Genetic diversity of two populations of the tufted puffin *Fratercula cirrhata* (Pallas, 1769)

Shizuka Sakurayama¹, Daiki Nojima², Madoka Yoshizawa³, Tomohiro Takeuchi⁴,
Makiko Ito⁴ and Takashi Kitano^{1*}

¹Graduate School of Science and Engineering, Ibaraki University, Hitachi, Ibaraki 316-8511, Japan

²Tokyo Sea Life Park, Edogawa-ku, Tokyo 134-8587, Japan

³Ueno Zoological Gardens, Taito-ku, Tokyo 110-8711, Japan

⁴Aqua World Oarai, Oarai, Ibaraki 311-1301, Japan

(Received 23 September 2020, accepted 1 April 2021; J-STAGE Advance published date: 17 June 2021)

The tufted puffin *Fratercula cirrhata* (Charadriiformes: Alcidae) is distributed throughout the boreal and low Arctic areas of the North Pacific, from California, USA to Hokkaido, Japan. Few studies have investigated the genetic diversity of this species. Therefore, we analyzed the genetic diversity of two captive populations using nucleotide sequences of two mitochondrial loci (*COX1* and *D-loop*) and one nuclear locus (*RHBG*). We sequenced these loci for birds from Tokyo Sea Life Park (Kasai Rinkai Suizokuen), originally from Alaska, and birds from Aqua World Oarai, originally from far eastern Russia. We found five *COX1* haplotypes and 17 *D-loop* haplotypes for the mitochondrial loci, and obtained 14 predicted haplotypes for the nuclear *RHBG* locus. The major haplotypes of all three loci occurred in individuals from both populations. Thus, there were no clear genetic differences between the populations with respect to these three loci. Although the breeding range of the tufted puffin covers the boreal and low Arctic from California to Hokkaido, our results suggest that the species has not genetically diverged within its breeding range.

Key words: haplotype, nucleotide sequence, polymorphism, seabird

INTRODUCTION

The tufted puffin *Fratercula cirrhata* (Pallas, 1769) is a pelagic seabird that belongs to the family Alcidae, in the order Charadriiformes. This species was first described by Peter Simon Pallas in 1769 with the scientific name *Alca cirrhata*, although it was later renamed *Lunda cirrhata* (Pallas, 1811). The first to fifth editions of the Check-list of North American Birds used the genus name *Lunda*. Then, in the 34th supplement (Eisenmann et al., 1982) of the fifth edition, the genus name *Fratercula* was adopted, and this use has continued to the current seventh edition (American Ornithologists' Union, 1998).

The tufted puffin has a thick bill with vertical grooves. During the winter, it is mostly black, with a gray face and amber feet. In the summer, its feet become bright red, its face becomes white with yellow tufts, and its bill becomes a bright orange. A common name of the species in Japan is “etupirka”, from “etu”, meaning nose (likely representing the bill of the bird), and “pirka”, meaning beautiful, in the Ainu language spoken by the Ainu people of northern Japan.

The tufted puffin is found throughout the boreal and low Arctic areas of the North Pacific, from Big Sur, California to Hokkaido, Japan, but it is abundant only from British Columbia to the Sea of Okhotsk. Small numbers penetrate the Chukchi Sea and as far north as Wrangel Island (Gaston and Jones, 1998). Although the tufted puffin has no recognized subspecies (Gaston and Jones, 1998), subtle differences in adult body mass across the species' range have been observed. The body mass of both males and females from the Bering Sea is larger than that of birds from the Aleutian Islands or the Alaska Peninsula, and individuals from the Sea of Okhotsk are smaller than those from elsewhere (Piatt and Kitaysky, 2002).

Edited by Yoko Satta

* Corresponding author: takashi.kitano.evolution@vc.ibaraki.ac.jp
DOI: <https://doi.org/10.1266/ggs.20-00051>

Copyright: ©2021 The Author(s). This is an open access article distributed under the terms of the Creative Commons BY 4.0 International (Attribution) License (<https://creativecommons.org/licenses/by/4.0/legalcode>), which permits the unrestricted distribution, reproduction and use of the article provided the original source and authors are credited.

The genetic diversity of this species has received scant attention in the literature. Only 13 sequences for cytochrome c oxidase subunit I (*COX1*) are available in the Barcode of Life Data System v4 (BOLD) (Ratnasingham and Hebert, 2007) for genetic diversity analysis. Five of these are for birds from Japan (Saitoh et al., 2015), with four from Alaska (Hebert et al., 2004; Tavares and Baker, 2008), one from Canada (Kerr et al., 2007), two from Russia (Kerr et al., 2009) and one of unknown origin (Pereira and Baker, 2008). There are no clear regional differences in the *COX1* data in the BOLD system.

In this study, to improve our understanding of the genetic diversity of tufted puffins, we analyze the genetic diversity of two captive populations (originally from far eastern Russia and Alaska) using not only *COX1* sequences but also the *D-loop* locus, which is expected to exhibit more variability than *COX1* within the mitochondrial DNA (mtDNA). The *D-loop* locus has been effectively used for population studies in several Charadriiformes bird species (e.g., Liebers et al., 2001, 2004; Liebers and Helbig, 2002; Buehler and Baker, 2003, 2005; Miller et al., 2010, 2015; Sternkopf et al., 2010; Rheindt et al., 2011; Birt et al., 2012; Küpper et al., 2012; Verkuil et al., 2012; Pons et al., 2014; Trimbos et al., 2014; Pshenichnikova et al., 2015; Sonsthagen et al., 2015; Tigano et al., 2015; Wallace et al., 2015; Wang et al., 2019). In addition, we also selected one nuclear locus, Rh family B glycoprotein (*RHBG*), to sequence. *RHBG* is a member of the *RH* gene family, which was duplicated in the common ancestor of vertebrates (e.g., Huang and Peng, 2005; Peng and Huang, 2006; Kitano et al., 2010; Suzuki et al., 2014, 2017), during two rounds of whole-genome duplication (Dehal and Boore, 2005). A unique characteristic of the bird lineage is that it exhibits a higher rate of nonsynonymous substitutions in the *RHBG* genes, probably because the functional constraints on this gene are relaxed (Suzuki et al., 2017). We therefore expected to obtain more variable sequence data in the analysis of this nuclear gene.

In Japan, as of 2020, five aquariums (Aquamarine Fukushima, Kamogawa Seaworld, Osaka Aquarium Kaiyukan, Aqua World Oarai and Tokyo Sea Life Park) hold tufted puffins. The birds at four of these aquariums originally came from far eastern Russia, while those at Tokyo Sea Life Park originally came from Alaska. We believe that the genetic data provided in this study will also be helpful for the management of captive tufted puffins in aquariums.

