

RESEARCH ARTICLE

# The mitochondrial genome of pin-tailed snipe *Gallinago stenura*, and its implications for the phylogeny of Charadriiformes

Chaochao Hu<sup>1</sup>✉, Chenling Zhang<sup>2</sup>✉, Lei Sun<sup>3</sup>, Yi Zhang<sup>4</sup>, Wenli Xie<sup>4</sup>, Baowei Zhang<sup>5</sup>, Qing Chang<sup>4</sup>\*

**1** Analytical and Testing Center, Nanjing Normal University, Nanjing, Jiangsu, People's Republic of China, **2** Faculty of Life Science and Chemical Engineering, Jiangsu Second Normal University, Nanjing, Jiangsu, People's Republic of China, **3** Department of Biological Engineering, Utah State University, 4105 Old Main Hill, Logan, UT, United States of America, **4** Jiangsu Key Laboratory for Biodiversity and Biotechnology, College of Life Sciences, Nanjing Normal University, Nanjing, Jiangsu, People's Republic of China, **5** Anhui Key Laboratory of Eco-engineering and Bio-technique, School of Life Sciences, Anhui University, Hefei, Anhui, People's Republic of China

✉ These authors contributed equally to this work.

\* [qingchangnj@163.com](mailto:qingchangnj@163.com)



**OPEN ACCESS**

**Citation:** Hu C, Zhang C, Sun L, Zhang Y, Xie W, Zhang B, et al. (2017) The mitochondrial genome of pin-tailed snipe *Gallinago stenura*, and its implications for the phylogeny of Charadriiformes. PLoS ONE 12(4): e0175244. <https://doi.org/10.1371/journal.pone.0175244>

**Editor:** Bi-Song Yue, Sichuan University, CHINA

**Received:** January 4, 2017

**Accepted:** March 22, 2017

**Published:** April 6, 2017

**Copyright:** © 2017 Hu et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** The complete mitogenome of *Gallinago stenura* is deposited in GenBank of NCBI under accession number KY056596.

**Funding:** This research was supported by grants from the Program of National Natural Science Foundation of China (#31071901 and #41271057) and Natural Science Research of Jiangsu Higher Education Institutions of China (#16KJB180012).

**Competing interests:** The authors have declared that no competing interests exist.

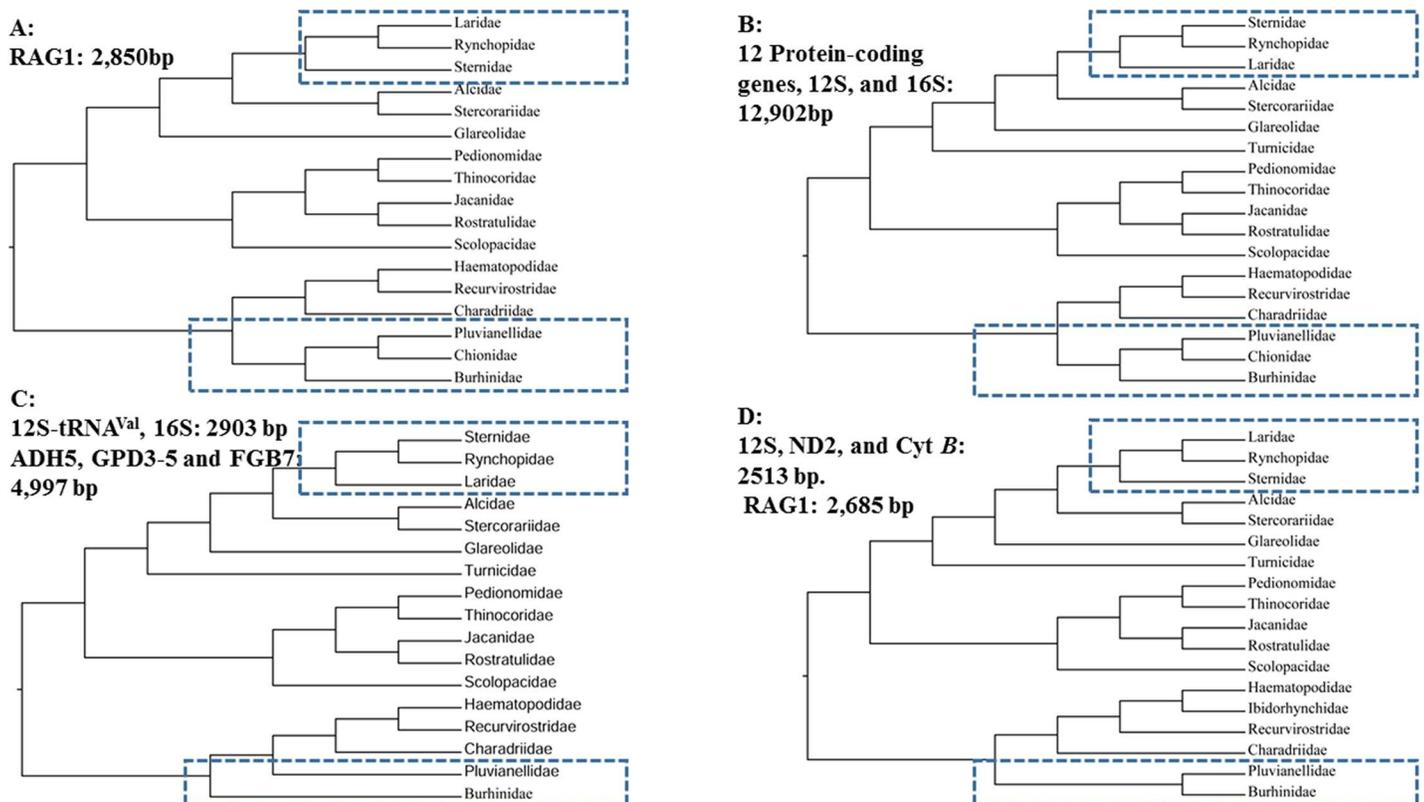
## Abstract

The Charadriiformes, among the most diverse orders of bird, is a good source to research on evolution. The mitochondrial genome sequencing database has rapidly increased in recent years, while Charadriiformes mitogenome has not been well studied. In this research, we determined the complete mitogenome sequence of *Gallinago stenura*, and comparatively analysed 20 mitogenomes of Charadriiformes. The mitogenomes display moderate size variation, and most of variation due to mutations in the control region. In 13 protein-coding genes, we found: 1. The GC skews are always negative, while the negative AT skews are found in 5 genes, 2. The average uncorrected pairwise distances reveal heterogeneity of evolutionary rate for each gene, 3. The ATG and TAA, respectively, are observed the most commonly start and stop codon. The highest dN/dS is detected for ATP8 (0.16) among Charadriiformes, while the lowest for COI (0.01), indicating that 13 protein-coding genes are evolving under the purifying selection. Predicted secondary structures of tRNAs indicate that the sequences and structures of anticodon, amino acceptor, and T $\psi$ C arms are highly conserved, and most nucleotide variation is restricted to dihydrouridine arms with obvious indel polymorphisms. A total of 15 conserved sequence boxes were recognized in the control regions, and the 4 bp (5' -AAAC-3') and 7 bp (5' -AAACAAC-3') repeat sequences occurred frequently. Phylogenomic analysis based on the nearly complete mitochondrial genomes strongly supported the monophyly of the order, and the suborder Charadrii is at the basal of Charadriiformes. Moreover, our results well resolved the complexity family-level relationships and clearly depicted the evolutionary processes of Charadriiformes, based on 12 mitochondrial protein-coding genes from 18 families. This study improves our understanding of mitogenomic structure and evolution, which can provide further insights into our understanding of phylogeny and taxonomy in Charadriiformes.

### Introduction

The order Charadriiformes (Vertebrate: Aves), known as shorebirds, is a diverse order containing approximately 384 species belonging to 94 genera of 19 families, and has members distributing in all parts of the world. Because of the diversity in their ecology, morphology, behavior, distribution and other adaptive characters, Charadriiformes provide potential insight into evolutionary forces acting on life-history traits [1–4]. Because of convergent evolution, the use of morphological and biochemical methods to reconstruct Charadriiformes phylogeny is unresolved and poorly supported [5, 6]. Based on the genetic analysis (mitochondrial and/or nuclear genes), the majority of previous analyses revealed that the order can be divided taxonomically into three major clades (Charadrii, Scolopaci, and Lari), and place the Charadrii as sister to the other two groups [7–11]. However, the complexity family-level relationships of Charadriiformes have not been well resolved, especially in the deep nodes of the phylogenetic trees (Fig 1). For example, in the major clade of Lari, many studies suggest that the Sternidae and Rynchopidae are nearest phylogenetic neighbors (Fig 1B and 1C), others place the Sternidae as sister to the other two families (Fig 1A and 1D). Another potential conflict in topology is the major clade of Charadrii, the relationship of Chionidae, Pluvianellidae and Burhinidae (Fig 1).

Animal mitogenome is an approximately 16 kb circular molecules [14]. This genome typically contains 13 protein-coding genes, 2 ribosomal RNAs (12S rRNA and 16S rRNA), 22 transfer RNAs (tRNAs), and a non-coding control region (CR) [15, 16]. The genome-level characteristics of mitogenomes include the base composition, codon usage, gene order



**Fig 1. Family-level phylogenetic relationships of Charadriiformes.** (A) based on RAG-1 [12], (B) based on 14 mitochondrial genes [8], (C) based on 3 mitochondrial genes and 3 nuclear genes [11], and (D) based on 3 mitochondrial genes and 1 nuclear gene [13].

<https://doi.org/10.1371/journal.pone.0175244.g001>

arrangement, tRNA and rRNA gene secondary structures, and the signaling genes control of replication and transcription [14, 17]. Such information has been used to show potential in resolving ancient patterns of evolutionary history, and also for developing conservation management programs for endangered species [8, 18, 19]. In the past decade, the number of sequenced mitogenomes have accumulated and these sequences have been widely used in biological research, especially when studying molecular evolution, genetic structure, and phylogeography [20–27]. Especially, a number of studies performed using mitochondrial genome sequences to elucidate avian phylogeny [21, 22, 28, 29].

