Kimball RT, Braun EL and Ligon JD (2000) Mitochondrial DNA phylogeny and speciation in the tragopans. Auk 117:1003-1015.

Reyes A, Gissi C, Catzeflis F, Nevo E, Pesole G and Saccone C (2004) Congruent mammalian trees from mitochondrial and nuclear genes using Bayesian methods. Mol Biol Evol 21:397-403.

Ritchie PA and Lambert DM (2000) A repeat complex in the mitochondrial control region of *Adelie penguins* from Antarctica. Genome 43:613-618.

Roques S, Godoy JA, Negro JJ and Hiraldo F (2004) Organization and variation of the mitochondrial control region in two vulture species, *Gypaetus barbatus* and *Neophron percnopterus*. J Hered 95:332-337.

Russell RD and Beckenbach AT (2008) Recoding of translation in turtle mitochondrial genomes: Programmed frameshift mutations and evidence of a modified genetic code. J Mol Evol 67:682-695.

Saccone C, Pesole G and Sbisa E (1991) The main regulatory region of mammalian mitochondrial DNA: Structure-function model and evolutionary pattern. J Mol Evol 33:83-91.

Sato A, O'hUigin C, Figueroa F, Grant PR, Grant BR, Tichy H and Klein J (1999) Phylogeny of Darwin's finches as revealed by mtDNA sequences. Proc Natl Acad Sci USA 96:5101-5106.

Sbisà E, Tanzariello F, Reyes A, Pesole G and Saccone C (1997) Mammalian mitochondrial D-loop region structural analysis: Identification of new conserved sequences and their functional and evolutionary implications. Gene 205:125-140.

Shadel GS and Clayton DA (1997) Mitochondrial DNA maintenance in vertebrates. Annu Rev Biochem 66:409-435.

Sheldon FH (1987) Rates of single-copy DNA evolution in herons. Mol Biol Evol 4:56-69.

Sheldon FH and Kinnarney M (1993) The effects of sequence removal on DNA-hybridization estimates of distance, phylogeny, and rates of evolution. Syst Biol 42:32-48.

Sheldon FH, McCracken KG and Stuebing KD (1995) Phylogenetic relationships of the zigzag heron (*Zebrilus undulatus*) and white-crested bittern (*Tigriornis leucolophus*) estimated by DNA-DNA hybridization. Auk 112:672-679.

Sheldon FH, Jones CE and McCracken KG (2000) Relative patterns and rates of evolution in heron nuclear and mitochondrial DNA. Mol Biol Evol 17:437-450.

Shen YY, Shi P, Sun YB and Zhang YP (2009) Relaxation of selective constraints on avian mitochondrial DNA following the degeneration of flight ability. Genome Res 19:1760-1765.

Shi Y, Shan X, Li J, Zhang X and Zhang H (2002) Sequence and organization of the complete mitochondrial genome of the Indian muntjac (*Muntiacus muntjak*). Acta Zool Sin 49:629-636.

Sibley CG and Monroe BL (1990) Distribution and Taxonomy of Birds of the World. Yale University Press, New Haven and London, pp 302-310.

Sibley CG, Ahlquist JE and Monroe BL (1988) A classification of the living birds of the world based on DNA-DNA hybridization studies. Auk 105:409-423.

Song GH, Lin Q, Yue WB, Liu TF and Hu SN (2012) Sequence analysis of the complete mitochondrial genome and molecular evolution of *Cricetulus griseus*. Acta Lab Anim Sci Sin 20:70-75 (in Chinese with English abstract).

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

Southern SO, Southern PJ and Dizon AE (1988) Molecular characterization of a cloned dolphin mitochondrial genome. J Mol Evol 28:32-42.

Sullivan J and Joyce P (2005) Model selection in phylogenetics. Annu Rev Ecol Evol Syst 36:445-466.

Suzuki H, Nunome M, Kinoshita G, Aplin KP, Vogel P, Kryukov AP, Jin ML, Han SH, Maryanto I, Tsuchiya K, *et al.* (2013) Evolutionary and dispersal history of Eurasian house mice *Mus musculus* clarified by more extensive geographic sampling of mitochondrial DNA. Heredity 111:375-390.

Taanman JW (1999) The mitochondrial genome: Structure, transcription, translation and replication. Biochim Biophys Acta 1410:103-123.

Tamura K, Peterson D, Peterson N, Stecher G, Nei M and Kumar S (2011) MEGA5: Molecular Evolutionary Genetics Analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol Biol Evol 28:2731-2739.

Walberg MW and Clayton DA (1981) Sequence and properties of the human KB cell and mouse L cell D-loop regions of mitochondrial DNA. Nucleic Acids Res 9:5411-5421.

Wang C, Chen Q, Lu G, Xu J, Yang Q and Li S (2008) Complete mitochondrial genome of the grass carp (*Ctenopharyngodon idella*, Teleostei): Insight into its phylogenic position within Cyprinidae. Gene 424:96-101.

Wenink PW, Baker AJ and Tilanus MG (1994) Mitochondrial control-region sequences in two shorebird species, the Turnstone and the Dunlin, and their utility in population genetic studies. Mol Biol Evol 11:22-31.

Wolstenholme DR (1992) Animal mitochondrial DNA: Structure and evolution. Int Rev Cytol 141:173-216.

Wu X, Wang Y, Zhou K, Zhu W, Nie J and Wang C (2003) Complete mitochondrial DNA sequence of Chinese alligator, *Alligator sinensis*, and phylogeny of crocodiles. Chin Sci Bull 48:2050-2054.

Yamamoto Y, Murata K, Matsuda H, Hosoda T, Tamura K and Furuyama JI (2000) Determination of the complete nucleotide sequence and haplotypes in the D-loop region of the mitochondrial genome in the oriental white stork, *Ciconia boyciana*. Genes Genet Syst 75:25-32.

Yang R, Wu X, Yan P, Su X and Yang B (2010) Complete mitochondrial genome of *Otis tarda* (Gruiformes, Otididae) and phylogeny of Gruiformes inferred from mitochondrial DNA sequences. Mol Biol Rep 37:3057-3066.

Zhang P and Wake DB (2009) Higher-level salamander relationships and divergence dates inferred from complete mitochondrial genomes. Mol Phylogenet Evol 53:492-508.

Zhang L, Wang L, Gowda V, Wang M, Li X and Kan X (2012) The mitochondrial genome of the Cinnamon Bittern, *Ixobrychus cinnamomeus* (Pelecaniformes, Ardeidae): Sequence, structure and phylogenetic analysis. Mol Biol Rep 39:8315-8326.

Zheng ZX (1997) Fauna Sinica Aves. Vol. 1. Science Press, Beijing, pp 138-140.

Zheng GM (2002) A Checklist on the Classification and Distribution of the Birds of the WorId. Science Press, Beijing, 11 p.

Zhong D, Zhao GJ, Zhang ZS and Xu AL (2002) Advance in the entire balance and local unbalance of base distribution in genome. Hereditas 24:351-355.

Zwickl DJ and Hillis DM (2002) Increased taxon sampling greatly reduces phylogenetic error. Syst Biol 51:588-598.

Supplementary Material

The following online material is available for this article:

Table S1 - Species examined in this study.

Table S2 - Codon usage in the mitochondrial genome of E. garzetta.

Table S3 - Pairwise distances of 11 species inferred from the mitochondrial genome.

This material is available as part of the online article from http://www.scielo.br/gmb.

Associate Editor: Houtan Noushmehr