Pakistan J. Zool., pp 1-9, 2021. **DOI**: https://dx.doi.org/10.17582/journal.pjz/20200221020245

Complete Mitochondrial Genome of the Eurasian Oystercatcher *Haematopus ostralegus* and Comparative Genomic Analyses in Charadriiformes

Chaochao Hu[1,2], Xue Xu[1], Wenjia Yao[1], Wei Liu[3], Deyun Tai[1], Wan Chen[4] and Qing Chang[1]*

[1]*Jiangsu Key Laboratory for Biodiversity and Biotechnology, College of Life Sciences, Nanjing Normal University, Nanjing 210023, Jiangsu, China.*
[2]*Analytical and Testing Center, Nanjing Normal University, Nanjing 210023, Jiangsu, China.*
[3]*Nanjing Institute of Environmental Sciences, Ministry of Environmental Protection, Nanjing 210042, Jiangsu, China.*
[4]*College of Environment and Ecology, Jiangsu Open University (The City Vocational College of Jiangsu), Nanjing 210036, Jiangsu, China.*

Article Information
Received 21 Februaary 2020
Revised 02 May 2020
Accepted 20 May 2020
Available online 10 December 2020

Authors' Contribution
CCH and QC conceived and designed the study. XX, WJY, DYT, WL and WC collected the data. CCH, XX and WL analyzed the data and wrote the first draft of the manuscript. QC commented on the manuscript.

Key words
Mitogenome, Charadriiformes, *Haematopus ostralegus*

ABSTRACT

The Eurasian oystercatcher *Haematopus ostralegus* is a relatively large migratory wader with an extremely large distribution range. The objective of this study was to determine the complete mitogenome sequence of *H. ostralegus*, and illustrating mitogenomes structure and investigating their evolutionary relationship by comparing 31 species in Charadriiformes. The complete mitochondrial genome of *H. ostralegus* is a circular molecule of 16,798 bp in length and the overall nucleotide composition of H-strand was A: 31.45%, T: 23.46%, C: 31.30%, G: 13.79%. In Charadriiformes mitogenomes, AT skews values were positive, while the values of GC skew were negative. We found a significant negative correlation between CBI and ENC, and a significant positive correlation was found between CBI and G + Cc and G + C3s. The ratio of Ka/Ks of all PCGs indicated that all these genes evolved under purifying selection. Evidence of positive selection was obtained for three (ND2, ND4 and ND6) genes by at least one of the methods. This study provides a valuable resource facilitating further study of phylogenetic and evolutionary analysis for wader, and improves our understanding of the evolutionary and taxonomic research within Charadriiformes.

INTRODUCTION

The resources of mitogenome have rapidly accumulated in recent years thanks to the advanced genomic sequencing, while the mitogenome has not been well studied in Charadriiformes (Baker *et al*., 2007; Gibson and Baker, 2012; Friesen, 2015; Hu *et al*., 2017). The Eurasian oystercatcher, *Haematopus ostralegus* (Charadriiformes, Haematopodidae), is a relatively large wader with an extremely large distribution ranging from Europe to Siberia. This species has been categorized as near threatened in red list of threatened species. Eurasian oystercatcher feeds on mussels, crabs, earth worms and all kinds of invertebrates, rarely on fish. Recent studies of *H. ostralegus* pay more attention on habitat ecology, foraging behavior (Schwemmer *et al*., 2016, 2017; Bailey *et al*., 2019). Based on eight polymorphic microsatellite loci, no significant genetic differentiation was observed between two groups of 'residents' and 'leapfrogs', which were divided by the observational data on the dispersal behaviour of breeding individuals in the island of Schiermonnikoog, Netherlands (Van *et al*., 2010). Few population genetics have been carried out, which might be the lack of basic genetics data (or molecular markers) of *H. ostralegus*.

Animal mitogenome typically contains 13 protein-coding genes, 2 ribosomal RNAs (12S rRNA and 16S rRNA), 22 transfer RNAs (tRNAs), and a non-coding control region (Ruokonen and Kvist, 2002). During the past decade, mitogenome provides a valuable resource for further study of molecular systematics, species identification, population genetics, phylogeny, taxonomy and so on (Li *et al*., 2016; Du *et al*., 2019; Hu *et al*., 2020). In this study, we sequenced the complete

* Corresponding author: Qingchangnj@163.com
0030-9923/2021/0001-0001 $ 9.00/0

mitogenome of *H. ostralegus*, which may provide a valuable resource facilitating further study of population genetics and biogeography of *H. ostralegus*, and provide useful information for understanding the evolutionary and taxonomic research within Charadriiformes.

MATERIALS AND METHODS

Ethics statement

Our experimental procedures complied with the current laws on animal welfare and research in China, and were specifically approved by Nanjing Normal University's Animal Care and Use Committee (Permit #2011-04-008).