Although some studies have used a few mitochondrial and/or nuclear genes in reconstructing family-level phylogeny within Charadriiformes [4, 8, 11–13, 30], mitochondrial genome sequences and structures have proven to be useful for elucidation of evolutionary relationships, and the comprehensive study of the entire mitogenomes within Charadriiformes are still not well studied. Despite recent rapid increases in the available information on bird mitogenomes, only 19 complete mitogenomes sequences have been reported in database of Charadriiformes (Table 1). Of these sequenced mitogenomes, 5 were from species in the suborder Charadrii (Haematopodidae, Recurvirostridae and Charadriidae), 5 in Scolopaci (Jacanidae and Scolopacidae), and 9 in Lari (Laridae, Stercorariidae and Alcidae).

In this study, we sequenced and annotated the complete mitogenome of the pin-tailed snipe (*Gallinago stenura*) from the family of Scolopacidae, and this species is most common migrant snipe. We compared the newly sequenced mitogenome with previously published sequences, we tried to address: (1) features of Charadriiformes mitogenomes, (2) rates and patterns of mitochondrial gene evolution within Charadriiformes, and (3) the adaptation of the Charadriiformes mitogenome to different environment based on dN/dS analysis. And then, we have collected sequences of 12 protein-coding genes from 18 families to assess the family-

**Table 1. List of the complete mitogenomes sequences included in this study.**

Suborder	Family	Species	Accession no.	Size (bp)
Charadrii	Haematopodidae	<i>Haematopus ater</i>	AY074886	16791
	Recurvirostridae	<i>Recurvirostra avosetta</i>	KP757766	16897
	Charadriidae	<i>Vanellus cinereus</i>	KM404175	17074
		<i>Vanellus cinereus</i>	KM873665	17135
		<i>Vanellus vanellus</i>	KM577158	16795
Scolopaci	Jacanidae	<i>Jacana jacana</i>	KJ631049	16975
		<i>Jacana spinosa</i>	KJ631048	17079
	Scolopacidae	<i>Scolopax rusticola</i>	KM434134	16984
		<i>Arenaria interpres</i>	AY074885	16725
		<i>Eurynorhynchus pygmeus</i>	KP742478	16707
		<i>Gallinago stenura</i>	KY056596 <sup>a</sup>	16899
Lari	Laridae	<i>Chroicocephalus brunnicephalus</i>	JX155863	16769
		<i>Chroicocephalus ridibundus</i>	KM577662	16807
		<i>Chroicocephalus saundersi</i>	JQ071443	16725
		<i>Larus crassirostris</i>	KM507782	16746
		<i>Larus dominicanus</i>	AY293619	16701
		<i>Ichthyaetus relictus</i>	KC760146	16586
		<i>Sternula albifrons</i>	KT350612	16357
		Stercorariidae	<i>Stercorarius maccormicki</i>	KM401546
	Alcidae	<i>Synthliboramphus antiquus</i>	AP009042	16730

<sup>a</sup> The mitogenome sequence (KY056596) was obtained in this study.

level relationships of Charadriiformes. This study contributes to more molecular information on mitogenomic structure and evolution, which can provide further insights into our understanding of phylogeny and taxonomy in Charadriiformes.

## Materials and methods

### Ethical statement

No specific permits were required for the specimens collected for this study. The specimen was common in China and the field studies did not involve endangered or protected species. Our experimental procedures complied with the current laws on animal welfare and research in China, and were specifically approved by the Animal Research Ethics Committee of Nanjing Normal University.

### Sample collection and DNA extraction

The muscle tissue sample of *G. stenura* was collected from a derelict and abandoned mist net in Rudong, Jiangsu Province, China (32°32'43.41" N, 121°06'09.01" E). The tissue was preserved in absolute ethanol and stored at -80°C in our laboratory at Nanjing Normal University (specimen voucher: NJNU-Gst01). Total genomic DNA was extracted using standard phenol-chloroform methods [31].

### Primer design, PCR amplification and sequencing

The primers were designed based on sequence-conserved regions, which were identified using multiple alignments of the complete mitogenomes of 19 closely related sequences downloaded from GenBank (Table 1). Using the primer pairs newly designed for this study, PCR amplifications were carried out in a final reaction volume of 25 µL, which contained 10 × buffer 3 µL, 25 mmol/L MgCl<sub>2</sub> 2 µL, 2 mmol/L dNTPs 2 µL, 10 µM of each primer (forward and reverse) 1 µL, 1 µL template DNA, 1U Taq DNA polymerase (Takara, Japan), and sterile double-distilled water (ddH<sub>2</sub>O) to make up a final volume of 25 µL. Amplifications were performed in an Applied Biosystems 2720 thermal cycler with heated lid under the following conditions: 95°C for 5 min (initial denaturation); then 35 cycles of 95°C for 30 s (denaturation), 50–55°C for 30 s (annealing), and 72°C for 1 min (extension); and a final extension at 72°C for 8 min and 4°C hold. Annealing temperature was changed from 50 to 55°C to improve the quality of PCR products as necessary. PCR products were excised by electrophoresis on a 1% agarose gel and visualized by Ethidium Bromide staining. And then the PCR products were purified using a gel extraction kit (Promega) and sequenced with each of the PCR primers on an ABI 377 sequencer.

### Genome annotation and sequence analysis

Sequences obtained were assembled and edited using the software SeqMan (DNASTar, Inc.) to generate complete mitochondrial DNA sequences. The protein-coding genes were determined by open reading frame finder implemented at the NCBI website with the vertebrate mitochondrial genetic code, and then finally confirmed by sequence comparisons with the reported Charadriiformes mitogenomes. The tRNAscan-SE 1.21 [32], MITOS [33], and ARWEN [34] were used to confirm tRNA annotation results. The skewing of the nucleotide composition was calculated according to the following formulas: AT skew = (A - T) / (A + T) and GC skew = (G - C) / (G + C) [35, 36]. The tandem repeats were searched in the CR using the Tandem Repeats Finder program [37].

The number of variable sites, the parsimony informative sites, the singleton, and the average uncorrected pairwise distances for each protein-coding gene were calculated by MEGA 6.0 [38]. The rates of non-synonymous substitutions ( $K_a$ ,  $\pi$  modified), and synonymous substitutions ( $K_s$ ,  $\pi$  modified) for each protein-coding genes were determined with DnaSP 5.0 [39]. The ratio of nonsynonymous substitution rate to synonymous substitution rate ( $\omega = dN/dS$ ) was calculated with Datamonkey [40], by choosing the vertebrate mitochondrial DNA genetic code, with in the SLAC method ( $P < 0.1$ ). The ratio of nonsynonymous substitution rate ( $dN$ ) to synonymous substitution rate ( $dS$ ) is widely used as an indicator of selective pressure at the sequence level among different species. It is commonly accepted that  $\omega > 1$ ,  $\omega = 1$ , and  $\omega < 1$  generally indicate positive selection (replacement substitutions increase fitness), neutral mutation, and negative selection (remove such substitutions from the gene pool), respectively [41].

## Phylogenetic analysis

Two datasets were assembled for phylogenetic analyses: (1) the first dataset (Dataset 1) for mitogenomic phylogeny of Charadriiformes, 13 protein-coding genes, 12S and 16S, of 20 Charadriiformes species were used (Table 1), (2) the second dataset (Dataset 2) for the family-level phylogeny of Charadriiformes, 12 protein-coding genes (with excluding ND6) of 40 species were used (S1 Table). The two aligned datasets were 13,965 and 10,773 bp in length for the Dataset 1 and Dataset 2 matrices, respectively. In all phylogenetic analyses, *Columba livia*, *Gallus gallus*, *Phodilus badius*, and *Zenaida auriculata* were used as outgroups [10, 12, 42, 43]. Before reconstructing the phylogenetic trees, sequence alignment was carried out. Each mitochondrial gene was aligned individually using Muscle in MEGA 6.0 [38]. The 13 protein-coding genes were translated into amino acids, and then aligned using Muscle in MEGA 6.0 [38] with default parameters for each gene, and finally retranslated into nucleotide sequences after removing the stop codons.

The phylogenetic trees were inferred under the maximum likelihood (ML) and Bayesian inference (BI), which were performed with RAxML 8.2.4 [44] and MrBayes 3.2.2 [45], respectively. To determine the optimal partitioning of the data, we explored three partition schemes in phylogenetic analyses. The data set was partitioned *a priori* on the basis of genes and codon positions. The three partitioning schemes are: (1) single (not partitioning), (2) each gene as one partition, (3) partitioned by each of the three codon positions in each gene. The best-fit partitioning scheme and the most appropriate nucleotide evolution model for each partition were implemented in PartitionFinder 1.1.1 [46]. Unpartitioned scheme using a GTR+ I + G substitution model. For ML analysis, node support was calculated with a GTRGAMMA model via rapid bootstrapping (-f a -x option) with ten runs and 1,000 replications to estimate the best topology. For the BI analysis, two independent runs of four Markov Chains Monte Carlo (MCMC) chains (one cold chain and three hot chains) were simultaneously run for  $2.0 \times 10^7$  generations, with sampling conducted every 1000 generations. The first 25% of the initial trees were discarded as “burn-in” and the remaining trees were used to calculate 50% majority rule consensus tree and Bayesian posterior probabilities. The MCMC runs were repeated twice to confirm consistent approximation of the posterior parameter distributions. The convergence of MCMC runs and effective sample sizes ( $ESS > 200$ ) were checked by plotting the log likelihood scores against the generation times using the program Tracer 1.6 (<http://beast.bio.ed.ac.uk/Tracer>).