MATERIALS AND METHODS

DNA extraction Tufted puffins from the Tokyo Sea Life Park (Kasai Rinkai Suizokuen, hereafter “Kasai”) and Aqua World Oarai (hereafter “Oarai”) were used. The Kasai individuals were originally from Kodiak Island

and the Triplet Islands in Alaska, USA, while the Oarai individuals were originally from Kamchatka, in far eastern Russia. Individuals were identified by means of implanted microchips. The family trees of the birds were constructed based on the Japanese studbook for tufted puffins (Figs. 1 and 2). Blood collection was carried out under manual restraint, with sufficient attention given to bird welfare. A blood sample was taken from the medial metatarsal vein of each individual and stored at -20°C until DNA extraction. All procedures for collecting samples from birds were performed in accordance with the guidelines of the Japanese Association of Zoos and Aquariums.

Genomic DNA was extracted from the blood samples using the conventional sodium dodecyl sulfate lysis method with proteinase K and phenol–chloroform. The quality of the extracted DNA was confirmed via 0.8% agarose gel electrophoresis, and DNA concentration was measured by a Beckman DU-7500 spectrophotometer (Beckman Coulter, Brea, CA, USA).

Polymerase chain reaction and DNA sequencing PCR was performed in a 5- μl reaction volume containing 20–50 ng of template DNA, $1 \times$ KAPA2G Fast Multiplex PCR Mix (KAPA Biosystems, Wilmington, MA, USA) and 0.2 μM each of the forward and reverse primers. Thermal cycling was performed in a Veriti 200 machine (Applied Biosystems, Waltham, MA, USA). A two-step PCR method was employed: 95°C for 3 min, followed by 30 cycles of denaturation at 95°C for 15 s, annealing at 68°C for 30 s, and extension at 72°C for 1 min.

The primers for *COX1* were fcCOI-1 (5'-TTGGC-GCATGAGCTGGCA-3') and fcCOI-2 (5'-AAAGAAT-CAGAAGAGGTGCTG-3'). These were designed based on *COX1* sequences from tufted puffins retrieved from the DDBJ/EMBL/GenBank International Nucleotide Sequence Database (accession numbers AB842787.1, AB842789.1, AB842790.1, AB843514.1, AB843515.1, AY666344.1, DQ433644.1, EF380329.1, EU525406.1, GQ481922.1 and GQ481923.1). The primers for *D-loop* were fcDloop-1 (5'-GCCTGAAAAGCCGCTGTTG-3') and fcDloop-2 (5'-TAGCCGACCAGATGTATCG-3'). These primers were designed to anneal to the tRNA-Glu (5' end of *D-loop*) and tRNA-Phe (3' end of *D-loop*) regions, respectively. These regions appeared to be relatively well conserved, based on a multiple alignment of the region [tRNA-Glu, *D-loop*, tRNA-Phe] for other Charadriiformes species retrieved from the sequence database (*Arenaria interpres*: AY074885.2; *Chroicocephalus ridibundus*: KM577662.1; *Eurynorhynchus pygmeus*: KP742478.1; *Gallinago stenura*: KY056596.1; *Gelochelidon nilotica*: MF582631.1, *Haematopus ater*: AY074886.2; *H. ostralegus*: KY419886.1; *Himantopus himantopus*: KY623656.1; *Ichthyophaga relictus*: KC760146.1; *Jacana*

Fig. 1. Family trees of tufted puffins from the Kasai population. Circles and squares represent females and males, respectively. Numbers in the circles and squares are individuals' identification numbers. The three known maternal groups are indicated by the colors red, orange and green; white indicates individuals without a recognized maternal group in this population. A black circle or square indicates a deceased individual. †, individuals for which all three loci (*COX1*, *D-loop* and *RHBG*) were sequenced; ‡, individuals for which two mtDNA loci (*COX1* and *D-loop*) were sequenced; §, an individual for which only the *D-loop* locus was sequenced.

Fig. 2. Family trees of tufted puffins from the Oarai population. Symbols are as for Fig. 1. The seven known maternal groups are indicated by the colors brown, red, yellow, green, pink, blue and orange; white indicates individuals without a recognized maternal group in this population.

jacana: KJ631049.1; *J. spinosa*: KJ631048.1; *Larus brun-*
nicephalus: JX155863.1; *L. crassirostris*: KM507782.1; *L.*
dominicanus: AY293619.1; *L. vegae*: KT943749.1; *Numen-*
ius phaeopus: KP308149.1; *Pluvialis fulva*: KX639757.1;
Recurvirostra avosetta: KP757766.1; *Saundersilarus*

saundersi: JQ071443.1; *Scolopax rusticola*: KM434134.1;
Stercorarius maccormicki: KM401546.1; *Sterna hirundo*:
 MF582632.1; *Sternula albifrons*: KT350612.1; *Synth-*
liboramphus antiquus: AP009042.1; *S. wumizusume*:
 KT592378.1; *Tringa erythropus*: KX230491.1; *T. ochro-*

pus: KX668223.1; *T. semipalmata*: MF036175.1; *Vanellus cinereus*: KM404175.1; *V. vanellus*: KM577158.1; *Xenus cinereus*: KX644890.1); the *D-loop* sequence for the tufted puffin was not in the DNA database.

The primers for *RHBG* were fcRHBG-3 (5'-TGCTCAACACCTACTTCTCGC-3') for putative exon 5 and fcRHBG-4 (5'-GGCTACCTTGGGCCACCAG-3') for putative exon 8. The putative exon numbers of the *RHBG* locus were assigned with reference to the chicken (*Gallus gallus*) genome sequence data. These primers were designed based on the *RHBG* sequence of the killdeer (*Charadrius vociferus*; accession number: XM_009892067.1), which belongs to the same order (Charadriiformes) as the tufted puffin and is the most closely related species for which whole-genome shotgun sequence data (Zhang et al., 2014) were available.

Based on the family trees (Figs. 1 and 2), three maternal groups were recognized in Kasai and seven in Oarai. One individual was selected for the sequencing of all three loci from each maternal group in Kasai (K02, K11 and K19) and Oarai (O08, O34, O36, O39, O43, O45 and O48). Since there was no information about the parents of three individuals (K01, K12 and K30) from Kasai (Fig. 1) and eight individuals (O01, O03, O10, O22, O28, O31, O38 and O42) from Oarai (Fig. 2), their maternal groups could not be determined. These individuals were therefore also selected for sequencing. For O48, only the mtDNA loci (*COX1* and *D-loop*) were sequenced, because one of its *RHBG* alleles should be identical to that of its father, O10. In total, six Kasai and 14 Oarai individuals were selected for the sequencing of all three loci. In addition, five individuals (K26, K44, K47, K51 and K57) from Kasai and one (O78) from Oarai were selected for the sequencing of the mtDNA loci only, because their maternal groups were unknown.