Sample and DNA extraction

The specimen of *H. ostralegus* was collected from a derelict and abandoned mist net in Dongtai, Yancheng City, Jiangsu Province, China (32°46'23.12" N, 120°57'41.68" E), in July 2017. After collection, the tissue was initially preserved in 95% ethanol in the field, and transferred to −80°C in our laboratory for long-term storage at Nanjing Normal University (specimen voucher number: NJNU-Host07). Total genomic DNA was extracted using standard phenol-chloroform methods (Sambrook and Russell, 1989). The quality of DNA was assessed through electrophoresis in a 1% agarose gel and staining with Gold View. The complete mitogenome of *H. ostralegus* was generated by amplification of overlapping polymerase chain reaction (PCR) fragments (Hu *et al*., 2017). 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. Sequences obtained were aligned and edited using the software SeqMan (DNAStar, Inc.) to generate complete mitochondrial DNA sequences.

PCR amplification and sequencing

All primers used in this study were taken from Hu *et al*. (2017). PCR was performed in a 30 μL system, which contained 2 × Taq PCR SuperMix (Tiangen, China) 15 μL, 10 μM of each primer (forward and reverse) 1 μL, 1 μL template DNA and 12 μL sterile double-distilled water (ddH2O). PCR profile was 5 min initial denaturation at 95 °C; 35 cycles of 30 s denaturation at 95 °C, 30 s annealing at 50–55 °C, and 1–2 min extension at 72 °C; and a final extension at 72 °C for 8 min and 4 °C hold. The PCR products were electrophoresed by 1% agarose gel and then purified and sequenced with each of the PCR primers on a DNA sequencer (ABI 3731XL).

Genome annotation and bioinformatics analysis

Sequences obtained were assembled and edited manually using the software SeqMan (DNAStar, Inc.) to generate complete mitochondrial DNA sequences. The protein-coding genes (PCGs) 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 (Lowe and Eddy, 1997), MITOS (Bernt *et al*., 2013), and ARWEN (Laslett and Canbäck, 2008) 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) (Perna and Kocher, 1995; Lobry, 1996). The tandem repeats were searched in the CR using the Tandem Repeats Finder program (Benson, 1999).

Codon usage and nucleotide composition statistics were estimated using DnaSP 5.1 (Librado and Rozas, 2009) and Microsoft Excel 2016. The number of variable sites, the parsimony informative sites, the singleton, and the average uncorrected pairwise distances for each PCG were calculated by MEGA 6.0 (Tamura *et al*., 2013). The rates of non-synonymous substitutions (Ka, π modified), synonymous substitutions (Ks, π modified), the effective number of codons (ENC) and the codon bias index (CBI) for each PCG was determined with DnaSP 5.0 (Librado and Rozas, 2009).

Phylogenetic analysis

For phylogenetic analysis, PCGs of mitochondrial genomes of 32 Charadriiformes species were used (Table II). Each mitochondrial gene was aligned individually using Muscle in MEGA X (Kumar *et al*., 2018), and subsequently edited and trimmed. Based on 15 mitochondrial genes (13 PCGs, 12S and 16S) of Charadriiformes species, phylogenetic analysis was performed using Bayesian Inference (BI) and Maximum likelihood analysis (ML), with *Columba livia* (KP168712) and *Gallus gallusa* (KM096864) as outgroups. To determine the optimal partitioning of the data, the best-fit partitioning scheme and the most appropriate nucleotide evolution model for each partition were implemented in Partition Finder 1.1.1 (Lanfear *et al*., 2012). BI method was performed using MrBayes 3.1.2 (Ronquist and Huelsenbeck, 2003). Four Markov Chains Monte Carlo (MCMC) chains were run for 1.0×10^6 generations. Two independent runs were performed to allow additional confirmation of the convergence of MCMC runs. ML analysis was performed with RAxML 8.0.0 (Stamatakis, 2014). The 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.

Table I. Characteristics of the mitochondrial genome of *Haematopus ostralegus*.