## Results and discussion

### Genome organization

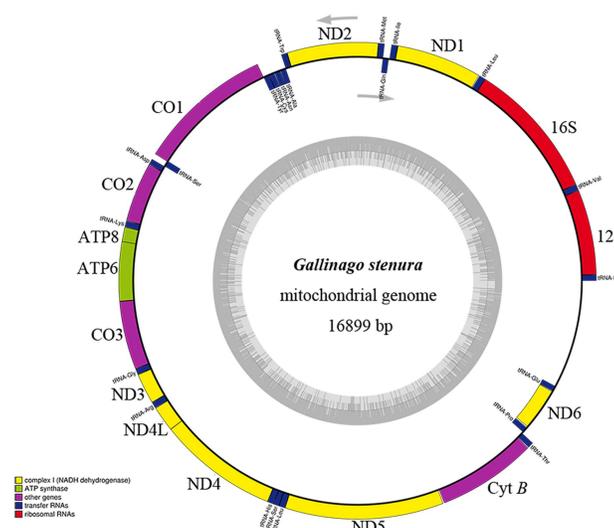
The mitogenome of *G. stenura* is a closed circular molecule of 16,899 bp in length (GenBank No. KY056596), and its map is shown in Fig 2. The completely sequenced mitogenomes from

Charadriiformes encode a CR, and a typical set of 37 mitochondrial genes containing 13 protein-coding genes, 2 rRNA genes (12S rRNA and 16S rRNA), and 22 tRNA genes. Gene arrangement is highly conserved within the Charadriiformes mitogenomes, without showing any structural rearrangement [10, 14]. The gene of ND6 and 8 tRNA are encoded on the light strand, whereas the other genes are located on the heavy strand (S2 Table). The mitogenomes of the twenty completely sequenced Charadriiformes species displayed moderate size variation, the mean size was 16,807 bp (SD = 179.66, n = 20), ranging from 16,357 bp (*Sternula albifrons*) to 17,135 bp (*Vanellus cinereus*) (Table 1). Most of the size variation primarily due to length mutation in the CR. In the *G. stenura* mitogenome, gene overlaps occurred 10 times, spanning 1–10 nucleotides, for a total of 35 nucleotides. The longest overlap (10 bp) existed between ATP8 and ATP6, and the pairs of COI/tRNA<sup>Ser</sup> overlap 9 nucleotides, whereas 5 pairs (16srRNA/tRNA<sup>Leu</sup>, tRNA<sup>Gln</sup>/tRNA<sup>Met</sup>, tRNA<sup>Cys</sup>/tRNA<sup>Tyr</sup>, ATP6/COIII and tRNA<sup>Ser</sup>/tRNA<sup>Leu</sup>) overlap only one nucleotide (S2 Table). The intergenic spacer regions occurred 17 times, spanning 1–12 bp, for a total of 69 bp (S2 Table).

### Nucleotide composition

The overall mean base composition of Charadriiformes mitogenomes was: A, 31.20%; C, 30.60%; T, 24.44% and G, 13.76%. Twenty Charadriiformes mitogenomes are consistently biased towards AT rich, ranging from 54.39% (*Chroicocephalus saundersi*) to 58.35% (*G. stenura*) (Table 2), which is consistent with previous avian mitogenomes [22]. Interestingly, higher A + T content is observed not only in the CR but also in protein-coding genes and rRNA (Fig 3). The lowest A + T content is found in the gene of 12S rRNA (19 species) or ND1 (only in *Scolopax rusticola*), whereas the highest A + T content is found in ATP8 (11 species) or CR (9 species).

AT and GC skews are a measure of compositional asymmetry. In Charadriiformes mitogenomes, AT skews values were always positive, while the values of GC skew were negative. The AT skew value observed is  $0.12 \pm 0.02$  (mean  $\pm$  SD), ranging from 0.09 (*Sternula albifrons*) to 0.15 (*Recurvirostra avosetta*). The GC skews value is  $-0.38 \pm 0.01$ , ranging from  $-0.40$  (*Jacana spinose*) to  $-0.36$  (*S. albifrons*) (Table 2). The lowest value of GC skew was always found in Jacanidae ( $-0.40$ ), and highest value was found in Laridae ( $-0.36$  to  $-0.37$ ). In



**Fig 2. Circular map of the mitochondrial genome of *Gallinago stenura*.**

<https://doi.org/10.1371/journal.pone.0175244.g002>

general, AT and GC skews in Charadriiformes mitogenomes are similar to patterns typically found in most animal mitogenomes, which positive AT skew and negative GC skew are found for H-strand, implying the specific bias toward A and C in nucleotide composition [47, 48]. The GC skew in Charadriiformes mitogenomes revealed that the ATP8 containing significantly higher skews than other regions, and the positive GC skew was not found in this work. However, there was a marked negative AT skew in ND1 (11 species), COI (three species), ND3 (five species), ND4L (one species) and CR (four species) on the H-strand (Fig 4).

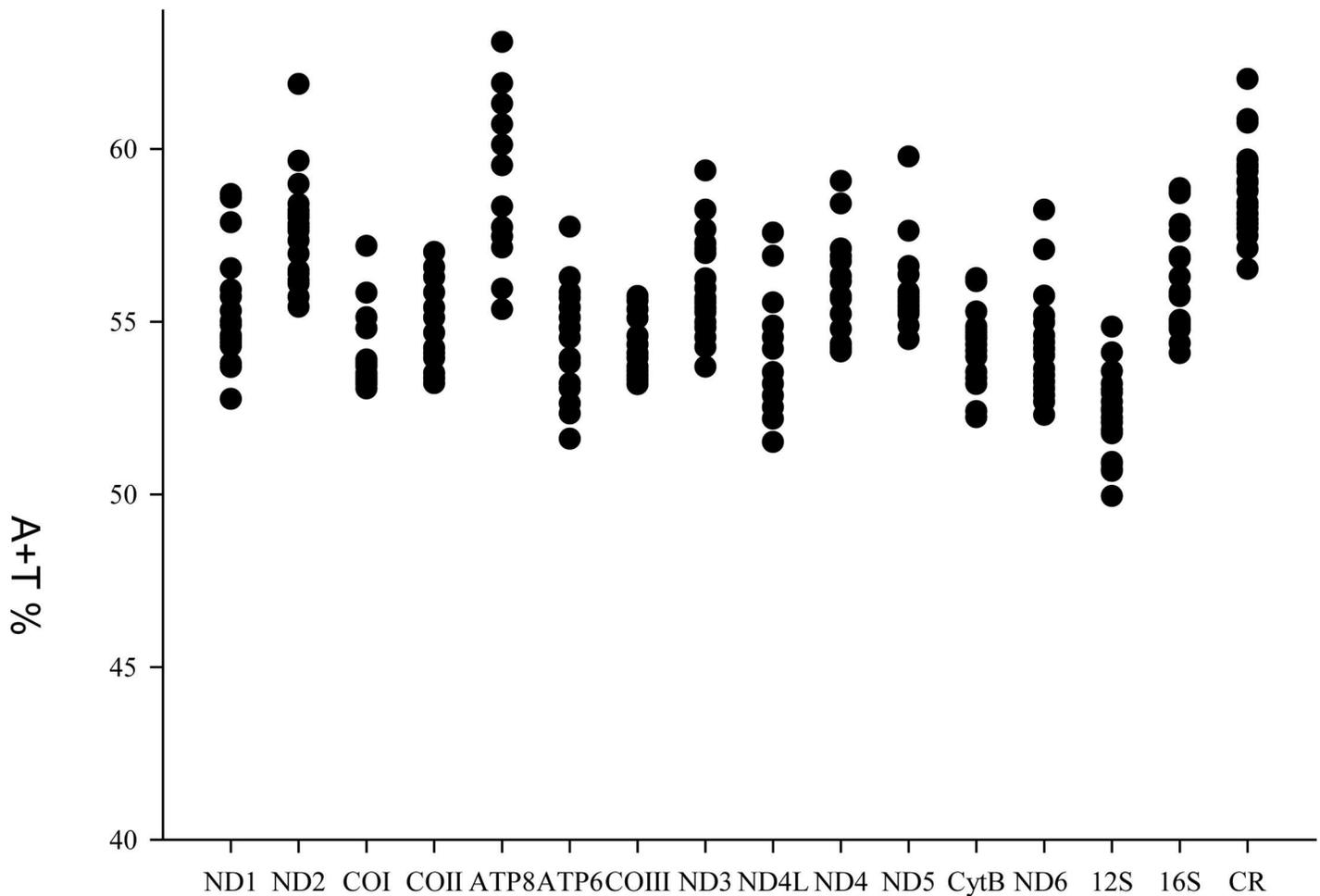
### Protein-coding genes evolutionary rate

Total length of the protein-coding genes in each species was 11,399 bp after removing termination codons and indels. Indels were only found in ATP8, COIII, ND5, and Cyt *b*, and ranged from two to five amino acids. Comparison of each protein-coding gene provides a better understanding of the evolutionary patterns of molecular evolution. The length of the 13 protein-coding genes of Charadriiformes genomes showed that the gene of ND5 and ATP8 being the longest and shortest, respectively. The percent of variable positions in each gene ranged from 33.91% (COI) to 53.57% (ATP8), and the parsimony informative sites ranged from 28.32% (COIII) to 51.6% (ATP8). The analysis of variable sites showed that the ATP8 contained more variable sites. The average uncorrected pairwise distances revealed the heterogeneity of evolutionary rate for each gene, the COIII (0.13), COI (0.14) and COII (0.14) had slow evolutionary rate, whereas those of ND6 (0.20), ND2 (0.20) and ATP8 (0.24) relatively faster (Table 3). So, we can infer that the COIII was the most conserved protein-coding gene, and ATP8 the least conserved. At the nucleotide levels, three genes (COI, COII and COIII) had the lowest evolutionary rates, suggesting that they are useful candidate barcoding markers. The nucleotide composition had a positive correlation with the observed variation in each gene,

