Ornithology AmericanOrnithology.org

Volume 139, 2022, pp. 1–14
https://doi.org/10.1093/ornithology/ukab081

RESEARCH ARTICLE

Distinctive mitogenomic lineages within populations of White-tailed Eagles

Charles Christian Riis Hansen,[1,*] Sina Baleka,[1] Sólveig Magnea Guðjónsdóttir,[1] Jacob Agerbo Rasmussen,[2,3] Jesus Adrian Chimal Ballesteros,[2,4] Gunnar Thor Hallgrimsson,[1] Robert A. Stefansson,[5] Menja von Schmalensee,[5] Kristinn Haukur Skarphédinsson,[6] Aili Lage Labansen,[7] Madis Leivits,[8] Kim Skelmose,[9] Christian Sonne,[10] Rune Dietz,[10] David Boertmann,[10] Igor Eulaers,[10] Michael D. Martin,[2] and Snæbjörn Pálsson[1]

[1] Faculty of Life and Environmental Sciences, University of Iceland, Reykjavik, Iceland
[2] Department of Natural History, NTNU University Museum, Norwegian University of Science and Technology (NTNU), Trondheim, Norway
[3] Section for EvoGenomics, The Globe Institute, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark
[4] Natural History Museum, University of Oslo, Oslo, Norway
[5] West-Iceland Nature Research Centre, Stykkisholmur, Iceland
[6] Icelandic Institute of Natural History, Garðabær, Iceland
[7] Department of Birds and Mammals, Greenland Institute of Natural Resources, Nuuk, Greenland
[8] Chair of Clinical Veterinary Medicine, Institute of Veterinary Medicine and Animal Sciences, Estonian University of Life Sciences, Tartu, Estonia
[9] Natural History Museum of Denmark, University of Copenhagen, Copenhagen, Denmark
[10] Department of Bioscience, Aarhus University, Roskilde, Denmark
* Corresponding author: ccr3@hi.is

Submission Date: March 22, 2021; Editorial Acceptance Date: October 26, 2021; Published December 22, 2021

ABSTRACT

Using whole mitochondrial DNA sequences from 89 White-tailed Eagles (*Haliaeetus albicilla*) sampled from Iceland, Greenland, Norway, Denmark and Estonia between 1990 and 2018, we investigate the mitogenomic variation within and between countries. We show that there is a substantial population differentiation between the countries, reflecting similar major phylogeographic patterns obtained previously for the control region of the mitochondria, which suggested two main refugia during the last glacial period. Distinct mitogenomic lineages are observed within countries with divergence times exceeding the end of the last glacial period of the Ice Age. Deviations from neutrality indicate that these lineages have been maintained by natural selection and there is an excess of segregating amino acids in comparison with number of fixations suggesting a large load of deleterious mutations. The maintenance of the distinct mitogenic lineages within countries inflates our estimates of divergence times.

Keywords: conservation, divergence time, phylogeography, selection, White-tailed Eagle

LAY SUMMARY

- Whole mitochondrial genomes were used to examine the population genetics of White-tailed Eagles.
- Large genetic differences potentially upheld by selection within populations were identified.
- Potential time of population splits during the last Ice age were identified.

Lignées mitogénomiques distinctes au sein de populations de *Haliaeetus albicilla*

RÉSUMÉ

En utilisant les séquences entières d'ADN mitochondrial de 89 individus de *Haliaeetus albicilla* échantillonnés en Islande, au Groenland, en Norvège, au Danemark et en Estonie entre 1990 et 2018, nous examinons la variation mitogénomique dans chaque pays et entre les pays. Nous montrons qu'il existe une différenciation substantielle des populations entre les pays, reflétant des patrons phylogéographiques principaux similaires obtenus préalablement pour la région de contrôle mitochondriale, ce qui suggérait la présence de deux refuges principaux au cours de la dernière période de glaciation. Des lignées mitogénomiques distinctes sont observées dans chaque pays avec des temps de divergence dépassant la

fin de la dernière période de glaciation de l'ère glaciaire. Les déviations de la neutralité indiquent que ces lignées ont été maintenues par la sélection naturelle et qu'il existe un excès d'acides aminés pouvant ségréguer par rapport au nombre de fixations, suggérant une grande charge de mutations délétères. Le maintien des lignées mitogènes distinctes au sein des pays gonfle nos estimations des temps de divergence.