The PCR products were confirmed via 1% agarose gel electrophoresis and purified using the High Pure PCR Product Purification Kit (Roche Diagnostics, Mannheim, Germany). DNA sequencing was performed on the PCR products using the BigDye Terminator v1.1/v3.1 Cycle Sequencing Kit and the ABI PRISM 310 Genetic Analyzer (Applied Biosystems). For *COX1*, the primers fcCOI-1 and fcCOI-2 were used for each sequence reaction. For *D-loop*, the primers fcDloop-2 and fcD-loop-7 (5'-CCGCCCAATGGAATCCGA-3'), instead of fcD-loop-1, were used for each sequence reaction, because all sequence reactions using fcDloop-1 failed due to a poly(C) tract. For *RHBG*, in addition to primers fcRHBG-3 and fcRHBG-4, fcRHBG-7 (5'-TCTCAGCCCCGGCTCTTC-3') for putative intron 6 and fcRHBG-8 (5'-GTTTGCAG-GTTTGGGAGC-3') for putative intron 6 were used for sequencing.

Phred/Phrap software (Ewing et al., 1998) was used for base-calling and sequence assembly, and to obtain quality scores for the assembled data. Editing of the assembled

sequences was performed using Consed (Gordon et al., 1998). Essentially, both strands of the PCR products were sequenced. If only one strand of a portion of the end region was sequenced, sites with high-quality scores (Phred quality ≥ 30) were used. Two overlapping peaks for one site in a sequence were interpreted as indicating heterozygosity, and the site was designated according to the Nomenclature Committee of the International Union of Biochemistry protocol (R: A or G; Y: C or T; K: G or T; M: A or C; S: C or G; and W: A or T).

Sequence data analysis Multiple alignments were performed using MUSCLE (Edgar, 2004), implemented with MEGA7 software (Kumar et al., 2016). Haplotypes of *RHBG* were predicted using PHASE 2.1.1 (Stephens

Table 1. Haplotypes of the three tufted puffin loci sequenced

Population	Individual	♂ ♀	Father	Mother	<i>COX1</i>	<i>D-loop</i>	<i>RHBG</i>
Kasai	K01 [†]	♂	nd	nd	I	1	c, f
	K02 [†]	♀	nd	nd	I	2	a, a
	K11 [†]	♀	nd	nd	III	3,4	h, k
	K12 [†]	♂	nd	nd	II	5	l, l
	K19 [†]	♀	nd	nd	I	4	m, m
	K26 [‡]	♂	K01	deceased	II	6	–
	K30 [†]	♂	nd	nd	I	7	a, n
	K44 [‡]	♂	K30	deceased	I	8,9	–
	K47 [‡]	♀	K30	deceased	I	7	–
	K51 [‡]	♂	K38	deceased	II	10	–
	K57 [‡]	♂	K44	deceased	II	11	–
	K62 [§]	♂	K44	K40	–	3,4	–
Oarai	O01 [†]	♂	nd	nd	I	4	a, d
	O03 [†]	♂	nd	nd	II	12	b, j
	O08 [†]	♀	nd	nd	II	6,13	e, e
	O10 [†]	♂	nd	nd	IV	14	e, i
	O22 [†]	♂	nd	nd	II	5	d, l
	O28 [†]	♀	nd	nd	I	3	a, b
	O31 [†]	♀	nd	nd	I	1,15	f, f
	O34 [†]	♀	nd	nd	I	16	b, b
	O36 [†]	♀	nd	nd	I	3	a, a
	O38 [†]	♂	nd	nd	II	5	a, g
	O39 [†]	♀	nd	nd	I	17	a, a
	O42 [†]	♂	nd	nd	V	4	a, a
	O43 [†]	♀	nd	nd	II	5	a, a
	O45 [†]	♀	nd	nd	I	7	l, l
	O48 [‡]	♀	O10	deceased	I	4	–
	O73 [§]	♂	deceased	O08	–	6,13	–
	O78 [‡]	♀	O01	deceased	I	7	–

Symbols (†, ‡, §) are as for Fig. 1.

nd, no data for parents; –, not analyzed.

et al., 2001). Genetic differentiation between two populations was estimated by analysis of molecular variance (AMOVA) using Arlequin 3.5 (Excoffier and Lischer, 2010). Character state-based phylogenetic networks for the haplotypes of each locus were constructed manually following Bandelt (1994) and Saitou and Yamamoto (1997). A phylogenetic tree was constructed using the neighbor-joining method (Saitou and Nei, 1987) with the distances corrected by the Kimura 2-parameter model (Kimura, 1980) in MEGA7. The bootstrap test was carried out with 1,000 replicates (Felsenstein, 1985).

RESULTS AND DISCUSSION

Nucleotide sequences of tufted puffin loci We obtained nucleotide sequences of 468 bp for *COX1*, 360 bp for *D-loop*, and 840 bp for *RHBG* in tufted puffins. For *RHBG*, a region containing part of exon 5 to part of exon 8 was sequenced. A section of intron 6 of *RHBG* could not be sequenced due to a poly(C) tract. All nucleotide sequence data obtained in this study were deposited in the DDBJ/EMBL/GenBank database under the accession numbers LC535135–LC535210.

We observed five haplotypes for *COX1* in the aquarium birds, in addition to two that were represented only in the BOLD data (Table 1, Fig. 3). In addition to the two populations sequenced in this study, we included sequences from BOLD in the phylogenetic network. In total, we found six segregating sites in the region; all were transitional and synonymous. Haplotype I was the major haplotype, and haplotypes III, IV and V each had one mutation differentiating them from haplotype I. Haplotype II was the second largest in this study, and haplotype VII differed by one mutation from haplotype II. Haplotype VI was located midway between haplotypes I and II.

We observed seventeen *D-loop* haplotypes in this study (Table 1, Fig. 4). Four individuals (K11, K44, O08 and O31) each had one heterogeneous site: K11 (158Y), K44 (48Y), O08 (86Y) and O31 (164R). To confirm these heterogeneous sites, we sequenced the *D-loop* of K62, a grandson of K11 (Fig. 1), and of O73, a son of O08 (Fig. 2). We obtained identical sequences, with the same heterogeneous sites, from K62 and K11 and from O73 and O08 (Fig. 4). It is therefore likely that these individuals are heteroplasmic, that is, they carry more than one type of mtDNA within each cell. Although O31 and K44

Fig. 3. Haplotype network for a section of the *COX1* locus sequence obtained for tufted puffins. The numbers on each branch are the positions of the mutations responsible for those edges. The circle at each node indicates the proportion of occurrences of that haplotype observed in the populations from Kasai (black); Oarai (gray); East Asia (Japan and far eastern Russia), taken from the BOLD database (dots); and North America (Canada and Alaska, USA), taken from the BOLD database (vertical lines). The size of each circle indicates the frequency of each haplotype (I: Kasai [6], Oarai [9], East Asia BOLD [3], North America BOLD [4]; II: Kasai [4], Oarai [5], East Asia BOLD [4], North America BOLD [1]; III: Kasai [1]; IV Oarai [1]; V: Oarai [1], East Asia BOLD [1]; VI: East Asia BOLD [1]; VII: North America BOLD [1]).