**Table 2. Nucleotide composition of mitochondrial genomes in Charadriiformes.**

Species	Accession no.	Proportion of nucleotides (%)				AT content	AT skew	GC skew
		A	T	G	C			
<i>Haematopus ater</i>	AY074886	31.59	23.62	13.67	31.12	55.21	0.14	-0.39
<i>Recurvirostra avosetta</i>	KP757766	31.72	23.59	13.56	31.13	55.31	0.15	-0.39
<i>Vanellus cinereus</i>	KM404175	31.63	23.53	13.77	31.08	55.15	0.15	-0.39
<i>Vanellus cinereus</i>	KM873665	31.46	23.69	13.98	30.87	55.14	0.14	-0.38
<i>Vanellus vanellus</i>	KM577158	31.44	24.03	13.76	30.77	55.47	0.13	-0.38
<i>Jacana jacana</i>	KJ631049	31.88	24.35	13.15	30.62	56.23	0.13	-0.40
<i>Jacana spinosa</i>	KJ631048	31.54	24.86	13.09	30.42	56.45	0.12	-0.40
<i>Scolopax rusticola</i>	KM434134	31.79	25.02	13.34	29.85	56.81	0.12	-0.38
<i>Arenaria interpres</i>	AY074885	30.64	24.71	13.94	30.72	55.34	0.11	-0.38
<i>Eurynorhynchus pygmeus</i>	KP742478	31.29	24.85	13.84	30.02	56.14	0.11	-0.37
<i>Gallinago stenura</i>	KY056596	32.21	26.14	12.88	28.77	58.35	0.10	-0.38
<i>Chroicocephalus brunnicephalus</i>	JX155863	30.70	24.03	14.16	31.11	54.73	0.12	-0.37
<i>Chroicocephalus ridibundus</i>	KM577662	30.81	23.97	14.16	31.06	54.78	0.12	-0.37
<i>Chroicocephalus saundersi</i>	JQ071443	30.41	23.98	14.39	31.21	54.39	0.12	-0.37
<i>Larus crassirostris</i>	KM507782	30.62	24.34	14.13	30.91	54.96	0.11	-0.37
<i>Larus dominicanus</i>	AY293619	30.54	24.45	14.13	30.88	54.98	0.11	-0.37
<i>Ichthyaetus relictus</i>	KC760146	30.62	24.38	14.07	30.93	55.00	0.11	-0.37
<i>Sternula albifrons</i>	KT350612	31.11	26.12	13.74	29.03	57.22	0.09	-0.36
<i>Stercorarius maccormicki</i>	KM401546	30.94	24.39	13.82	30.85	55.34	0.12	-0.38
<i>Synthliboramphus antiquus</i>	AP009042	31.10	24.75	13.56	30.60	55.85	0.11	-0.39

<https://doi.org/10.1371/journal.pone.0175244.t002>



**Fig 3. The A + T content in Charadriiformes mitogenomes.** Each point represents a species.

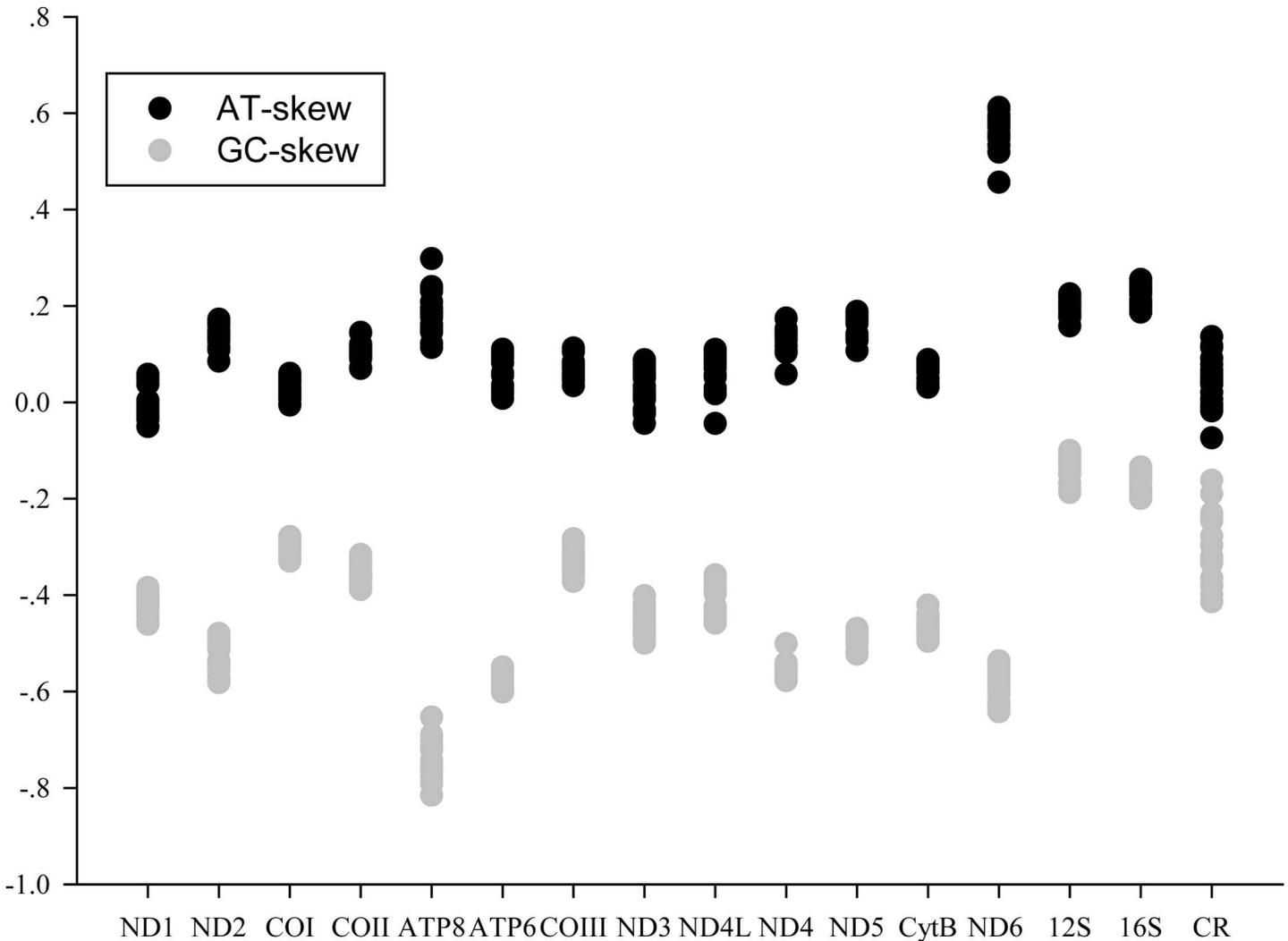
<https://doi.org/10.1371/journal.pone.0175244.g003>

indicating that the larger biased in nucleotide composition might be expected evolving more quickly (Fig 4, Table 3).

To better understand the evolutionary patterns among the 13 protein-coding genes and the role of selection in Charadriiformes species, the values of  $K_a$ ,  $K_s$ , and  $dN/dS$  ( $\omega$ ) were calculated for each protein-coding gene, respectively (Fig 5). The  $K_s$  of ND1 was the highest, while the values of  $K_a$  and  $\omega$  for ATP8 were the highest. The highest  $dN/dS$  was detected for the ATP8 gene (0.16) among Charadriiformes, while the lowest for the COI gene (0.01). The  $dN/dS$  values for all protein-coding genes were far lower than one ( $\leq 0.16$ ), indicating that these genes were evolving under the purifying selection.

### The usage of start and stop codon

Start and stop codons vary among the protein-coding genes in Charadriiformes, although ATG and TAA were the most commonly observed. Start and stop codon usage biases were shown in Fig 6. Five start codons (ATG, GTG, ATT, ATC, and ATA) were detected in the 13 protein-coding genes. The most common start codon was ATG, which accounts for 78.85% of the start codons, followed by GTG (11.15%). The start codon ATG appeared in 12 protein-coding genes (with the exception of ND3), and 8 genes (COII, ATP8, ATP6, COIII, ND4L,



**Fig 4. The AT skew and GC skew in Charadriiformes mitogenomes.** Each point represents a species.

<https://doi.org/10.1371/journal.pone.0175244.g004>

ND4, Cyt *b*, and ND6) only used ATG as start codon. The start codon GTG was commonly used in the COI and ND5. Then the start codons ATT and ATC were only used in the ND3. The start codon ATA were found in the ND1, ND2, and ND3 (Fig 6), which was frequently observed in other avian orders [20]. While in the suborder Lari, the start codons ATC and ATA were not common used, as they were only found in the ND3 of *S. antiquus* and *A. interpres* [10].

Four stop codons (TAA, TAG, AGG, and AGA) and an incomplete stop codon (T-) were detected. The most common stop codon TAA served as stop codon only in six genes (COII, ATP8, ATP6, ND3, ND4L, and Cyt *b*) (Fig 6). In the case of other bird orders, the stop codon TAA appears only in COII, ND4L, and Cyt *b* [20]. The stop codon TAG appeared in the ND2 and ND6. But the stop codon TAG has so far been described in the Cyt *b* of *Aceros corrugatus* (Coraciiformes; Bucerotidae) and the ND6 of a few Columbidae species [20]. The stop codon AGG was found in COI and ND1 (Fig 6). The stop codon AGA was found in ND5, and ND1. The stop codon T- only appeared in the COIII and ND4. Truncated stop codons are common and may be completed by a poly-adenylation of the 3'-end of the mRNA *via* post-transcriptional [49–51].

**Table 3. The mutational information and average distances among Charadriiformes calculated by gene.**

Gene	Length (bp)	%Vs	%Pis	%S	%Aupd
ND1	978	55.01	44.99	37.42	19.00
ND2	1041	51.30	48.70	38.33	20.30
COI	1551	66.09	33.91	28.69	14.40
COII	684	64.04	35.96	30.12	14.40
ATP8	168	46.43	53.57	43.45	23.50
ATP6	684	55.12	44.88	36.11	18.20
COIII	784	65.43	34.57	28.32	13.30
ND3	351	58.40	41.88	35.33	18.00
ND4L	297	53.20	46.80	37.37	17.90
ND4	1378	54.86	45.14	35.70	17.90
ND5	1818	51.76	48.07	37.24	18.80
Cyt <i>b</i>	1143	60.89	39.11	30.53	15.10
ND6	522	50.19	49.81	36.40	19.80

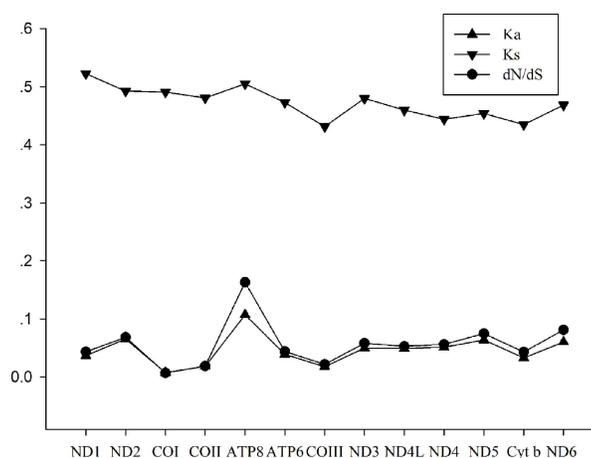
Vs: variable sites; Pis: parsimony informative sites; S: singleton; Aupd: The average uncorrected pairwise distances.