Gene	Strand	Position	Intergenic nucleotides	Overlapping nucleotides	Size (bp)	No. of codons	Anti-Codon	Start codon	Stop codon
$tRNA^{Phe}$	H	1–71			71		GAA		
12S rRNA	H	73–1040	1		968				
$tRNA^{Val}$	H	1041–1112			72		TAC		
16S rRNA	H	1113–2699			1587				
$tRNA^{Leu\ (UUR)}$	H	2700–2773			74		TAA		
ND1	H	2776–3753	2		978	325		ATG	AGG
$tRNA^{Ile}$	H	3752–3823		2	72		GAT		
$tRNA^{Gln}$	L	3833–3903	9		71		TTG		
$tRNA^{Met}$	H	3903–3971		1	69		CAT		
ND2	H	3972–5010			1039	345		ATG	T–
$tRNA^{Trp}$	H	5011–5080			70		TCA		
$tRNA^{Ala}$	L	5082–5150	1		69		TGC		
$tRNA^{Asn}$	L	5153–5225	2		73		GTT		
$tRNA^{Cys}$	L	5230–5296	4		67		GCA		
$tRNA^{Tyr}$	L	5296–5366		1	71		GTA		
COI	H	5368–6918	1		1551	516		GTG	AGG
$tRNA^{Ser\ (UCN)}$	L	6910–6983		9	74		TGA		
$tRNA^{Asp}$	H	6986–7054	2		69		GTC		
COII	H	7056–7739	1		684	227		ATG	TAA
$tRNA^{Lys}$	H	7741–7811	1		71		TTT		
ATP8	H	7813–7977	1		165	54		ATG	TAA
ATP6	H	7968–8651		10	684	227		ATG	TAA
COIII	H	8651–9434		1	784	261		ATG	T–
$tRNA^{Gly}$	H	9435–9503			69		TCC		
ND3	H	9504–9855			352	117		ATT	TAA
$tRNA^{Arg}$	H	9858–9926	2		69		TCG		
ND4L	H	9928–10224	1		297	98		ATG	TAA
ND4	H	10218–11595	7		1378	459		ATG	T–
$tRNA^{His}$	H	11596–11665			70		GTG		
$tRNA^{Ser\ (AGY)}$	H	11666–11731			66		GCT		
$tRNA^{Leu\ (CUN)}$	H	11731–11801		1	71		TAG		
ND5	H	11802–13616			1815	604		GTG	TAA
Cyt *b*	H	13630–14772	13		1143	380		ATG	TAA
$tRNA^{Thr}$	H	14778–14847	5		70		TGT		
$tRNA^{Pro}$	L	14862–14931	14		70		TGG		
ND6	L	14953–15474	11		522	173		ATG	T–
$tRNA^{Glu}$	L	15478–15549	3		72		TTC		
CR	H	15550–16798			1249				

Tests of selection

The effect of natural selection on the evolution of the mtDNA PCGs was assessed by comparing the number of nonsynonymous changes per nonsynonymous sites (dN) with that of synonymous changes per synonymous site (dS) (Yang *et al*., 2000). We further estimated the impact of selection along the mtDNA phylogeny of seabirds using codon models to assess the rates of synonymous

and nonsynonymous substitutions. Four methods [Single Likelihood Ancestral Counting (SLAC), Fixed Effects Likelihood (FEL), Random Effects Likelihood (REL), and the mixed effects model of evolution (MEME)] implemented on the DATAMONKEY web server (http://www.datamonkey.org/; last accessed March 12, 2019) were used (Delport *et al*., 2010), by choosing the vertebrate mitochondrial DNA genetic code. The phylogeny estimated in this work, unrooted and excluding the outgroup, was used in all analyses of selection.

RESULTS AND DISCUSSION

Genome organization

The circular mitogenome of *H. ostralegus* is 16,798 bp in length with 13 PCGs, 2 ribosomal RNAs (12S rRNA and 16S rRNA), 22 transfer RNA genes, and a non-coding region (Table I). The annotated mitogenome of *H. ostralegus* has been deposited in GenBank (accession number: MH727533). The graphical mitogenome map was visualized using the software OGDRAW (Lohse *et al*., 2013), and the map was shown in Figure 1. The overall nucleotide composition was A: 31.45%, T: 23.46%, C: 31.30%, G: 13.79%. Among the 37 genes, nine genes ($tRNA^{Gln}$, $tRNA^{Ala}$, $tRNA^{Asn}$, $tRNA^{Cys}$, $tRNA^{Tyr}$, $tRNA^{Ser}$, ND6, $tRNA^{Pro}$ and $tRNA^{Glu}$) were encoded on the light strand, and the remaining 28 genes were encoded on the heavy strand. Gene overlaps have been found at 7 gene junctions, spanning 1–10 nucleotides, for a total of 25 nucleotides. The longest overlap (10 bp) exists between ATP8 and ATP6. The intergenic spacer regions occurred 19 times, spanning 1–14 bp, for a total of 81 bp. The gene order of the *H. ostralegus* mitogenome is identical to that of Charadriiformes mitogenomes, without showing any structural rearrangement.

Nucleotide composition

The 32 mitogenomes within Charadriiformes were summarized and compared (Table II). The differences in length are almost due to the size variation of the non-coding region. The nucleotide composition showed highly similar nucleotide composition biases towards AT rich (mean= 55.74, SD= 0.86) (Table II), which is consistent with previous avian mitogenomes (Hu *et al*., 2017). AT and GC skews are a measure of compositional asymmetry. In Charadriiformes mitogenomes, AT skew values were positive, while the values of GC skew were negative. The AT and GC skew values observed were 0.12 ± 0.02 (mean ± SD) and −0.38 ± 0.01, respectively. In 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 (Hassanin *et al*., 2005; Hassanin, 2006).