Fig. 4. Phylogenetic tree of partial sequences of the *D-loop* from individual tufted puffins. Black circles and gray squares represent individuals from Kasai and Oarai, respectively. Seventeen *D-loop* haplotypes were identified. Six individuals (K11, K44, K62, O08, O31 and O73) had two haplotypes, probably through heteroplasmy, and both haplotypes are shown in the tree. *COX1* haplotype clusters [I, III, V], [IV] and [II] are also indicated. Bootstrap values were less than 70% for all nodes and are not shown. Scale bar: 0.001 mutations per site.

also had one heterogeneous site each, this heterogeneity could not be confirmed, because O31 has no relatives in the aquarium population and K44 has no relatives on his maternal side. Berg et al. (1995) observed heteroplasmic variation in the *D-loop* region in four Laridae, ten Alcidae, and one Scolopacidae species. It therefore seems that Charadriiformes species tend to exhibit heteroplasmy in the *D-loop* region. Berg et al. (1995) observed tandem repeat heteroplasmy, whereas we observed only point mutations in this study; however, we did not sequence the heteroplasmic tandem repeat region, which is located at the 5' end of the *D-loop* region. Heteroplasmic sites represented by point mutations have also been observed in the same *D-loop* region in Cassin's auklet (*Ptychoramphus aleuticus*; Wallace et al., 2015).

Two large *COX1* haplotype groups (I and II) were observed for this species (Fig. 3). *D-loop* sequences from individuals who had *COX1* haplotype II formed a cluster in the *D-loop* tree (Fig. 4), although the bootstrap value was low (36%). Although bootstrap values for all nodes are low, compatibility between the *COX1* and *D-loop* sequences shows the reliability of our DNA sequencing,

since mtDNA is haploid and consistent haplotype patterns are expected from anywhere in the mitochondrial genome. In total, 10 segregating sites of the *D-loop* sequences were observed (Fig. 5). Of these, the 125th site had three character states (A: 1-2-3-4-7-8-9-14-15-16-17; T: 5-11-12; C: 6-10-13). There was therefore a total of 11 mutations. The changes between A and T and between A and C at the 125th site were transversions. All other changes were transitions.

The predicted haplotypes of *RHBG* are shown in Supplementary Table S1. The most likely haplotype pairs for each individual were used for the subsequent analysis (Table 1, Fig. 6). In total, 18 segregating sites were observed. Sites 34, 35 and 64 were in exon 5; sites 94 and 182 were in intron 5; sites 357, 371 and 372 were in exon 6; site 441 was in intron 6; sites 599, 632, 674, 675 and 698 were in exon 7; and sites 766, 767, 820 and 904 were in intron 7. Four nonsynonymous (34, 64, 371 and 675) and seven synonymous (35, 357, 372, 599, 632, 674 and 698) changes were observed. The changes at sites 94 and 182 were transversions and all others were transitions.

With respect to polymorphism, there were no large differences in θ_W or π values between the two populations for the two mtDNA loci (*COX1* and *D-loop*; Table 2). The *RHBG* sequences from Kasai ($\theta_W = 0.0059$, $\pi = 0.0059$) showed slightly higher nucleotide diversity than those

from Oarai ($\theta_W = 0.0040$, $\pi = 0.0030$). The levels of polymorphism within the *D-loop* sequence tended to be larger than those for *COX1* or *RHBG*. Nucleotide diversities of *D-loop* of several species of Charadriiformes were observed from the lower 0.00014 of *Larus cachinnans mongolicus* (excluding introgressed haplotypes) (Liebers et al., 2001) to the upper 0.019 of *Aethia cristatella* (St.-Jonah population) (Pshenichnikova et al., 2015) (Supplementary Table S2). Therefore, the nucleotide diversity of tufted puffins (Kasai: 0.0082, Oarai: 0.0080) was considered to be moderate in Charadriiformes bird species.

Fig. 5. Seventeen-haplotype network for partial sequences of the *D-loop* obtained for tufted puffins. The numbers on each edge are the positions of the mutations responsible for those edges; boldface denotes transversions. The circle at each node indicates the proportion of occurrences of that haplotype observed in the Kasai (black) and Oarai (gray) populations, and the size of the circle indicates the frequency of each haplotype (see Table 1).

Table 2. Polymorphism in the three tufted puffin loci sequenced

Locus	Population	<i>N</i>	<i>S</i>	<i>P_s</i>	θ_W	π
<i>COX1</i>	Kasai	11	3	0.0064	0.0022	0.0026
	Oarai	16	4	0.0085	0.0026	0.0025
	Merged	27	5	0.0107	0.0028	0.0024
<i>D-loop</i>	Kasai	13	8	0.0222	0.0072	0.0082
	Oarai	18	8	0.0222	0.0065	0.0080
	Merged	31	10	0.0278	0.0070	0.0081
<i>RHBG</i>	Kasai	12	15	0.0179	0.0059	0.0059
	Oarai	28	13	0.0155	0.0040	0.0030
	Merged	40	18	0.0214	0.0050	0.0040

N, number of sequences; *S*, number of segregating sites; *P_s*, number of segregating sites per total number of sites; θ_W , nucleotide diversity based on the proportion of segregating sites in a sample; π , average number of nucleotide differences per site between two sequences.

The *D-loop* sequences from K62 and O73 were excluded for these estimates.

Fig. 6. Haplotype network for a section of the *RHBG* sequence obtained for tufted puffins. The numbers on each edge are the positions of the mutations responsible for those edges; boldface denotes transversions, underlining denotes nonsynonymous changes, and italicization denotes substitutions in introns. The circle at each node (a–n) indicates the proportion of occurrences of that haplotype observed in the Kasai (black) and Oarai (gray) populations, and the size of the circle indicates the frequency of each haplotype (see Table 1).

No clear genetic differences between the two tufted puffin populations

We observed no clear genetic differences between the Kasai (originally from Alaska) and Oarai (originally from Kamchatka, far eastern Russia) populations in the three loci we analyzed. The major haplotype of *COX1*, haplotype I, consists of 10 individuals (six Kasai and four BOLD) from North American birds and 12 (nine Oarai and three BOLD) from East Asian birds. The second largest haplotype, II, consists of five individuals (four Kasai and one BOLD) from North American birds and nine (five Oarai and four BOLD) from East Asian birds (Fig. 3). In the phylogenetic network of the *D-loop* locus (Fig. 5), the largest haplotype, '4', consists of three Kasai and three Oarai sequences, and the other major haplotype groups, including '3' (two Kasai and two Oarai), '5' (one Kasai and three Oarai) and '7' (two Kasai and two Oarai), also consist of individuals from both populations. For *RHBG* (Fig. 6), the largest haplotype, 'a', consists of three Kasai and 11 Oarai sequences, and the second largest, 'l', consists of two Kasai and three Oarai sequences. We also performed AMOVA within and between the two populations (Table 3). Although *P*-values are not statistically significant at the 1% level, low fixation indices were observed for three loci (-0.0623 for *COX1*, -0.0128 for *D-loop* and 0.1072 for *RHBG*). Thus, it is clear that the two populations have not diverged genetically.