<https://doi.org/10.1371/journal.pone.0175244.t003>

In the gene of ND3, an extra nucleotide (C: cytosine) was present at the position 174 of 16 species, which was not observed in four species (*Chroicocephalus ridibundus*, *C. saundersi*, *Synthliboramphus antiquus* and *Scolopax rusticola*). In the species of *C. saundersi*, Yoon et al. (2015) found the extra base in ND3, whereas Ryu and Kwang (2012) lacked it [10, 52]. The extra base insert in ND3 has been observed in some birds and turtles [53], but it appears to be randomly happened in closely related species, even within species [10, 52]. So, we can infer that there is no evolutionary significance for the extra insertion site.

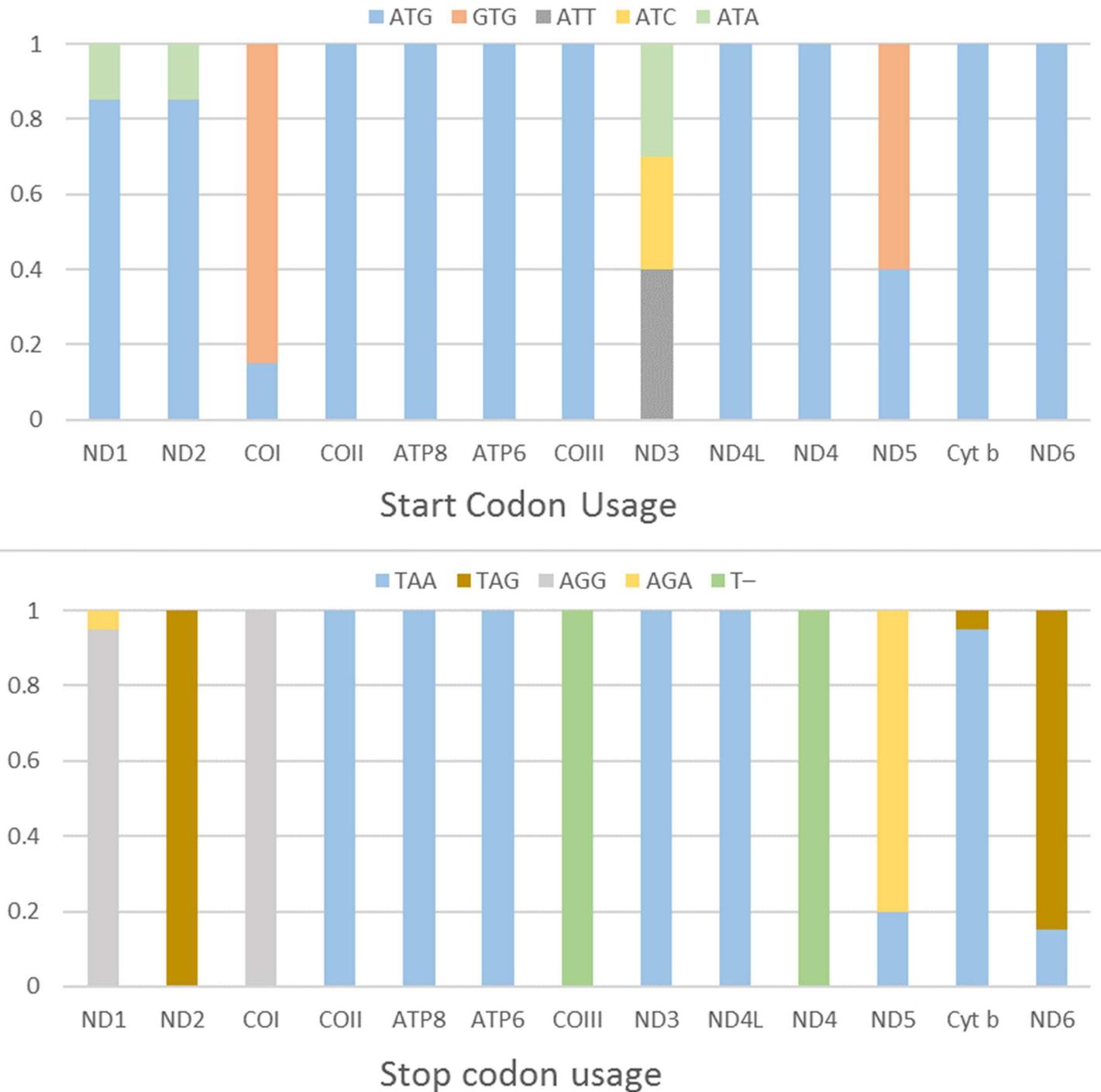
### Sequence features of rRNA and tRNA genes

As other avian rRNA genes, Charadriiformes mitogenomes had 2 rRNA (12S and 16S) that were identified on the heavy strand and located at a conserved position between tRNA<sup>Phe</sup> and



**Fig 5. Evolutionary rates of 13 protein-coding genes in Charadriiformes mitogenomes.** The Y-axis provides the substitution rate of mitochondrial gene. Synonymous nucleotide substitutions per synonymous site (Ks) and nonsynonymous nucleotide substitutions per nonsynonymous site (Ka) are calculated using Dnasp, and dN/dS is calculated using DataMonkey.

<https://doi.org/10.1371/journal.pone.0175244.g005>



**Fig 6. The usage of start and stop codons in the 13 mitochondrial protein-coding genes of the order Charadriiformes.** All genes are shown in the order of occurrence in the mitochondrial genome starting from ND1.

<https://doi.org/10.1371/journal.pone.0175244.g006>

tRNA<sup>Leu</sup> genes, and separated by tRNA<sup>Val</sup> gene (Fig 2; S2 Table). The length of 12S rRNA varied from 963 bp (*Haematopus ater*) to 981 bp (*Vanellus cinereus*) (KM404175), and the smallest and largest 16S rRNA genes were 1,589 bp (*Jacana spinose* and *J. jacana*) and 1,607 bp (*V. vanellus*). The average size of two combined rRNA genes was 2,566 bp ( $\pm 8$ , n = 20), and ranged

from 2,553 bp (*H. ater*) to 2,580 bp (*V. vanellus*). Therefore, there was no substantial size variation between rRNAs, since their sequence size varied slightly with few base pairs difference. In *G. stenura*, the combined size of the two rRNA genes was 2,571 bp.

Twenty-two tRNA genes were identified in Charadriiformes in the same genomic positions as observed in Laridae [10], with an average size ranging from  $65.0 \pm 0.9$  bp in tRNA<sup>Ser (AGY)</sup> to  $74.0 \pm 0.2$  bp in tRNA<sup>Leu (UUR)</sup>. The length difference was due to the mutation of stem and loop sizes of the dihydrouridine (DHU) and T $\Psi$ C arms. In *G. stenura* mitogenome, the length of 22 tRNA genes ranged from 66 bp [tRNA<sup>Ser(AGY)</sup>] to 74 bp [tRNA<sup>Ser</sup> and tRNA<sup>Glu</sup>] (S2 Table). Putative secondary structures of tRNAs in *G. stenura* indicated that 21/22 [except for tRNA<sup>Ser (AGY)</sup>, which appears to lack the DHU arm] possess a typical canonical cloverleaf secondary structure, which were similar to those of other Charadriiformes species, indicated their similar functions (S1 Fig). The sequences and structures of anticodon, amino acceptor, and T $\Psi$ C arms were highly conserved, while the nucleotide variation was mostly restricted to DHU arms with obvious indels polymorphisms (13/22 tRNA genes; S2 Fig).

## Control region

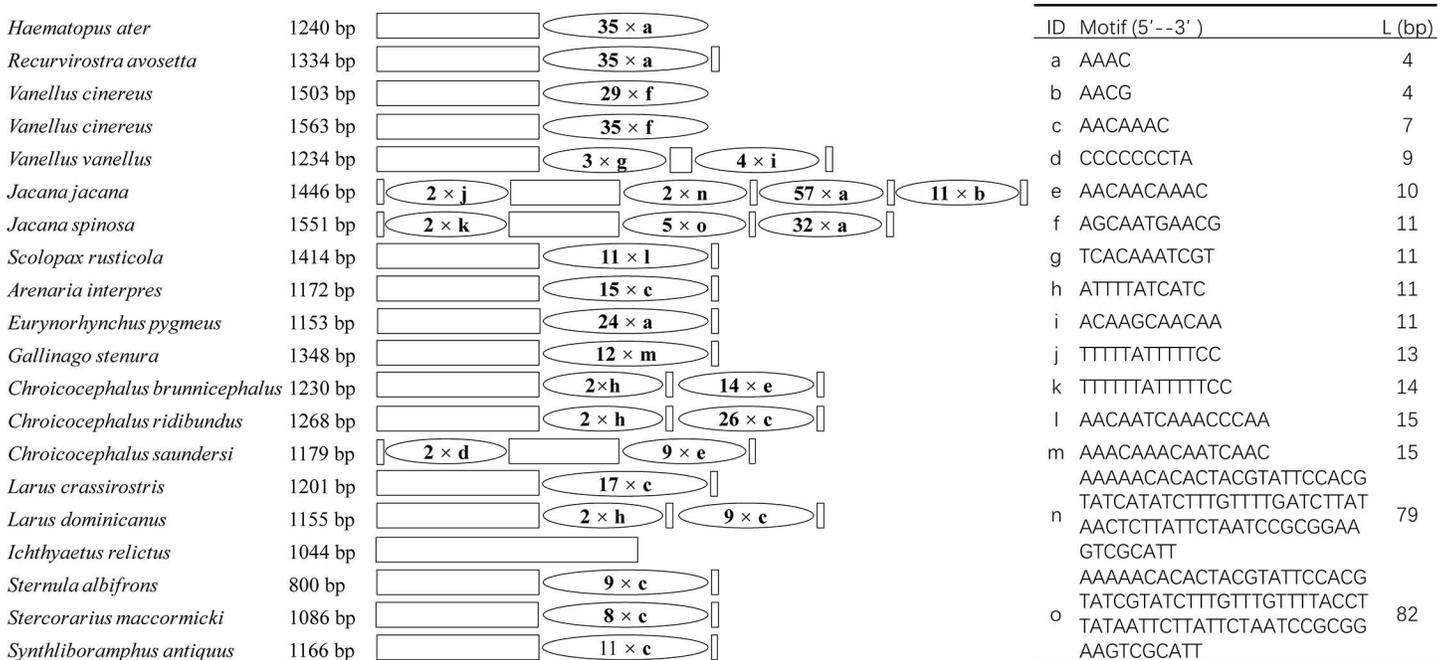
The CR of Charadriiformes mitogenomes were highly divergent, such as the length ranged from 800 bp in *Sternula albifrons* to 1,563 bp in *V. cinereus*. However, the CR duplication was not found in this study. In CR of *G. stenura*, the nucleotide composition was 31.97% A, 28.78% T, 11.80% G, and 27.45% C, and with an A + T bias of 60.76%. The AT skew of CR ranged from -0.07 in *Ichthyaetus relictus* to 0.12 in *V. cinereus*, and there were two species (-0.07 in *I. relictus* and -0.02 in *Larus dominicanus*) with a negative value. The mean value of GC skew was -0.31, ranging from -0.41 in *S. rusticola* to -0.16 in *V. cinereus* (KM873665).