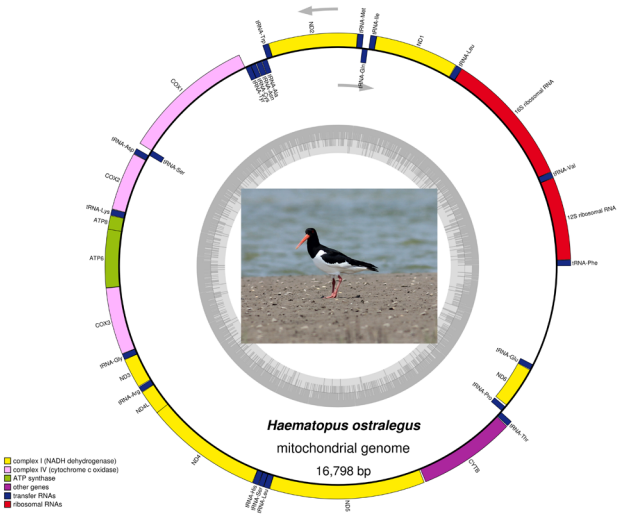

Fig. 1. The gene map of complete mitochondrial genome of the Eurasian Oystercatcher *Haematopus ostralegus*. Genes illustrated on the outside and inside of the main circle are encoded on the heavy (H) strand and the light (L) strand, respectively.

Protein-coding genes (PCGs)

The total length of 13 PCGs is 11,392bp accounting for 67.82% of the complete genome. The nucleotide composition of PCGs can be found (Table II). All PCGs began with the ATG start codon, except ND3 (started with ATT) and ND5 (started with GTG). Three types of termination codons are used in this mitogenome, two PCGs (ND1and COI) use the complete stop codon AGG, seven PCGs (COII, ATP8, ATP6, ND3, ND4L, ND5 and Cyt *b*) stop with TAA, but the remaining four PCGs (ND2, COIII, ND4, and ND6) terminate in the incomplete stop codons T–.

Nucleotide composition bias is also reflected in the codon usage pattern. Among 62 amino acid encoding codons, CUA-Leu (L), AUC- Ile (I), and UUC-Phe (F) were the most frequently used codons. The least frequent codons were CCG-Pro (P), ACG- Thr (T) and CGG- Arg (R). Relative synonymous codon frequencies (RSCU) values were summarized in Table III, revealed that the degenerate codon usage at the third codon positions is generally biased to use more As and Ts than Gs and Cs. The most common amino acids were Leu, Ile and Phe, which are invariably rich in mitochondrial proteins of other birds (Li *et al*., 2014). However, Charadriiformes species used more codon GGA for amino acid Gly than other three degenerate codons (GGG, GGC, and GGU) (Hu *et al*., 2017).

Table II. Comparative nucleotide compositions of 32 species used in Charadriiformes.

Family/Species	Accession no.	Size (bp)	Protein coding genes					
			A (%)	T (%)	G (%)	C (%)	AT skew	GC skew
Family: Haematopodidae								
Haematopus ater	AY074886	16791	31.06	23.57	13.31	32.06	0.14	-0.41
Haematopus ostralegus	MH727533	16789	30.79	23.37	13.57	32.27	0.14	-0.41
Family: Recurvirostridae								
Recurvirostra avosetta	KP757766	16897	31.24	23.40	13.27	32.09	0.14	-0.41
Himantopus himantopus	KY623656	17378	31.38	23.07	13.18	32.37	0.15	-0.42
Family: Charadriidae								
Vanellus cinereus	KM404175	17074	31.06	23.69	13.15	32.09	0.13	-0.42
Vanellus vanellus	KM577158	16795	31.09	23.86	13.33	31.71	0.13	-0.41
Pluvialis fulva	KX639757	16854	30.60	23.48	13.39	32.53	0.13	-0.42
Charadrius placidus	KY419888	16895	30.70	23.79	13.56	31.94	0.13	-0.40
Charadrius alexandrinus	MF565382	16903	30.98	24.41	13.02	31.58	0.12	-0.42
Family: Jacanidae								
Jacana jacana	KJ631049	16975	31.21	24.99	13.30	30.50	0.11	-0.39
Family: Scolopacidae								
Scolopax rusticola	KM434134	16984	30.10	24.84	13.75	31.31	0.10	-0.39
Arenaria interpres	AY074885	16725	30.70	25.04	13.64	30.63	0.10	-0.38
Eurynorhynchus pygmeus	KP742478	16707	31.65	26.43	12.72	29.19	0.09	-0.39
Gallinago stenura	KY056596	16899	30.20	25.55	13.80	30.45	0.08	-0.38
Numenius phaeopus	KP308149	17091	29.73	25.31	14.18	30.78	0.08	-0.37
Xenus cinereus	KX644890	16817	31.13	25.14	13.02	30.71	0.11	-0.40
Tringa erythropus	KX230491	16683	30.85	25.41	13.19	30.56	0.10	-0.40
Tringa semipalmata inornata	MF036175	16603	30.08	25.83	13.77	30.32	0.08	-0.38
Limosa lapponica	KX371106	16732	30.43	24.40	13.54	31.62	0.11	-0.40
Family: Stercorariidae								
Stercorarius maccormicki	KM401546	16669	30.66	24.97	13.21	31.16	0.10	-0.40
Family: Alcidae								
Synthliboramphus antiquus	AP009042	16730	30.47	24.85	13.36	31.32	0.10	-0.40
Synthliboramphus wumizusume	KT592378	16714	30.61	25.12	13.44	30.84	0.10	-0.39
Family: Laridae								
Chroicocephalus brunnicephalus	JX155863	16769	30.07	24.07	13.93	31.92	0.11	-0.39
Chroicocephalus ridibundus	KM577662	16807	30.04	24.10	13.96	31.90	0.11	-0.39
Chroicocephalus saundersi	JQ071443	16725	29.68	24.07	14.26	32.00	0.10	-0.38
Larus crassirostris	KM507782	16746	29.97	24.44	13.90	31.69	0.10	-0.39
Larus dominicanus	AY293619	16701	30.03	24.43	13.84	31.70	0.10	-0.39
Larus vegae	KT943749	16379	30.81	25.90	13.22	30.07	0.09	-0.39
Ichthyaetus relictus	KC760146	16586	30.19	24.31	13.70	31.81	0.11	-0.40
Gelochelidon nilotica	MF582631	16748	30.29	25.54	13.73	30.44	0.08	-0.38
Sterna hirundo	MF582632	16707	29.96	25.69	13.99	30.36	0.08	-0.37
Sternula albifrons	KT350612	16357	30.60	26.42	13.41	29.57	0.07	-0.38