Saitoh et al. (2015) analyzed *COX1* from more than 200 bird species around the Japanese archipelago and showed that the average genetic distance within species (estimated using the Kimura 2-parameter method (Kimura, 1980)) was 0.49% (range: 0–6.13%). They showed that the genetic distance for the tufted puffin was on average 0.21%, with a maximum of 0.45%, from seven individuals (deposited in the BOLD system). Based on the merged Kasai and Oarai data in our study (27 individuals), the

estimated genetic distance was on average 0.25% with a maximum of 0.65%. These values are comparable with the values found by Saitoh et al. (2015) and fit within the range of within-species genetic variability of Japanese birds.

Although tufted puffins are known to breed throughout the boreal and low Arctic, from California to Hokkaido, their range during the feeding season is uncertain. They are thought to generally feed offshore throughout the year, and often far from the continental shelf in winter, and lay one egg per pair per year (Gaston and Jones, 1998). Although adults return to breeding colonies *en masse* in spring (Wehle, 1980), it is not known whether all adults return to their natal colony. It is likely that pair-bonded birds occupy the same burrows year after year, but information on nest-site fidelity and the duration of pair-bonds is scarce (Piatt and Kitaysky, 2002). One observational study estimated that a minimum of 29% of banded tufted puffins returned to monitored nest sites on Ugaiushak Island, Alaska (Wehle, 1980). Our molecular data suggest that the tufted puffin has not diverged genetically within its breeding range.

According to The IUCN Red List of Threatened Species (<https://dx.doi.org/10.2305/IUCN.UK.2018-2.RLTS.T22694934A132582357.en>), the worldwide population of tufted puffins is extremely large, and their Red List status is Least Concern. However, the current population trend is rated as “decreasing”. In Japan, which represents the southern limit of the breeding range, tufted puffin numbers have declined steeply (Senzaki et al., 2020). In 1959 and 1960, about 250 individuals for each year were observed on Moyururi island, Hokkaido (Fujimaki, 1961). In 1972, 111 individuals were observed on Moyururi and neighboring Yururi islands (Japan Environmental Agency, 1973), whereas Haga (1973) estimated a total of 61 individuals on Moyururi and Yururi islands. In

Table 3. AMOVA for *COX1*, *D-loop* and *RHBG* based on pairwise nucleotide differences

	<i>d.f.</i>	Sum of squares	Variance components	% of variation	Fixation index (<i>P</i>)
<i>COX1</i>					
Between populations	1	0.139	-0.0346 Va	-6.23	-0.0623 ($0.93939+ -0.00717$)
Within populations	25	14.75	0.5900 Vb	106.23	
Total	26	14.889	0.5554		
<i>D-loop</i>					
Between populations	1	1.18	-0.01843 Va	-1.28	-0.0128 ($0.50635+ -0.01566$)
Within populations	29	42.303	1.45874 Vb	101.28	
Total	30	43.484	1.44031		
<i>RHBG</i>					
Between populations	1	4.827	0.19211 Va	10.72	0.1072 ($0.03226+ -0.00595$)
Within populations	38	60.798	1.59994 Vb	89.28	
Total	39	65.625	1.79205		

d.f., degrees of freedom; *P*, *P*-value.

any case, until the 1970s, tufted puffins were observed breeding not only on Moyururi and Yururi islands but also in other regions of Japan (summarized in Osa and Watanuki (2002)). More recently, however, only about 15 breeding pairs were observed, and only on Yururi and Moyururi islands (Osa and Watanuki, 2002). In 2019, unusually high mortality of this species in the eastern Bering Sea was reported (Jones et al., 2019). This may have been the result of starvation caused by a warming climate. The present study has provided genetic diversity data for three loci (*COX1*, *D-loop* and *RHBG*) of the tufted puffin. Although the number of populations analyzed in the study was limited, we established that tufted puffins maintain a moderate level of genetic diversity by comparisons of nucleotide diversities in the *D-loop* region for several species of Charadriiformes. These genetic markers will facilitate further assessment of tufted puffin population dynamics.

We thank Ms. Yoko Osuka for giving us the opportunity to conduct this study. We are grateful to Mr. Sotaro Kawakami for assisting with the blood collection at Tokyo Sea Life Park. We thank two anonymous reviewers and the editor for their valuable comments.