Mots-clés: conservation, *Haliaeetus albicilla* temps de divergence, phylogéographie, sélection,

INTRODUCTION

The White-tailed Eagle (*Haliaeetus albicilla*) is a top predator distributed over most of northern Eurasia, from Greenland and Iceland in the west to Japan in the east, and as far south as North Africa (BirdLife International 2016). Phylogeographic studies based on sequences of the mitochondrial control region and microsatellites have described a major split between European and Asian populations, with a large contact zone (Hailer et al. 2007, Honnen et al. 2010, Langguth et al. 2013). Two mitochondrial haplotypes have been observed in the control region in samples from Greenland and Iceland, which are only separated by a single mutation, one which is unique to the two islands and the other which is common in northern Europe, where more general variation was also observed (Hailer et al. 2007). The low diversity in the island populations may reflect both a possible bottleneck during colonization of these islands and small population sizes. The number of White-tailed Eagles plummeted in Iceland in the late 19th century as in many other countries in Europe (Langguth et al. 2013, Treinys et al. 2016). The population in Iceland experienced a reduction from ~150 pairs in the mid-nineteenth century to ~20 pairs or less in the early 20th century; and despite a conservation law passed in 1915, the population did not increase in numbers until after 1967, following a ban on fox poisoning in 1964 (Petersen 1998, Skarphéðinsson 2013). As with other European populations, the Icelandic population has increased in number to ~80 pairs (Skarphéðinsson 2013). The Greenlandic population is larger than the Icelandic population, with 150–200 breeding pairs (Greenland Institute of Natural Resources 2020); however, the Greenlandic population also experienced a reduction in the 20th century and was estimated to 50–75 pairs around 1950 (Hansen 1979). The mainland population in Europe consists of approximately 17,900–24,500 pairs (BirdLife International 2016), with the Norwegian population being the largest of all European countries, counting ~2,000 breeding pairs (Jais 2020). With the population in mainland Europe expanding after successful conservation programs, White-tailed Eagles recolonized Denmark in 1995 after being extinct for more than 50 years, and in 2016, 61 breeding pairs were recorded (Kongeaastien 2020). In Estonia the population decreased from 400–500 breeding pairs (Lõhmus 1998) to ~20 pairs in the late 19th century, having years without successful breeding (Randla and Õun 1980). Currently, the Estonian White-tailed Eagle breeding population is estimated to be 290–330 pairs with a strong positive trend (Elts et al. 2019).

White-tailed eagles are to a large extent sedentary and display philopatry (Hailer et al. 2007), even though they have been shown to migrate for more than 2,000 km outside the breeding season (Ueta et al. 1998, Bragin et al. 2018). Between mainland Europe, Iceland, and Greenland there are no documented migrants. The mitochondrial DNA (mtDNA) diversity and an earlier classification of White-tailed Eagles in Greenland (and potentially Iceland) as a subspecies, primarily due to their large size (Salomonsen 1979) suggests that the 2 island populations have been isolated for centuries. Comparison among the mtDNA sequences can be used to date the colonization of the island populations and to determine their divergence. However, such time estimates and other predictions including effective population size based on the neutral theory of molecular variation can be affected by natural selection (Zink and Barrowclough 2008). Furthermore, application of a molecular clock rate obtained from comparison between species may also lead to an overestimate of the divergence time within species (Ho et al. 2007).