Table III. Codon usage in charadriiformes mitochondrial protein-coding genes. A total of 3,763 codons for analyzed, excluding the start and stop codons. AA, amino acid; RSCU, relative synonymous codon usage; n= frequency of each codon; %= n/3737.

AA	Codon	Count	Percentage (%)	RSCU
Phe (F)	UUU	62.5	1.66	0.56
Phe (F)	UUC	161.1	4.28	1.44
Leu2 (L2)	UUA	93.2	2.48	0.85
Leu2 (L2)	UUG	20.8	0.55	0.19
Leu1 (L1)	CUU	69.4	1.84	0.63
Leu1 (L1)	CUC	141.5	3.76	1.29
Leu1 (L1)	CUA	301.9	8.02	2.75
Leu1 (L1)	CUG	31.1	0.83	0.28
Ile (I)	AUU	89.5	2.38	0.62
Ile (I)	AUC	200.5	5.33	1.38
Met (M)	AUA	143.8	3.82	1.75
Met (M)	AUG	21	0.56	0.25
Val (V)	GUU	34.6	0.92	0.83
Val (V)	GUC	45.5	1.21	1.09
Val (V)	GUA	69.6	1.85	1.66
Val (V)	GUG	17.7	0.47	0.42
Ser2 (S2)	UCU	39.1	1.04	0.83
Ser2 (S2)	UCC	83	2.21	1.76
Ser2 (S2)	UCA	99.2	2.64	2.10
Ser2 (S2)	UCG	7.6	0.20	0.16
Pro (P)	CCU	34.3	0.91	0.63
Pro (P)	CCC	71.7	1.91	1.31
Pro (P)	CCA	106.3	2.82	1.94
Pro (P)	CCG	6.4	0.17	0.12
Thr (T)	ACU	61.9	1.64	0.71
Thr (T)	ACC	147.7	3.93	1.69
Thr (T)	ACA	134.6	3.58	1.54
Thr (T)	ACG	6.2	0.16	0.07
Ala (A)	GCU	60.4	1.61	0.84
Ala (A)	GCC	122.6	3.26	1.70
Ala (A)	GCA	98.9	2.63	1.37
Ala (A)	GCG	7.1	0.19	0.10
Tyr (Y)	UAU	34.2	0.91	0.61
Tyr (Y)	UAC	77.5	2.06	1.39
His (H)	CAU	30.4	0.81	0.56
His (H)	CAC	77.9	2.07	1.44
Gln (Q)	CAA	87.5	2.33	1.83
Gln (Q)	CAG	8.3	0.22	0.17
Asn (N)	AAU	25.3	0.67	0.39
Asn (N)	AAC	106.1	2.82	1.61
Lys (K)	AAA	79.2	2.10	1.84
Lys (K)	AAG	7	0.19	0.16
Asp (D)	GAU	18.4	0.49	0.59
Asp (D)	GAC	44	1.17	1.41
Glu (E)	GAA	81.2	2.16	1.74
Glu (E)	GAG	12	0.32	0.26
Cys (C)	UGU	10.2	0.27	0.70
Cys (C)	UGC	18.8	0.50	1.30
Trp (W)	UGA	96.3	2.56	1.80
Trp (W)	UGG	10.8	0.29	0.20
Arg (R)	CGU	10.1	0.27	0.57
Arg (R)	CGC	18.5	0.49	1.04
Arg (R)	CGA	37.8	1.00	2.13
Arg (R)	CGG	4.6	0.12	0.26
Ser (S1)	AGU	8.6	0.23	0.18
Ser (S1)	AGC	46.1	1.23	0.98
Gly (G)	GGU	36.7	0.98	0.67
Gly (G)	GGC	64.3	1.71	1.16
Gly (G)	GGA	88.9	2.36	1.61
Gly (G)	GGG	30.9	0.82	0.56

Continued on next column...