REFERENCES

- American Ornithologists' Union (1998) Check-list of North American birds, 7th ed. American Ornithologists' Union, Washington, D.C.
- Bandelt, H.-J. (1994) Phylogenetic networks. *Verh. Naturwiss. Ver. Hamburg* **34**, 51–71.
- Berg, T., Moom, T., and Johansen, S. (1995) Variable numbers of simple tandem repeats make birds of the order Ciconiiformes heteroplasmic in their mitochondrial genomes. *Curr. Genet.* **27**, 257–262.
- Birt, T. P., Carter, H. R., Whitworth, D. L., McDonald, A., Newman, S. H., Gress, F., Palacios, E., Koepke, J. S., and Friesen, V. L. (2012) Rangewide population genetic structure of Xantus's murrelet (*Synthliboramphus hypoleucus*). *Auk* **129**, 44–55.
- Buehler, D. M., and Baker, A. J. (2003) Characterization of the red knot (*Calidris canutus*) mitochondrial control region. *Genome* **46**, 565–572.
- Buehler, D. M., and Baker, A. J. (2005) Population divergence times and historical demography in red knots and dunlins. *Condor* **107**, 497–513.
- Dehal, P., and Boore, J. L. (2005) Two rounds of whole genome duplication in the ancestral vertebrate. *PLoS Biol.* **3**, e314.
- Edgar, R. C. (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **32**, 1792–1797.
- Eisenmann, E., Monroe, B. L., Parkes, K. C., Short, L. L., Banks, R. C., Howell, T. R., Johnson, N. K., and Storer, R. W. (1982) Thirty-fourth supplement to the American Ornithologists' Union check-list of North American birds. *Auk* **99**, 1CC–16CC.
- Ewing, B., Hillier, L., Wendl, M. C., and Green, P. (1998) Base-calling of automated sequencer traces using *Phred*. I. Accuracy assessment. *Genome Res.* **8**, 175–185.
- Excoffier, L., and Lischer, H. E. L. (2010) Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Mol. Ecol. Res.* **10**, 564–567.
- Felsenstein, J. (1985) Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* **39**, 783–791.
- Fujimaki, Y. (1961) Observations on sea birds of Moyururi Island, eastern Hokkaido. *Jap. J. Ornithol.* **16**, 387–398 (in Japanese).
- Gaston, A. J., and Jones, I. L. (1998) *The Auks: Alcidae*. Oxford University Press, New York, USA.
- Gordon, D., Abajian, C., and Green, P. (1998) Consed: a graphical tool for sequence finishing. *Genome Res.* **8**, 195–202.
- Haga, R. (1973) Avifauna of Yururi and Moyururi Islands, Hokkaido. In *Report on the Survey of Natural Community at Yururi and Moyururi Islands*. (ed.: Nemuro City Board of Education), pp. 31–61. Nemuro City Board of Education, Nemuro. (in Japanese)
- Hebert, P. D., Stoeckle, M. Y., Zemlak, T. S., and Francis, C. M. (2004) Identification of birds through DNA barcodes. *PLoS Biol.* **2**, e312.
- Huang, C.-H., and Peng, J. (2005) Evolutionary conservation and diversification of Rh family genes and proteins. *Proc. Natl. Acad. Sci. USA* **102**, 15512–15517.
- Japan Environmental Agency (1973) Report on the survey of rare species of wild birds. Japan Environmental Agency, Tokyo (in Japanese).
- Jones, T., Divine, L. M., Renner, H., Knowles, S., Lefebvre, K. A., Burgess, H. K., Wright, C., and Parrish, J. K. (2019) Unusual mortality of Tufted puffins (*Fratercula cirrhata*) in the eastern Bering Sea. *PLoS One* **14**, e0216532.
- Kerr, K. C. R., Birks, S. M., Kalyakin, M. V., Red'kin, Y. A., Koblik, E. A., and Hebert, P. D. N. (2009) Filling the gap - COI barcode resolution in eastern Palearctic birds. *Front. Zool.* **6**, 29.
- Kerr, K. C. R., Stoeckle, M. Y., Dove, C. J., Weigt, L. A., Francis, C. M., and Hebert, P. D. N. (2007) Comprehensive DNA barcode coverage of North American birds. *Mol. Ecol. Notes* **7**, 535–543.
- Kimura, M. (1980) A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *J. Mol. Evol.* **16**, 111–120.
- Kitano, T., Satou, M., and Saitou, N. (2010) Evolution of two Rh blood group-related genes of the amphioxus species *Branchiostoma floridae*. *Genes Genet. Syst.* **85**, 121–127.
- Kumar, S., Stecher, G., and Tamura, K. (2016) MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol. Biol. Evol.* **33**, 1870–1874.
- Küpper, C., Edwards, S. V., Kosztolányi, A., Alrashidi, M., Burke, T., Herrmann, P., Argüelles-Tico, A., Amat, J. A., Amezian, M., Rocha, A., et al. (2012) High gene flow on a continental scale in the polyandrous Kentish plover *Charadrius alexandrinus*. *Mol. Ecol.* **21**, 5864–5879.
- Liebers, D., de Knijff, P., and Helbig, A. J. (2004) The herring gull complex is not a ring species. *Proc. R. Soc. Lond. B* **271**, 893–901.
- Liebers, D., and Helbig, A. J. (2002) Phylogeography and colonization history of Lesser Black-backed Gulls (*Larus fuscus*) as revealed by mtDNA sequences. *J. Evol. Biol.* **15**, 1021–1033.
- Liebers, D., Helbig, A. J., and de Knijff, P. (2001) Genetic differentiation and phylogeography of gulls in the *Larus cachinnans-fuscus* group (Aves: Charadriiformes). *Mol. Ecol.* **10**, 2447–2462.
- Miller, M. P., Haig, S. M., Gratto-Trevor, C. L., and Mullins, T. D. (2010) Subspecies status and population genetic structure

- in Piping Plover (*Charadrius melodus*). *Auk* **127**, 57–71.
- Miller, M. P., Haig, S. M., Mullins, T. D., Ruan, L., Casler, B., Dondua, A., Gates, H. R., Johnson, J. M., Kendall, S., Tomkovich, P. S., et al. (2015) Intercontinental genetic structure and gene flow in Dunlin (*Calidris alpina*), a potential vector of avian influenza. *Evol. Appl.* **8**, 149–171.
- Osa, Y., and Watanuki, Y. (2002) Status of seabirds breeding in Hokkaido. *J. Yamashina Inst. Ornithol.* **33**, 107–141.
- Pallas, P. S. (1811) Zoographia Rosso-Asiatica, sistens omnium animalium in extenso Imperio Rossico et adjacentibus Maribus observatorum recensionem, domicilia, mores et descriptiones, anatomen atque icones plurimorum, vol. 2, pp. 363–365. In officina Caes. Academiae Scientiarum Impress, Petropoli.
- Peng, J., and Huang, C. H. (2006) Rh proteins vs Amt proteins: an organismal and phylogenetic perspective on CO₂ and NH₃ gas channels. *Transfus. Clin. Biol.* **13**, 85–94.
- Pereira, S. L., and Baker, A. J. (2008) DNA evidence for a Paleocene origin of the Alcidae (Aves: Charadriiformes) in the Pacific and multiple dispersals across northern oceans. *Mol. Phylogenet. Evol.* **46**, 430–445.
- Piatt, J. F., and Kitaysky, A. S. (2002) Tufted puffin. *Birds of North America* **708**, 1–32.
- Pons, J.-M., Sonsthagen, S., Dove, C., and Crochet, P.-A. (2014) Extensive mitochondrial introgression in North American Great Black-backed Gulls (*Larus marinus*) from the American Herring Gull (*Larus smithsonianus*) with little nuclear DNA impact. *Heredity* **112**, 226–239.
- Pshenichnikova, O. S., Klenova, A. V., Sorokin, P. A., Zubakin, V. A., Konyukhov, N. B., Kharitonov, S. P., and Artukhin, Y. B. (2015) The Crested Auklet, *Aethia cristatella* (Alcidae, Charadriiformes), does not vary geographically in genetics, morphology or vocalizations. *Mar. Biol.* **162**, 1329–1342.
- Ratnasingham, S., and Hebert, P. D. N. (2007) BOLD: The Barcode of Life Data System (www.barcodinglife.org). *Mol. Ecol. Notes* **7**, 355–364.
- Rheindt, F. E., Székely, T., Edwards, S. V., Lee, P. L. M., Burke, T., Kennerley, P. R., Bakewell, D. N., Alrashidi, M., Kosztolányi, A., Weston, M. A., et al. (2011) Conflict between genetic and phenotypic differentiation: the evolutionary history of a ‘lost and rediscovered’ shorebird. *PLoS One* **6**, e26995.
- Saitoh, T., Sugita, N., Someya, S., Iwami, Y., Kobayashi, S., Kamigaichi, H., Higuchi, A., Asai, S., Yamamoto, Y., and Nishiumi, I. (2015) DNA barcoding reveals 24 distinct lineages as cryptic bird species candidates in and around the Japanese Archipelago. *Mol. Ecol. Resour.* **15**, 177–186.
- Saitou, N., and Nei, M. (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* **4**, 406–425.
- Saitou, N., and Yamamoto, F. I. (1997) Evolution of primate ABO blood group genes and their homologous genes. *Mol. Biol. Evol.* **14**, 399–411.
- Senzaki, M., Terui, A., Tomita, N., Sato, F., Fukuda, Y., Kataoka, Y., and Watanuki, Y. (2020) Long-term declines in common breeding seabirds in Japan. *Bird Conserv. Int.* **30**, 434–446.
- Sonsthagen, S. A., Tibbitts, T. L., Gill, R. E., Williams, I., and Talbot, S. L. (2015) Spatial genetic structure of bristle-thighed curlews (*Numenius tahitiensis*): breeding area differentiation not reflected on the non-breeding grounds. *Conserv. Genet.* **16**, 223–233.
- Stephens, M., Smith, N. J., and Donnelly, P. (2001) A new statistical method for haplotype reconstruction from population data. *Am. J. Hum. Genet.* **68**, 978–989.
- Sternkopf, V., Liebers-Helbig, D., Ritz, M. S., Zhang, J., Helbig, A. J., and de Knijff, P. (2010) Introgressive hybridization and the evolutionary history of the herring gull complex revealed by mitochondrial and nuclear DNA. *BMC Evol. Biol.* **10**, 348.
- Suzuki, A., Endo, K., and Kitano, T. (2014) Phylogenetic positions of RH blood group-related genes in cyclostomes. *Gene* **543**, 22–27.
- Suzuki, A., Komata, H., Iwashita, S., Seto, S., Ikeya, H., Tabata, M., and Kitano, T. (2017) Evolution of the RH gene family in vertebrates revealed by brown hagfish (*Eptatretus atami*) genome sequences. *Mol. Phylogenet. Evol.* **107**, 1–9.
- Tavares, E. S., and Baker, A. J. (2008) Single mitochondrial gene barcodes reliably identify sister-species in diverse clades of birds. *BMC Evol. Biol.* **8**, 81.
- Tigano, A., Damus, M., Birt, T. P., Morris-Pocock, J. A., Artukhin, Y. B., and Friesen, V. L. (2015) The arctic: glacial refugium or area of secondary contact? Inference from the population genetic structure of the Thick-Billed Murre (*Uria lomvia*), with implications for management. *J. Hered.* **106**, 238–246.
- Trimbos, K. B., Doorenweerd, C., Kraaijeveld, K., Musters, C. J. M., Groen, N. M., de Knijff, P., Piersma, T., and de Snoo, G. R. (2014) Patterns in nuclear and mitochondrial DNA reveal historical and recent isolation in the Black-tailed Godwit (*Limosa limosa*). *PLoS One* **9**, e83949.
- Verkuil, Y. I., Piersma, T., Jukema, J., Hooijmeijer, J. C. A. W., Zwarts, L., and Baker, A. J. (2012) The interplay between habitat availability and population differentiation: a case study on genetic and morphological structure in an inland wader (Charadriiformes). *Biol. J. Linn. Soc.* **106**, 641–656.
- Wallace, S. J., Wolf, S. G., Bradley, R. W., Laurie Harvey, A., and Friesen, V. L. (2015) The influence of biogeographical barriers on the population genetic structure and gene flow in a coastal Pacific seabird. *J. Biogeogr.* **42**, 390–400.
- Wang, X., Que, P., Heckel, G., Hu, J., Zhang, X., Chiang, C.-Y., Zhang, N., Huang, Q., Liu, S., Martinez, J., et al. (2019) Genetic, phenotypic and ecological differentiation suggests incipient speciation in two *Charadrius* plovers along the Chinese coast. *BMC Evol. Biol.* **19**, 135.
- Wehle, D. H. S. (1980) The breeding biology of the puffins: tufted puffin (*Lunda cirrhata*), horned puffin (*Fratercula corniculata*), common puffin (*F. arctica*), and rhinoceros auklet (*Cerorhinca monocerata*). PhD dissertation, University of Alaska, Fairbanks, Alaska.
- Zhang, G., Li, C., Li, Q., Li, B., Larkin, D. M., Lee, C., Storz, J. F., Antunes, A., Greenwold, M. J., Meredith, R. W., et al. (2014) Comparative genomics reveals insights into avian genome evolution and adaptation. *Science* **346**, 1311–1320.