The small Icelandic and potentially other White-tailed Eagle populations are predicted to have little variation in the haploid, maternally inherited mtDNA, and this variation can be further reduced by directional selection either due to background selection (Charlesworth et al. 1993) or genetic hitchhiking (Smith and Haigh 1974) due to lack of recombination. In bird populations, it is expected that such linked effect is augmented due to linkage of mtDNA with the W-chromosome through shared maternal transmission; thus, selection on the W-chromosome will also affect the mitochondrial DNA. Accordingly, lower variation has been observed in avian mtDNA than in mammalian species, which do not have this linkage with the heterogametic sex chromosome (Berlin et al. 2007). In birds, however, Berlin et al. (2007) also observed a higher neutrality index (NI) than in mammals. The NI quantifies the direction of the selection, a positive NI indicates a negative selection or increased mutational load, where more deleterious mutations are segregating within species than would be expected by the number of fixed differences between species, possibly due to reduced efficacy of selection in small populations (Lynch 1996, Neiman and Taylor 2009).

Here we investigate the genetic variation using complete mtDNA genomes of White-tailed Eagles from Iceland and Greenland in comparison with the much larger mainland population in Norway, Estonia, and the recently established population in Denmark. We evaluate signals of selection, the impact of different population sizes and recent

bottlenecks, and investigate the split between the island populations of Iceland and Greenland from the mainland populations with the aim to reconstruct the population history and origin of these populations of White-tailed Eagles based on mtDNA lineages.

MATERIAL AND METHODS

Samples and Sequencing

Samples (blood or muscle) were obtained from 89 individuals from 5 countries; Greenland (n = 12; 1990–2017), Iceland (n = 42; 2003–2011), Norway (n = 21; 2001–2015), Denmark (n = 11; 2015–2018), and Estonia (n = 3; 2015), sampled during 1990–2018. An overview of sampling sites and year of sampling for all individuals studied is presented in Figure 1 (and in more detail in Supplementary Material Table S1). Blood samples from Iceland were collected from 42 chicks from Northwest Iceland, from 27 nest territories (64.38°N–65.87°N, 23.51°E–21.33°E), in an ongoing monitoring of the White-tailed Eagles in Iceland (led by the Icelandic Institute of Natural History). Three to 10 mL of blood were extracted from each chick. Blood was stored in EDTA buffer at −20°C until DNA extraction. Whole blood and muscle tissue samples from the Estonian, Danish, and Greenland individuals were stored at −20°C until DNA extraction, and were provided by the Department of Bioscience, Arctic Research Centre, AU, Roskilde, Denmark (Estonian, Danish, and Greenland samples); Natural History Museum of Denmark, University of Copenhagen, Copenhagen, Denmark (Danish samples); and the Greenland Institute of Natural Resources, Nuuk, Greenland (Greenland samples).

DNA from the 68 individuals from Iceland, Greenland, Denmark, and Estonia was extracted using Thermo Fisher GeneJET Whole Blood Genomics DNA Purification Mini Kit (Thermo Fisher Scientific, Waltham, Massachussetts, USA) following the standard protocol (Thermo Fisher 2006) for blood samples and Macherey-Nagel NucleoMag 96 Tissue kit (Macherey-Nagel 2014) for tissue at the University of Iceland. The concentration of the DNA was >20 ng μL^{-1} estimated using the NanoDrop 1000 (Thermo Fisher). The 68 samples were sequenced at BGI Hong Kong using DNBseq Normal DNA library construction and DNBseq PE150.