In order to study the codon usage bias among Charadriiformes, we also analyzed the correlations between ENC (effective number of codons), CBI (codon bias index), the G + C content of all codons (G + Cc), and the G + C content of the third codon position (G + C3s). We found a significant negative correlation between CBI and ENC (R = 0.94, P < 0.05), and a significant positive correlation was found between CBI and G + Cc (R = 0.50, P < 0.05) and G + C3s (R = 0.58, P < 0.05) (Fig. 2). However, other pairs were not correlated with each other pairs. These results were consistent with the neutral mutational theories, in which the G + C content of mitochondrial genome was reported to be the most significant factor in determining codon bias among organisms (Plotkin and Kudla, 2011).

Phylogenetic analysis

The phylogenetic analysis resolved a well-supported clade of Charadriiformes (Fig. 3), which showed great mitochondrial divergence within the Charadriiformes. Relationships of the phylogeny strongly support monophyly

of the order Charadriiformes, which has a long evolutionary history dating back at least to the late Cretaceous (Baker *et al*., 2007). This phylogeny was in agreement with previous family hypotheses for shorebirds (Paton and Baker, 2006; Hu *et al*., 2017), with the exceptions of relationships of *Larus vegae* (GenBank No. KT943749). This sample (*Larus vegae* KT943749) was sequenced on Next Generation Sequencing platform of IonTorrent, and the de novo assemblies were conducted with CLC genomics workbench v. 8.5.1 with the coverage of 1070.46. This study provides a valuable resource facilitating further study of population genetics of *H. ostralegus* and improves our understanding of the evolutionary and taxonomic research within Charadriiformes.

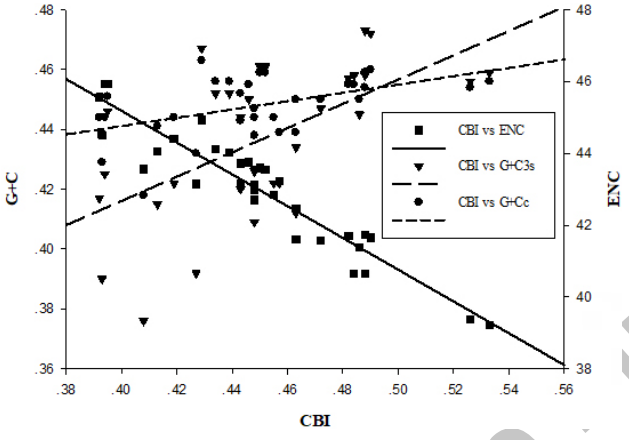

Fig. 2. Evaluation of codon bias in the mitochondrial genomes of 32 species. ENC, effective number of codons; CBI, codon bias index; G + Cc, G + C content of all codon positions; G + C3s, G + C content of the third codon positions.

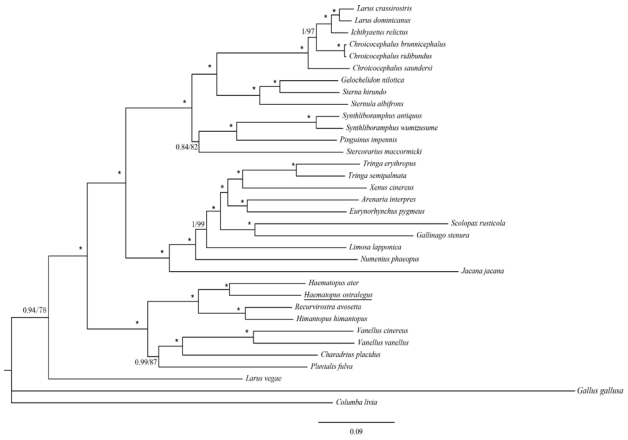

Fig. 3. The molecular phylogenetic relationships of *Haematopus ostralegus* within the order Charadriiformes based on sequence of 13 PCGs and 2 ribosomal RNAs (12S rRNA and 16S rRNA) by using partitioned Bayesian analysis. The GenBank accession numbers for each species are shown in parentheses. Values on the branches represent bootstrap values/Bayesian posterior probabilities. "*" represented the nodes received the values of 1.00/100 unless otherwise labeled. The species which was sequenced in this study was underlined.

Test of selection

We compared Ka/Ks ratios for all 13 PCGs in 32 Charadriiformes species. The average value of Ka and Ks ranged from 0.01 of COI to 0.09 of ATP8 and from 0.53 of ATP8 to 0.88 of ND1, respectively (Table IV). The ratio of Ka/Ks of all PCGs was far lower than one (≤ 0.17), indicating that all these genes evolved under purifying selection.

Table IV. The mutational information and the test of selection results among Charadriiformes were calculated by each protein coding genes.