Supplementary Table S1. Predicted haplotypes of a section of the tufted puffin *RHBG* gene

Individual	Predicted haplotype 1	Predicted haplotype 2	Probability
K01	CGGAACCGGCGCGGCACC	CGGAACTGGCGCGGCATT	0.170
	CGGAACCGGCGCGGCATC	CGGAACTGGCGCGGCACT	0.141
	CGGAACCGGCGCGGCATT	CGGAACTGGCGCGGCACC	0.689 (c, f)
K02	CGGAACCGGCGCGGCACC	CGGAACCGGCGCGGCACC	1.000 (a, a)
K11	CGGAACCGGCGCGGCGCC	CGAAATTGGTGTGACGCC	0.011
	CGGAACCGGCGTGGCGCC	CGAAATTGGTGCACGCC	0.018
	CGGAACCGGTGCGACGCC	CGAAATTGGCGTGGCGCC	0.094
	CGGAACCGGTGTGACGCC	CGAAATTGGCGCGGCGCC	0.028
	CGGAACTGGCGCGGCGCC	CGAAATCGGTGTGACGCC	0.051
	CGGAACTGGCGTGGCGCC	CGAAATCGGTGCGACGCC	0.089
	CGGAACTGGTGTGACGCC	CGAAATCGGCGTGGCGCC	0.173
	CGGAACTGGTGTGACGCC	CGAAATCGGCGCGGCGCC	0.057
	CGGAATCGGCGTGGCGCC	CGAAACTGGTGTGACGCC	0.026
	CGGAATCGGTGCGACGCC	CGAAACTGGCGTGGCGCC	0.028
	CGGAATCGGTGTGACGCC	CGAAACTGGCGCGGCGCC	0.010
	CGGAATTGGCGCGGCGCC	CGAAACCGGTGTGACGCC	0.079
	CGGAATTGGCGTGGCGCC	CGAAACCGGTGCGACGCC	0.233 (h, k)
	CGGAATTGGTGTGACGCC	CGAAACCGGCGTGGCGCC	0.061
	CGGAATTGGTGTGACGCC	CGAAACCGGCGCGGCGCC	0.033
K12	CAGCACCAGTGTGACGCC	CAGCACCAGTGTGACGCC	1.000 (l, l)
K19	CAGCACCAGTGTGATGCC	CAGCACCAGTGTGATGCC	1.000 (m, m)
K30	CGGAACCGGCGCGGCACC	TGAAACCGGCACGGCACC	0.400 (a, n)
	CGGAACCGGCACGGCACC	TGAAACCGGCGCGGCACC	0.172
	CGAAACCGGCGCGGCACC	TGGAACCGGCACGGCACC	0.259
	CGAAACCGGCACGGCACC	TGGAACCGGCGCGGCACC	0.170
O01	CGGAACCGGCGCGGCACC	CGGAACCGGCGCAGCATT	0.990 (a, d)
O03	CGGAACCGGCGCGGCACC	CGAAACCGGCGCGGCATC	0.389
	CGGAACCGGCGCGGCATC	CGAAACCGGCGCGGCACC	0.611 (b, j)
O08	CGGAACCAGCGCGGCACC	CGGAACCAGCGCGGCACC	1.000 (e, e)
O10	CGGAACCGGCGCGGCACC	CGGACCCAGCGCGGCATT	0.076
	CGGAACCGGCGCGGCATC	CGGACCCAGCGCGGCACT	0.066
	CGGAACCGGCGCGGCATT	CGGACCCAGCGCGGCACC	0.369
	CGGAACCAGCGCGGCACC	CGGACCCGGCGCGGCATT	0.401 (e, i)
	CGGAACCAGCGCGGCACT	CGGACCCGGCGCGGCATC	0.035
	CGGAACCAGCGCGGCATT	CGGACCCGGCGCGGCACC	0.052