DNA from the 21 Norwegian individuals was processed at the NTNU University Museum's standard molecular genetics laboratories, and extractions were performed with a Qiagen DNeasy Blood & Tissue kit (Qiagen, Valencia, California, USA). The manufacturer's protocol was followed (Qiagen 2006) except that the amount of proteinase K used in the lysis step was doubled, and the lysis step incubation at 56°C was extended to 15 hr. The DNA solutions were incubated at 37°C for 10 min prior to elution. A blank control extraction was performed alongside all tissue extractions. The DNA extracts were sheared to a mean fragment size of 350 base pairs (bp) using a Covaris LE220 focused ultrasonicator instrument, and then short fragments were removed via size-selection using SPRI beads prepared as in Rohland and Reich (2012). Blunt-end Illumina library preparation was conducted on the 21 Norwegian DNA extractions using the Blunt-End-Sigle-Tube (BEST) protocol (Carøe et al. 2018), during which customized blunt-end adapters (Kircher et al. 2012) were ligated to the genomic DNA fragments. Sample-specific, dual-indexing barcodes were incorporated into each library using custom, indexed primers during the indexing polymerase chain reactions (PCR). Library indexing and amplification was carried out in 100 µL PCR reactions with 7 µL library template, 0.25 mM each dNTP, 0.25 µM forward primer, 0.25 µM reverse primer, 1 µL Herculase II Fusion DNA polymerase, 20 µl 5X Herculase II Reaction Buffer, and 65.8 µL molecular biology H_2O. For each library, an optimal number of indexing PCR cycles was determined using quantitative PCR. The indexing PCRs were performed under the following conditions: 95°C for 3 min, 13–21 cycles of 95°C for 20 s, 60°C for 20 s, 72°C for 40 s, with a final extension of 72°C for 5 min. The amplified libraries were purified using SPRI beads (Rohland and Reich 2012) prior to elution into EB buffer (Qiagen). Purified, indexed libraries were pooled and sequenced over 2 runs on the Illumina HiSeq 4000 platform at the NTNU Genomics Core Facility and over 1 run on the Illumina NovaSeq 6000 platform at the University of Oslo Norwegian National Sequencing Centre.

Mapping

Quality of the 89 individual FASTQ files was checked with FastQC (Babraham Bioinformatics 2010), and then run through AdapterRemoval 2 (Schubert et al. 2016) to remove potential adapters when appropriate, using standard setting with the following customizing arguments: —collapse and—trimns. Individuals were mapped to the White-tailed Eagle mitochondrial genome of a Korean specimen (NCBI: NC_040858.1) (Kim et al. 2019) with BWA-MEM and converted to BAM files using *samtools view and sort* (Li and Durbin 2009, Li et al. 2009). BAM files were transformed to variant call format (VCF) and FASTA files using *samtools mpileup* with quality filters –q 30 and –Q 20, and *bcftools call* (Li et al. 2009). The Korean reference sequence includes a nuclear mitochondrial DNA segment (NUMT) in the 12s rRNA from position 492-538 (both included) that was removed from the alignment, resulting in total sequence length of 17,461 bp. Quality, depth, ratio of the common allele, and binomial probabilities for all segregating sites were evaluated from the VCF file. A recently published mitochondrial genome of a German specimen (NCBI: MN356434.1) was retrieved from GenBank and compared to our sequences (Feng et al.