Gene	Length (bp)	%Vs	%Pis	%S	ts/tv	Ks	Ka	Ka/Ks	SLAC	FEL	REL	MEME
ND1	969	47.57%	40.76%	6.81%	4.67	0.88	0.04	0.05	0	0	0	4 (4, 89,194,253)
ND2	1023	54.15%	43.99%	10.17%	3.63	0.78	0.07	0.09	1(277)	1(277)	2(89,276)	
COI	1545	36.76%	32.88%	3.88%	4.08	0.75	0.01	0.02	0	0	0	1(407)
COII	678	40.71%	35.10%	5.60%	4.26	0.75	0.02	0.03	1(128)	0	0	0
ATP8	162	60.49%	50.62%	10.49%	3.10	0.53	0.09	0.17	0	0	1(50)	1(17)
ATP6	678	49.71%	40.86%	8.85%	3.07	0.76	0.04	0.06	0	0	1(193)	1(115)
COIII	780	39.23%	32.44%	6.79%	3.74	0.67	0.02	0.03	0	0	0	1(153)
ND3	345	47.25%	41.74%	5.51%	4.12	0.74	0.06	0.08	0	0	0	0
ND4L	291	50.17%	42.96%	7.22%	3.43	0.73	0.05	0.07	0	0	0	1(72)
ND4	1374	50.51%	41.85%	8.66%	3.09	0.69	0.06	0.08	3(13,88,96)	1(96)	0	3(96,187,448)
ND5	1800	52.22%	41.56%	10.67%	3.26	0.66	0.07	0.10	2(439,520)	0	0	8(22,276,345,351,438,475,524,600)
Cyt *b*	1128	45.83%	36.17%	9.66%	2.89	0.70	0.03	0.04	1(212)	0	0	1(212)
ND6	516	53.88%	42.64%	11.24%	4.41	0.73	0.07	0.09	1(115)	1(115)	0	4(64,97,115,134)

We then applied several methods to detect evidence of positive selection affecting mtDNA PCGs throughout the phylogenetic tree of seabird. Several positive selection sites were found based on site-specific analyses (Table IV). There are 12 different codons in eight genes were suggested to have evolved under positive selection by REL, FEL, or SLAC. Only three codons (in ND2, ND4 and ND6) were concordant in two of the methods (Table IV). MEME, which is particularly sensitive to events of episodic selection, suggested that 27 sites of 11 genes evolved under positive selection, which may have affected several lineages along the evolution of seabirds.

ACKNOWLEDGEMENTS

This research was supported by grants from the Program of Nature Science Fund of Jiangsu Province (# BK20181076) and Qing Lan Project of Jiangsu High School (# 2019SZJS-003).

Statement of conflict of interest

The authors declare no conflict of interest.

REFERENCES

Bailey, L.D., Ens, B.J., Both, C., Heg, D., Oosterbeek, K. and van de Pol, M., 2019. Habitat selection can reduce effects of extreme climatic events in a long-lived shorebird. *J. Anim. Ecol.*, **88**: 1474–1485. https://doi.org/10.1111/1365-2656.13041

Baker, A.J., Pereira, S.L. and Paton, T.A., 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

Benson, G., 1999. Tandem repeats finder: A program to analyze DNA sequences. *Nucl. Acids Res.*, **27**: 573–580. https://doi.org/10.1093/nar/27.2.573

Bernt, M., Donath, A., Jühling, F., Externbrink, F., Florentz, C., Fritzsch, G., Pütz, J., Middendorf, M. and Stadler, P.F., 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

Delport, W., Poon, A.F.Y., Frost, S.D.W. and Pond, S.L.K., 2010. Datamonkey 2010: A suite of phylogenetic analysis tools for evolutionary biology. *Bioinformatics*, **26**: 2455–2457. https://doi.org/10.1093/bioinformatics/btq429

Du, Z., Hasegawa, H., Cooley, J.R., Simon, C., Yoshimura, J., Cai, W., Sota, T. and Li, H., 2019. Mitochondrial genomics reveals shared phylogeographic patterns and demographic history among three periodical cicada species groups. *Mol. Biol. Evol.*, **36**: 1187–1200. https://doi.org/10.1093/molbev/msz051

Friesen, V.L., 2015. Speciation in seabirds: Why are there so many species…and why aren't there more? *J. Ornithol.*, **156**: 1–13. https://doi.org/10.1007/s10336-015-1235-0

Gibson, R. and 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

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

Hassanin, A., Léger, N. and 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. *Syst. Biol.*, **54**: 277–298. https://doi.org/10.1080/10635150590947843

Hu, C.C., Zhang, C.L., Sun, L., Zhang, Y., Xie, W.L., Zhang, B.W. and Chang, Q., 2017. The mitochondrial genome of pin-tailed snipe gallinago stenura, and its implications for the phylogeny of charadriiformes. *PLoS One*, **12**: e0175244. https://doi.org/10.1371/journal.pone.0175244