Supplementary Table S1. Continued

Individual	Predicted haplotype 1	Predicted haplotype 2	Probability
O22	CGGAACCGGCGCGGCATT	CAGCACCAAGTGCAACGCC	0.020
	CGGAACCGGCGCAGCATT	CAGCACCAAGTGCGACGCC	0.979 (d, l)
O28	CGGAACCGGCGCGGCACC	CGGAACCGGCGCGGCATC	1.000 (a, b)
O31	CGGAACCTGGCGCGGCACC	CGGAACCTGGCGCGGCACC	1.000 (f, f)
O34	CGGAACCGGCGCGGCATC	CGGAACCGGCGCGGCATC	1.000 (b, b)
O36	CGGAACCGGCGCGGCACC	CGGAACCGGCGCGGCACC	1.000 (a, a)
O38	CGGAACCGGCGCGGCACC	CGGAACCTGACGCGGCACC	0.506 (a, g)
	CGGAACCGACGCGGCACC	CGGAACCTGGCGCGGCACC	0.494
O39	CGGAACCGGCGCGGCACC	CGGAACCGGCGCGGCACC	1.000 (a, a)
O42	CGGAACCGGCGCGGCACC	CGGAACCGGCGCGGCACC	1.000 (a, a)
O43	CGGAACCGGCGCGGCACC	CGGAACCGGCGCGGCACC	1.000 (a, a)
O45	CAGCACCAAGTGCGACGCC	CAGCACCAAGTGCGACGCC	1.000 (l, l)

Letters in parentheses in the Probability column represent the predicted haplotype.

Supplementary Table S2. Nucleotide diversity (π) in *D-loop* regions of Charadriiformes species

Family	Species	π	Reference	Note
Alcidae	<i>Aethia cristatella</i>	0.013	Pshenichnikova et al., 2015	Talan population
	<i>Aethia cristatella</i>	0.013	Pshenichnikova et al., 2015	Kamchatka population
	<i>Aethia cristatella</i>	0.013	Pshenichnikova et al., 2015	Kuril area population
	<i>Aethia cristatella</i>	0.012	Pshenichnikova et al., 2015	Medny population
	<i>Aethia cristatella</i>	0.019	Pshenichnikova et al., 2015	St.-Jonah population
	<i>Fratercula cirrhata</i>	0.0082	This study	Kasai
	<i>Fratercula cirrhata</i>	0.0080	This study	Oarai
	<i>Ptychoramphus aleuticus</i>	0.006	Wallace et al., 2015	Northern group
	<i>Ptychoramphus aleuticus</i>	0.007	Wallace et al., 2015	Southern group
	<i>Synthliboramphus hypoleucus craveri</i>	0.00795	Birt et al., 2012	
	<i>Synthliboramphus hypoleucus hypoleucus</i>	0.00596	Birt et al., 2012	
	<i>Synthliboramphus hypoleucus scrippsi</i>	0.00715	Birt et al., 2012	
	<i>Uria lomvia</i>	0.013	Tigano et al., 2015	
Charadriidae	<i>Charadrius alexandrinus</i>	0.0047	Küpper et al., 2012	Estimate of all Mainland
	<i>Charadrius alexandrinus</i>	0.00181	Wang et al., 2019	
	<i>Charadrius dealbatus</i>	0.00086	Wang et al., 2019	
	<i>Charadrius melodus</i>	0.0039	Miller et al., 2010	Prairie Canada population
	<i>Charadrius melodus</i>	0.0028	Miller et al., 2010	U.S. Great Plains population
	<i>Charadrius melodus</i>	0.002	Miller et al., 2010	Great Lakes population
	<i>Charadrius melodus</i>	0.0032	Miller et al., 2010	Atlantic Canada population
	<i>Charadrius melodus</i>	0.0056	Miller et al., 2010	Atlantic United States population
Laridae	<i>Larus cachinnans armenicus</i>	0.00171	Liebers et al., 2001	Excluding introgressed haplotypes
	<i>Larus cachinnans atlantis</i>	0.00529	Liebers et al., 2001	Excluding introgressed haplotypes
	<i>Larus cachinnans barabensis</i>	0.00081	Liebers et al., 2001	Excluding introgressed haplotypes

Supplementary Table S2. Continued

Family	Species	π	Reference	Note
Laridae	<i>Larus cachinnans cachinnans</i>	0.00991	Liebers et al., 2001	
	<i>Larus cachinnans michahellis</i>	0.00407	Liebers et al., 2001	
	<i>Larus cachinnans mongolicus</i>	0.00014	Liebers et al., 2001	Excluding introgressed haplotypes
	<i>Larus fuscus fuscus</i>	0.00333	Liebers and Helbig, 2002	
	<i>Larus fuscus graellsii</i>	0.00175	Liebers et al., 2001	
	<i>Larus fuscus graellsii</i>	0.00175	Liebers and Helbig, 2002	
	<i>Larus fuscus heuglini</i>	0.00223	Liebers and Helbig, 2002	
	<i>Larus fuscus heuglini</i>	0.00188	Liebers et al., 2001	
	<i>Larus fuscus intermedius</i>	0.00232	Liebers and Helbig, 2002	
	<i>Larus fuscus taimyrensis</i>	0.00421	Liebers and Helbig, 2002	
	<i>Larus fuscus taimyrensis</i>	0.0044	Liebers et al., 2001	
	<i>Larus marinus</i>	0.00877	Pons et al., 2014	Europe
	<i>Larus marinus</i>	0.00738	Pons et al., 2014	North America non-introgressed
Scolopacidae	<i>Calidris alpina</i>	0.0036	Buehler and Baker, 2005	Average estimate
	<i>Calidris alpina actites</i>	0.0019	Miller et al., 2015	
	<i>Calidris alpina arctica</i>	0.0015	Miller et al., 2015	
	<i>Calidris alpina hudsonia</i>	0.0023	Miller et al., 2015	
	<i>Calidris alpina kistchinski</i>	0.0028	Miller et al., 2015	
	<i>Calidris alpina pacifica</i>	0.0023	Miller et al., 2015	
	<i>Calidris alpina sakhalina</i>	0.0058	Miller et al., 2015	
	<i>Calidris canutus</i>	0.002	Buehler and Baker, 2005	Average estimate
	<i>Numenius tahitiensis</i>	0.004	Sonsthagen et al., 2015	Seward Peninsula
	<i>Numenius tahitiensis</i>	0.006	Sonsthagen et al., 2015	Nulato Hills
	<i>Numenius tahitiensis</i>	0.006	Sonsthagen et al., 2015	Oceania
	<i>Philomachus pugnax</i>	0.0069	Verkuil et al., 2012	All