We observed large tandem repeats with two or more copies in CR with the exception of *I. relictus*. In most species (12/20), tandem repeat sequences only had one type of repeat unit. Tandem repeat sequences contained two types of repeat units in 5 species (*V. vanellus*; *C. brunnicephalus*, *C. ridibundus*, *C. saundersi*, and *L. dominicanus*). The sequences contained three and four types in *Jacana spinosa* and *J. jacana* respectively (Fig 7). Total of 15 kinds of tandem repeats were found in CR. The motifs of 5' -AAACAAC-3' (occurs in 7 species) and 5' -AAAC-3' (in 5 species) are most common. However, 10/15 motifs were only found in one species, suggesting the fast rate evolution in the CR. Two long motifs of n (79 nucleotides with 2 tandem repeats) and o (82 nucleotides with 5 tandem repeats) only occurred in *Jacava* (Fig 7). Interestingly, the five motifs (a, c, f, l and m) had large copies (Fig 7), which had a strictly conserved common sequence (5' -AAAC-3'). The CR of Charadriiformes mitogenome shows distinct sequence and structural characteristics, such as varying sizes and diverse tandem repeats, which are taxon-specific and can potentially be used as genetic markers for evolution and population genetic studies.

## Phylogenetic analysis

Phylogenetic analysis with two inference methods (ML and BI) of 13 mitochondrial protein-coding genes, 12S and 16S (Dataset 1) for 20 Charadriiformes species revealed the identical topologies, which were highly supported by bootstrap and posterior probabilities at most nodes (Fig 8). Relationships of the order Charadriiformes based on the nearly complete mtDNA strongly support monophyly of the order Charadriiformes.

The topologies recovered using ML and BI of the combined 12 protein-coding genes (Dataset 2, 10,773 bp after alignment) were consistent (Fig 9). These relationships were well supported, as the posterior probabilities in the Bayesian tree are 1 at all nodes in the tree excepting two nodes, and similarly of the maximum likelihood bootstrap values. The order was divided



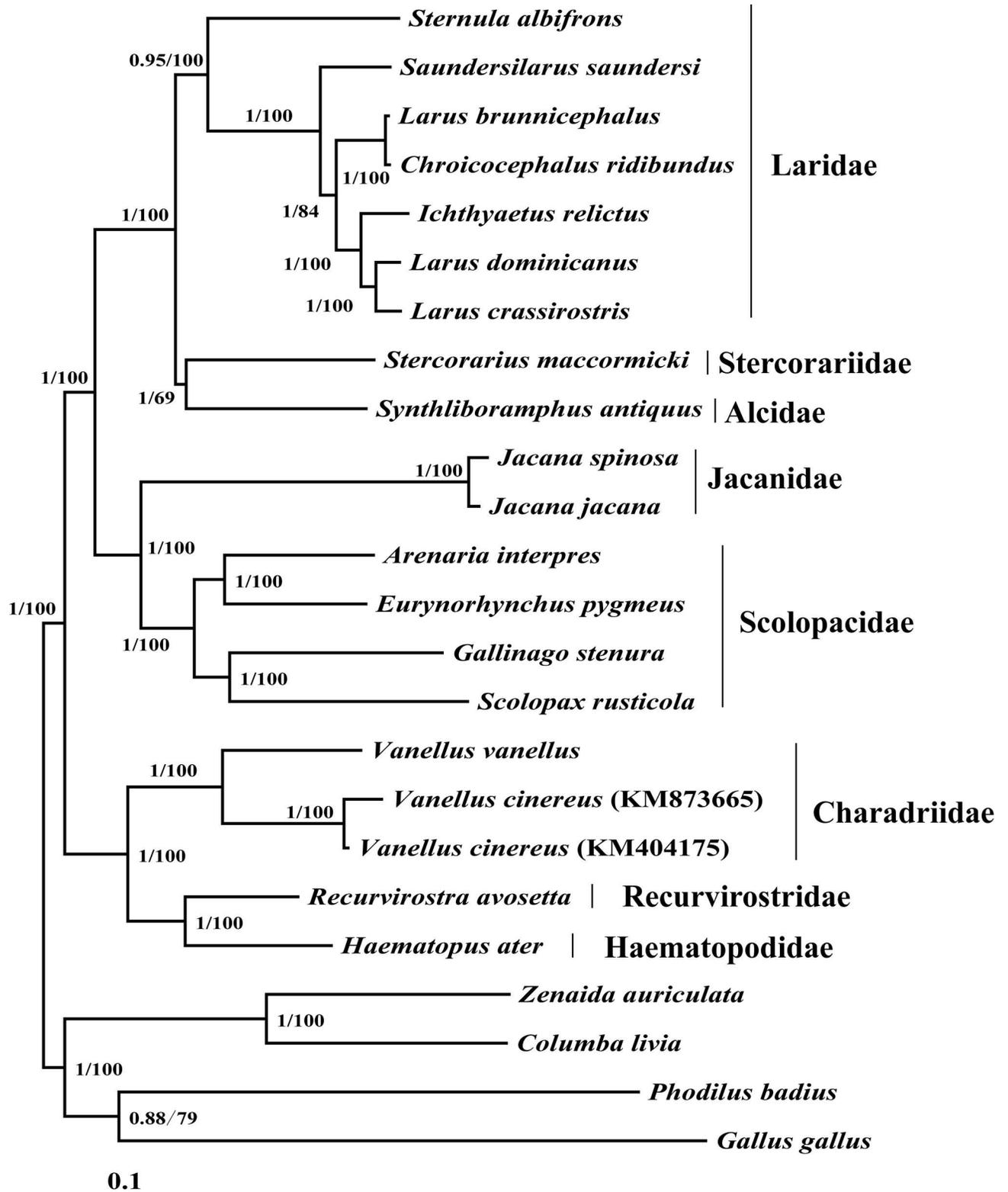
**Fig 7. Organization of the control region in Charadriiformes mitochondrial genomes.** The location and copy number of tandem repeats are shown by oval with Arabic numerals inside. Non-repeat regions are indicated by square box. L: Length.

<https://doi.org/10.1371/journal.pone.0175244.g007>

into three major clades including Lari (Laridae, Sternidae, Rynchopidae, Alcidae, Stercorariidae, Glareolidae, and Turnicidae) and its sister Scolopaci (Scolopacidae, Jacanidae, Rostratulidae, Pedionomidae, and Thinocoridae), which were in turn sister to the suborder Charadrii (Charadriidae, Recurvirostridae, Haematopodidae, Chionidae, Pluvianellidae, and Burhinidae) (Fig 9), as previous analysis based on both nuclear and mitochondrial gene have indicated [8, 11–13, 30]. Within Lari, our phylogenies indicated that Laridae was paraphyletic to Sternidae plus Rynchopidae. This result is in congruence with previous comprehensive phylogenetic analyses in Charadriiformes, using a combination of mitochondrial and nuclear genes [8, 11]. The Alcidae plus Stercorariidae comprise the sister group to Laridae, Sternidae and Rynchopidae. And Glareolidae and Turnicidae are successive sister groups to these clades instead as basal members of the major clade Lari. Within Scolopaci, Rostratulidae and Jacanidae as a sister group to the Pedionomidae plus Thinocoridae, and Scolopacidae is the basal to the other. This result is in congruence with more and more recent studies [11, 13, 54]. Within Charadrii, Burhinidae is sister to Chionidae plus Pluvianellidae, Charadriidae is paraphyletic to Haematopodidae plus Recurvirostridae (Fig 9). According to the phylogenetic results, the monophyly of Charadriiformes was strongly supported in ML and BI analyses.