Hu, Y., Thapa, A., Fan, H., Ma, T., Wu, Q., Ma, S., Zhang, D., Wang, B., Li, M. and Yan, L., 2020. Genomic evidence for two phylogenetic species and long-term population bottlenecks in red pandas. *Sci. Adv.*, **6**: eaax5751. https://doi.org/10.1126/sciadv.aax5751

Kumar, S., Stecher, G., Li, M., Knyaz, C. and Tamura, K., 2018. Mega x: Molecular evolutionary genetics analysis across computing platforms. *Mol. Biol. Evol.*, **35**: 1547–1549. https://doi.org/10.1093/molbev/msy096

Lanfear, R., Calcott, B., Ho, S.Y. and 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

Laslett, D. and 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

Li, Q., Wei, S.J., Tang, P., Wu, Q., Shi, M., Sharkey, M.J. and Chen, X.X., 2016. Multiple lines of evidence

from mitochondrial genomes resolve phylogenetic relationships of parasitic wasps in braconidae. *Genome Biol. Evol.*, **8**: 2651–2662. https://doi.org/10.1093/gbe/evw184

Li, X., Huang, Y. and Lei, F., 2014. Comparative mitochondrial genomics and phylogenetic relationships of the crossoptilon species (Phasianidae, Galliformes). *BMC Genomics*, **16**: 1–12. https://doi.org/10.1186/s12864-015-1234-9

Librado, P. and Rozas, J., 2009. DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. *Bioinformatics*, **25**: 1451–1452. https://doi.org/10.1093/bioinformatics/btp187

Lobry, J.R., 1996. Asymmetric substitution patterns in the two DNA strands of bacteria. *Mol. Biol. Evol.*, **13**: 660–665. https://doi.org/10.1093/oxfordjournals.molbev.a025626

Lohse, M., Drechsel, O., Kahlau, S. and Bock, R., 2013. OrganellarGenomeDRAW-a suite of tools for generating physical maps of plastid and mitochondrial genomes and visualizing expression data sets. *Nucleic Acids Res.*, **41**: 575–581. https://doi.org/10.1093/nar/gkt289

Lowe, T.M. and Eddy, S.R., 1997. tRNAscan-SE: A program for improved detection of transfer rna genes in genomic sequence. *Nucleic Acids Res.*, **25**: 955–964. https://doi.org/10.1093/nar/25.5.955

Paton, T.A. and Baker, A.J., 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

Perna, N.T. and Kocher, T.D., 1995. Patterns of nucleotide composition at fourfold degenerate sites of animal mitochondrial genomes. *J. mol. Evol.*, **41**: 353–358. https://doi.org/10.1007/BF01215182

Plotkin, J.B. and Kudla, G., 2011. Synonymous but not the same: The causes and consequences of codon bias. *Nat. Rev. Genet.*, **12**: 32–42. https://doi.org/10.1038/nrg2899

Ronquist, F. and Huelsenbeck, J.P., 2003. Mrbayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics*, **19**: 1572–1574. https://doi.org/10.1093/bioinformatics/btg180

Ruokonen, M. and Kvist, L., 2002. Structure and evolution of the avian mitochondrial control region. *Mol. Phylogenet. Evol.,* **23**: 422–432. https://doi.org/10.1016/S1055-7903(02)00021-0

Sambrook, J. and Russell, D.W., 1989. *Molecular cloning: A laboratory manual*. V.3. Cold Spring Harbor Laboratory Press.

Schwemmer, P., Güpner, F., Adler, S., Klingbeil, K. and Garthe, S., 2016. Modelling small-scale foraging habitat use in breeding eurasian oystercatchers (*Haematopus ostralegus*) in relation to preydistribution and environmental predictors. *Ecol. Model.*, **320**: 322–333. https://doi.org/10.1016/j.ecolmodel.2015.10.023

Schwemmer, P., Weiel, S. and Garthe, S., 2017. A fundamental study revisited: Quantitative evidence for territory quality in oystercatchers (*Haematopus ostralegus*) using gps data loggers. *Ecol. Evol.*, **7**: 285–294. https://doi.org/10.1002/ece3.2581

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

Tamura, K., Stecher, G., Peterson, D., Filipski, A. and Kumar, S., 2013. Mega6: Molecular evolutionary genetics analysis version 6.0. *Mol. Biol. Evol.*, **30**: 2725–2729. https://doi.org/10.1093/molbev/mst197

Van, T.R., Bijlsma, R., Tinbergen, J.M., Heg, D. and Van, d.Z.L., 2010. Genetic analysis of the population structure of socially organized oystercatchers (*Haematopus ostralegus*) using microsatellites. *Mol. Ecol.*, **8**: 181–187. https://doi.org/10.1046/j.1365-294X.1999.00548.x

Yang, Z., Nielsen, R., Goldman, N. and Pedersen, A.M., 2000. Codon-substitution models for heterogeneous selection pressure at amino acid sites. *Genetics*, **155**: 431–449.