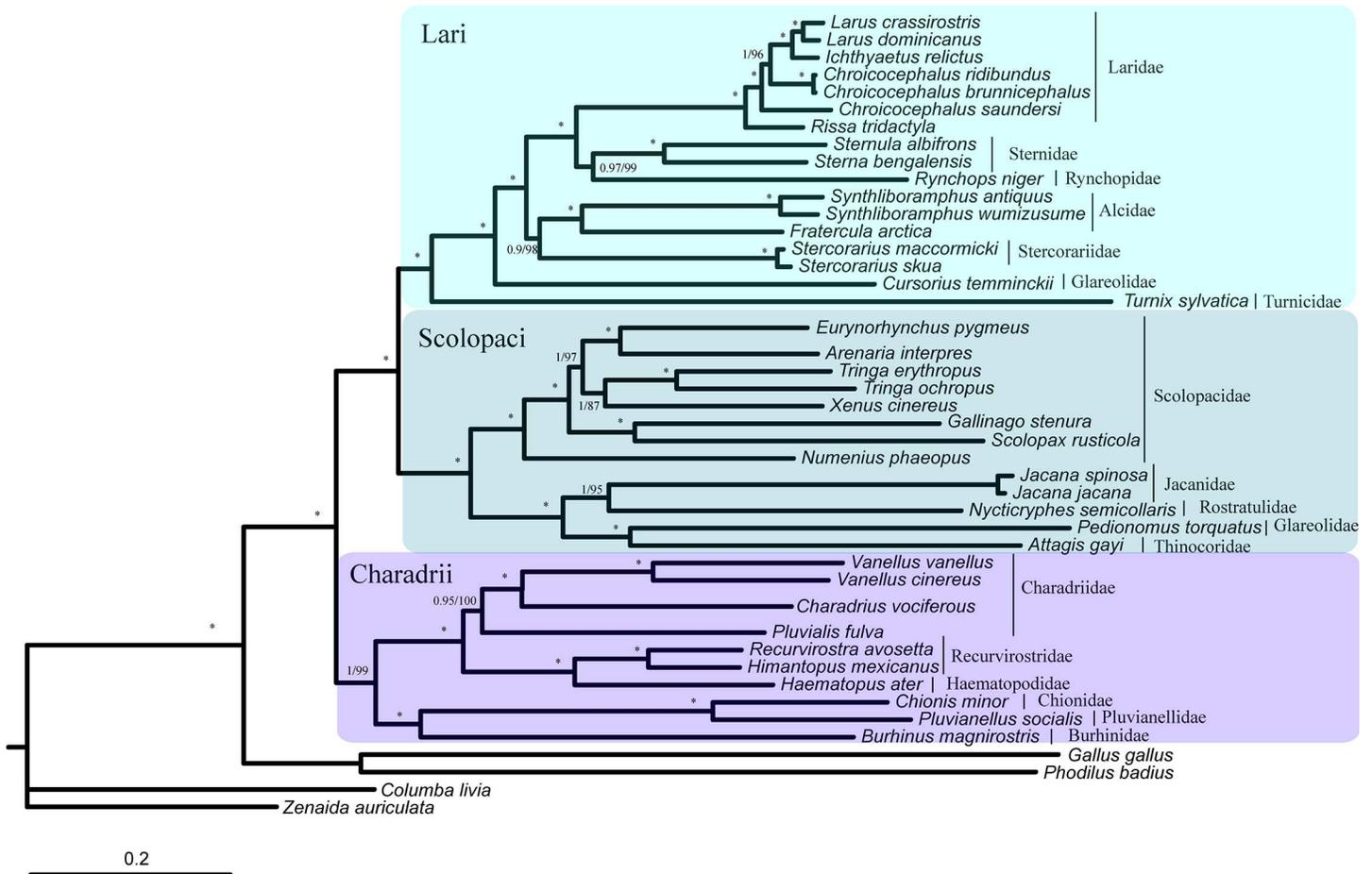
## Conclusion

In this study, we sequenced and annotated the complete mitogenome of *Gallinago stenura*. We compared 20 Charadriiformes mitogenomes to illustrate mitogenomes structure and investigate their evolutionary relationship. The Charadriiformes mitogenomes display moderate size variation, the mean size was 16,807 bp (SD = 179.66, n = 20), and most of the size variation due to mutations in the CR. Nucleotide composition is consistently biased towards AT rich, and the A + T content also varies for each protein-coding genes, such as the variation in ATP8 is the highest, whereas the variation in COIII is the lowest. In the 16 gene regions, the GC skew



**Fig 8. The phylogenetic trees of Charadriiformes based on the 13 mitochondrial protein-coding genes, 12S and 16S dataset using BI and ML methods.** Branch lengths and topologies were obtained using Bayesian inference analysis, and the numbers above each branches are the posterior probabilities and ML bootstrap support.

<https://doi.org/10.1371/journal.pone.0175244.g008>



**Fig 9. The phylogenetic trees of Charadriiformes based on the concatenated 12 mitochondrial protein-coding genes using BI and ML methods.** Branch lengths and topologies were obtained using Bayesian inference analysis. Asterisks indicate bootstrap values of 100% and posterior probability of 1.

<https://doi.org/10.1371/journal.pone.0175244.g009>

is always negative, with the ATP8 containing significantly higher value than other regions. However, the marked negative AT skew is found in 5 gene regions (ND1, COI, ND3, ND4L and CR). The average uncorrected pairwise distances reveal heterogeneity of evolutionary rate for each gene, the COIII, COI and COII have slow evolutionary rate, while it is relative fast for the ND6, ND2 and ATP8. In 13 protein-coding genes, Charadriiformes mitogenomes contain 5 kinds of start codons (ATG, GTG, ATT, ATC, and ATA), four stop codons (TAA, TAG, AGG, and AGA), and an incomplete stop codon (T-). The codon usage analysis revealed that the ATG and TAA are the most common start and stop codon, respectively. The start codon ATG appears in 12 protein-coding genes (with the exception of ND3) and 8 genes only used ATG as start codon. The stop codon TAA appears in 8 genes, and 6 genes (COII, ATP8, ATP6, ND3, ND4L, and Cyt *b*) exclusively used TAA as stop codon. Putative secondary structures of tRNAs indicate that the sequences and structures of anticodon, amino acceptor, and TψC arms are highly conserved, and most of the nucleotide variation is restricted to DHU arms with obvious indel polymorphisms. In the CR, there are 15 kinds of tandem repeats, 4 bp (5' -AAAC-3') and 7 bp (5' -AAACAAC-3') repeat sequences occur frequently. Phylogenomic analysis based on the nearly complete mitochondrial genomes strongly supports the monophyly of the order Charadriiformes, with the suborder Lari sister to the Scolopaci, which is in turn a sister group to the suborder Charadrii. Moreover, our results well resolved the complexity family-

level relationships and clearly depicted the evolutionary processes of Charadriiformes, based on 12 mitochondrial protein-coding genes from 18 families. In future, sequencing more mitogenomes from various taxonomic levels will significantly improve our understanding of mitogenomic evolution and phylogenetic relationships in Charadriiformes.

## Supporting information

**S1 Table. Taxon sampling and GenBank accession numbers.**  
(XLSX)

**S2 Table. Gene organization of *Gallinago stenura* mitochondrial genome.**  
(DOCX)

**S1 Fig. Predicted secondary structures of the 22 tRNA genes identified in the mitogenome of *G. stenura*.** All tRNA genes are shown in the order of occurrence in the mitochondrial genome starting from tRNA<sup>Phe</sup>.  
(TIF)

**S2 Fig. The alignment and predicted secondary structures of the 22 tRNA genes identified in the Charadriiformes mitogenomes.**  
(TIF)

## Acknowledgments

The authors alone are responsible for the content and writing of the article. We thank Wei Liu and Fangzheng Zhao for assistance in the data analysis, and three anonymous referees for valuable comments on the manuscript.

## Author Contributions

**Conceptualization:** CH QC.

**Data curation:** CH YZ WX.

**Formal analysis:** CH CZ.

**Funding acquisition:** CH QC.

**Investigation:** CH QC.

**Methodology:** CH BZ.

**Project administration:** CH QC.

**Resources:** QC.

**Software:** CH CZ BZ.

**Supervision:** CH CZ QC.

**Validation:** CH QC.

**Visualization:** CH QC WX.

**Writing – original draft:** CH QC.

**Writing – review & editing:** CH LS QC.

## References

1. Barbosa A, Moreno E (1999) Hindlimb morphology and locomotor performance in waders: an evolutionary approach. *Biol J Linn Soc* 67: 313–330.
2. Friesen VL (2015) Speciation in seabirds: why are there so many species... and why aren't there more? *J Ornithol* 156: 1–13.
3. Reynolds JD, Székely T (1997) The evolution of parental care in shorebirds: Life histories, ecology, and sexual selection. *Behav Ecol* 8: 126–134.
4. Smith NA, Clarke JA (2015) Systematics and evolution of the Pan-Alcidae (Aves, Charadriiformes). *J Avian Biol* 46: 125–140.
5. Björklund M (1994) Phylogenetic relationships among Charadriiformes: reanalysis of previous Data. *Auk* 111: 825–832.
6. Chu PC (1995) Phylogenetic reanalysis of strauch's osteological data set for the Charadriiformes. *Condor* 97: 174–196.
7. Livezey BC (2010) Phylogenetics of modern shorebirds (Charadriiformes) based on phenotypic evidence: analysis and discussion. *Zool J Linn Soc* 160: 567–618.
8. Paton TA, Baker AJ (2006) Sequences from 14 mitochondrial genes provide a well-supported phylogeny of the Charadriiform birds congruent with the nuclear RAG-1 tree. *Mol Phylogenet Evol* 39: 657–667. <https://doi.org/10.1016/j.ympev.2006.01.011> PMID: 16531074
9. Tuinen MV, Waterhouse D, Dyke GJ (2004) Avian molecular systematics on the rebound: a fresh look at modern shorebird phylogenetic relationships. *J Avian Biol* 35: 191–194.
10. Yoon KB, Cho CU, Park YC (2015) The mitochondrial genome of the Saunders's gull *Chroicocephalus saundersi* (Charadriiformes: Laridae) and a higher phylogeny of shorebirds (Charadriiformes). *Gene* 572: 227–236. <https://doi.org/10.1016/j.gene.2015.07.022> PMID: 26165451
11. Fain MG, Houde P (2007) Multilocus perspectives on the monophyly and phylogeny of the order Charadriiformes (Aves). *BMC Evol Biol* 7: 324.
12. Paton TA, Baker AJ, Groth JG, Barrowclough GF (2003) RAG-1 sequences resolve phylogenetic relationships within Charadriiform birds. *Mol Phylogenet Evol* 29: 268–278. PMID: 13678682
13. Baker AJ, Pereira SL, Paton TA (2007) Phylogenetic relationships and divergence times of Charadriiformes genera: multigene evidence for the Cretaceous origin of at least 14 clades of shorebirds. *Biol Lett* 3: 205–209. <https://doi.org/10.1098/rsbl.2006.0606> PMID: 17284401
14. Boore JL (1999) Animal mitochondrial genomes. *Nucleic Acids Res* 27: 1767–1780. PMID: 10101183
15. Ruokonen M, Kvist L (2002) Structure and evolution of the avian mitochondrial control region. *Mol Phylogenet Evol* 23: 422–432. PMID: 12099796
16. Wolstenholme DR (1992) Animal mitochondrial DNA: structure and evolution. *Int Rev Cytol* 141: 173–216. PMID: 1452431
17. San MD, Gower DJ, Zardoya R, Wilkinson M (2006) A hotspot of gene order rearrangement by tandem duplication and random loss in the vertebrate mitochondrial genome. *Mol Biol Evol* 23: 227–234. <https://doi.org/10.1093/molbev/msj025> PMID: 16177229
18. Morgan-Richards M, Trewick SA, Bartosch-Härlid A, Kardailsky O, Phillips MJ, Mclenachan PA et al. (2008) Bird evolution: testing the Metaves clade with six new mitochondrial genomes. *BMC Evol Biol* 8: 20. <https://doi.org/10.1186/1471-2148-8-20> PMID: 18215323
19. Slack KE, Delsuc F, Mclenachan PA, Arnason U, Penny D (2007) Resolving the root of the avian mitochondrial tree by breaking up long branches. *Mol Phylogenet Evol* 42: 1–13. <https://doi.org/10.1016/j.ympev.2006.06.002> PMID: 16854605
20. 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. *Mol Biol Evol* 28: 1927–1942. <https://doi.org/10.1093/molbev/msr014> PMID: 21242529
21. Jiang L, Chen J, Wang P, Ren Q, Yuan J, Qian C, et al. (2014) The Mitochondrial Genomes of *Aquila fasciata* and *Buteo lagopus* (Aves, Accipitriformes): Sequence, Structure and Phylogenetic Analyses. *PLoS ONE* 10.
22. Li X, Huang Y, Lei F (2014) Comparative mitochondrial genomics and phylogenetic relationships of the *Crossoptilon* species (Phasianidae, Galliformes). *BMC Genomics* 16: 1–12.
23. Pozzi L, Hodgson JA, Burrell AS, Sterner KN, Raaum RL, Disotell TR (2014) Primate phylogenetic relationships and divergence dates inferred from complete mitochondrial genomes. *Mol Phylogenet Evol* 75: 165–183. <https://doi.org/10.1016/j.ympev.2014.02.023> PMID: 24583291
24. Timmermans MJTN, Lees DC, Simonsen TJ (2014) Towards a mitogenomic phylogeny of Lepidoptera. *Mol Phylogenet Evol* 79: 169–178. <https://doi.org/10.1016/j.ympev.2014.05.031> PMID: 24910155

25. Yuan ML, Zhang QL, Guo ZL, Wang J, Shen YY (2015) Comparative mitogenomic analysis of the superfamily Pentatomoidea (Insecta: Hemiptera: Heteroptera) and phylogenetic implications. *BMC Genomics* 16: 1–16.
26. Zhang L, Sechi P, Yuan M, Jiang J, Dong Y, Qiu J (2015) Fifteen new earthworm mitogenomes shed new light on phylogeny within the *Pheretima* complex. *Sci Rep* 6: 28–44.
27. Ballard JWO, Pichaud N (2013) Mitochondrial DNA: more than an evolutionary bystander. *Funct Ecol* 28: 218–231.
28. Zhou X, Lin Q, Fang W, Chen X (2014) The complete mitochondrial genomes of sixteen ardeid birds revealing the evolutionary process of the gene rearrangements. *BMC Genomics* 15: 1–9.
29. Zhou X, Yao C, Lin Q, Fang W, Chen X (2016) Complete mitochondrial genomes render the Night Heron genus *Gorsachius* non-monophyletic. *J Ornithol* 157: 1–9.
30. Ericson PG, Envall I, Irestedt M, Norman JA (2003) Inter-familial relationships of the shorebirds (Aves: Charadriiformes) based on nuclear DNA sequence data. *BMC Evol Biol* 3: 149–154.
31. Sambrook J, Russell DW (1989) *Molecular cloning: a laboratory manual*. v.3: Cold Spring Harbor Laboratory Press.
32. Lowe TM, Eddy SR (1997) tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res* 25: 955–964. PMID: [9023104](https://pubmed.ncbi.nlm.nih.gov/9023104/)
33. Bernt M, Donath A, Jühling F, Externbrink F, Florentz C, Fritzsche G, et al. (2013) MITOS: Improved de novo metazoan mitochondrial genome annotation. *Mol Phylogenet Evol* 69: 313–319. <https://doi.org/10.1016/j.ympev.2012.08.023> PMID: [22982435](https://pubmed.ncbi.nlm.nih.gov/22982435/)
34. Laslett D, Canbäck B (2008) ARWEN: a program to detect tRNA genes in metazoan mitochondrial nucleotide sequences. *Bioinformatics* 24: 172–175. <https://doi.org/10.1093/bioinformatics/btm573> PMID: [18033792](https://pubmed.ncbi.nlm.nih.gov/18033792/)
35. Lobry JR (1996) Asymmetric substitution patterns in the two DNA strands of bacteria. *Mol Biol Evol* 13: 660–665. PMID: [8676740](https://pubmed.ncbi.nlm.nih.gov/8676740/)
36. Perna NT, Kocher TD (1995) Patterns of nucleotide composition at fourfold degenerate sites of animal mitochondrial genomes. *J Mol Evol* 41: 353–358. PMID: [7563121](https://pubmed.ncbi.nlm.nih.gov/7563121/)
37. Benson G (1999) Tandem repeats finder: a program to analyze DNA sequences. *Nucleic Acids Res* 27: 573–580. PMID: [9862982](https://pubmed.ncbi.nlm.nih.gov/9862982/)
38. Tamura K, Stecher G, Peterson D, Filipowski A, Kumar S (2013) MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Mol Biol Evol* 30: 2725–2729. <https://doi.org/10.1093/molbev/mst197> PMID: [24132122](https://pubmed.ncbi.nlm.nih.gov/24132122/)
39. Librado P, Rozas J (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* 25: 1451. <https://doi.org/10.1093/bioinformatics/btp187> PMID: [19346325](https://pubmed.ncbi.nlm.nih.gov/19346325/)
40. Delpont W, Poon AFY, Frost SDW, Pond SLK (2010) Datamonkey 2010: a suite of phylogenetic analysis tools for evolutionary biology. *Bioinformatics* 26: 2455–2457. <https://doi.org/10.1093/bioinformatics/btq429> PMID: [20671151](https://pubmed.ncbi.nlm.nih.gov/20671151/)
41. Yang Z, Nielsen R, Goldman N, Pedersen AM (2000) Codon-substitution models for heterogeneous selection pressure at amino acid sites. *Genetics* 155: 431–449. PMID: [10790415](https://pubmed.ncbi.nlm.nih.gov/10790415/)
42. Groth JG, Barrowclough GF (1999) Basal divergences in birds and the phylogenetic utility of the nuclear RAG-1 gene. *Mol Phylogenet Evol* 12: 115–123. <https://doi.org/10.1006/mpev.1998.0603> PMID: [10381315](https://pubmed.ncbi.nlm.nih.gov/10381315/)
43. Paton T, Haddrath O, Baker AJ (2002) Complete mitochondrial DNA genome sequences show that modern birds are not descended from transitional shorebirds. *P Roy Soc B Biol Sci* 269: 839–846.
44. Stamatakis A (2014) RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30: 1312–1313. <https://doi.org/10.1093/bioinformatics/btu033> PMID: [24451623](https://pubmed.ncbi.nlm.nih.gov/24451623/)
45. Ronquist F, Teslenko M, Mark PVD, Ayres DL, Darling A, Höhna S, et al. (2012) MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Systemat Biol* 61: 539–542.
46. Lanfear R, Calcott B, Ho SY, Guindon S (2012) Partitionfinder: combined selection of partitioning schemes and substitution models for phylogenetic analyses. *Mol Biol Evol* 29: 1695–1701. <https://doi.org/10.1093/molbev/mss020> PMID: [22319168](https://pubmed.ncbi.nlm.nih.gov/22319168/)
47. Hassanin A (2006) Phylogeny of Arthropoda inferred from mitochondrial sequences: Strategies for limiting the misleading effects of multiple changes in pattern and rates of substitution. *Mol Phylogenet Evol* 38: 100–116. <https://doi.org/10.1016/j.ympev.2005.09.012> PMID: [16290034](https://pubmed.ncbi.nlm.nih.gov/16290034/)

48. Hassanin A, Léger N, Deutsch J (2005) Evidence for multiple reversals of asymmetric mutational constraints during the evolution of the mitochondrial genome of metazoa, and consequences for phylogenetic inferences. *Systemat Biol* 54: 277–298.
49. Anderson S, Bankier AT, Barrell BG, De Bruijn MHL, Coulson AR, Drouin J, et al. (1981) Sequence and organization of the human mitochondrial genome. *Nature* 290: 457–465. PMID: [7219534](#)
50. Lavrov DV, Boore JL, Brown WM (2002) Complete mtDNA sequences of two millipedes suggest a new model for mitochondrial gene rearrangements: duplication and nonrandom loss. *Mol Biol Evol* 19: 163–169. PMID: [11801744](#)
51. Ojala D, Montoya J, Attardi G (1981) TRNA punctuation model of RNA processing in human mitochondrial. *Nature* 290: 470–474. PMID: [7219536](#)
52. Ryu SH, Hwang UW (2012) Complete mitochondrial genome of Saunders's gull *Chroicocephalus saundersi* (Charadriiformes, Laridae). *Mitochondr DNA* 23: 134–136.
53. Mindell DP, Sorenson MD, Dimcheff DE (1998) An Extra Nucleotide is not Translated in Mitochondrial ND3 of Some Birds and Turtles. *Mol Biol Evol* 15: 1568–1571. PMID: [12572620](#)
54. Gibson R, Baker A (2012) Multiple gene sequences resolve phylogenetic relationships in the shorebird suborder Scolopaci (Aves: Charadriiformes). *Mol Phylogenet Evol* 64: 66–72. <https://doi.org/10.1016/j.ympev.2012.03.008> PMID: [22491071](#)