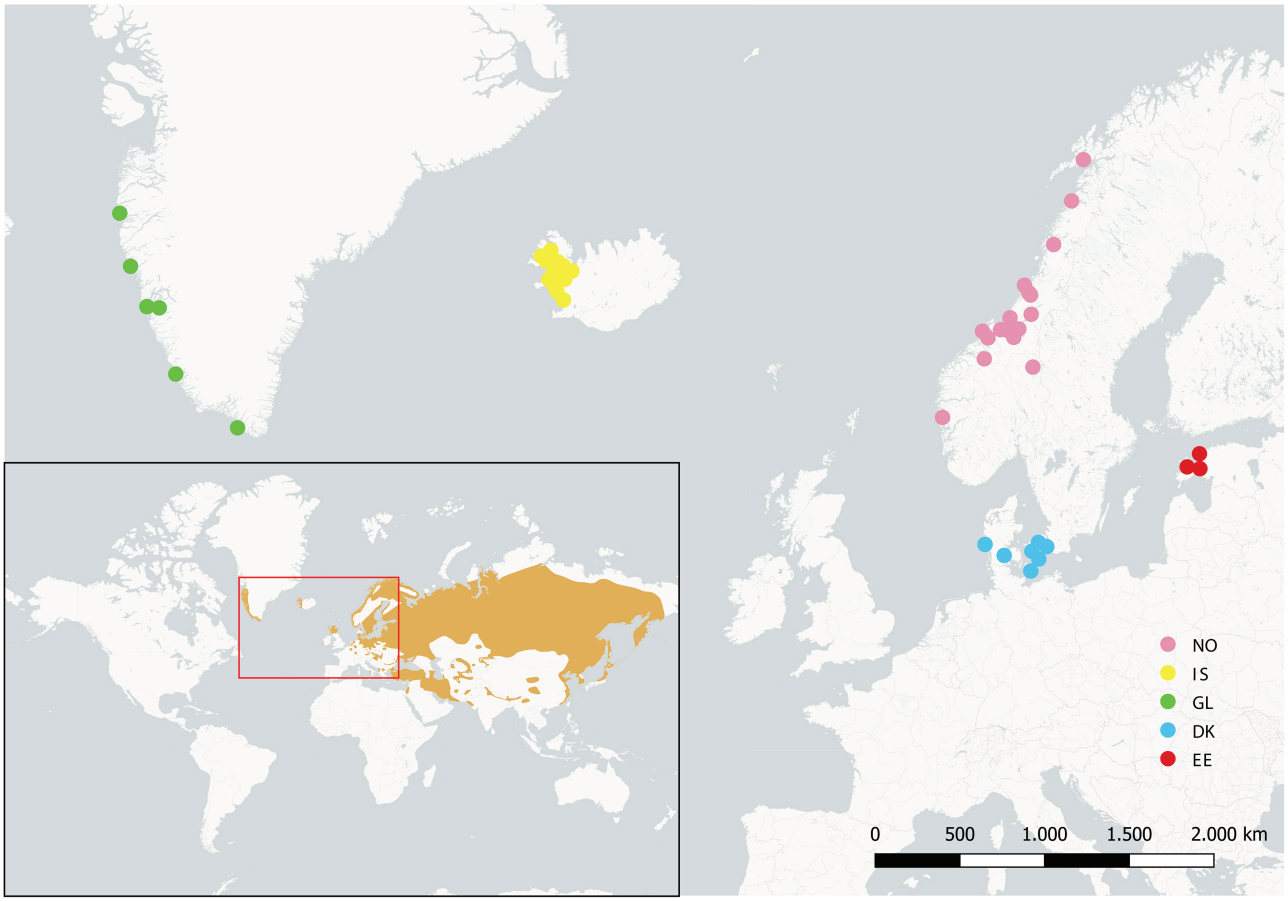

FIGURE 1. Map of distribution and sample origin of the White-tailed Eagles from 5 different countries analyzed here. Small map in black box displays the world and the White-tailed Eagle distribution in orange (without regards to where they breed and only use as passage etc.; Birdlife International 2020). The red box is the section of the world that makes up the larger map. The dots represent the known sample origin for the samples analyzed. Some dots represent more than one sample, and some dots are overlapping. Greenland (GL) is in green (n = 12), Iceland (IS) in yellow (n = 42), Norway (NO) in pink (n = 21), Denmark (DK) in blue (n = 11), Estonia (EE) in red (n = 3).

2020) but it was not included in subsequent analyzes unless specifically stated.

Data Analysis

Unless otherwise stated, R and Rstudio were used for all calculations (3.6.3 and 1.2.5033, respectively) (RStudio Team 2019, R core Team 2020). Due to the small sample size from Estonia, it was omitted from the statistical test described below assessing differences in variation within countries. Molecular diversity per country was summarized based on number of haplotypes, segregating sites, haplotype diversity (Nei and Tajima 1981), nucleotide diversity (π) (Nei 1987) using the *pegas* package (Paradis 2010), and haplotype richness (to account for difference in sample sizes) using the *hierfstat* package (Goudet 2005) that uses the rarefaction function, with standard settings and the smallest sample size of 11 (Paradis 2010). Differences in genetic diversity among the countries were summarized with 3 estimators of the population parameter Theta (θ = Neu), which may vary in sensitivity to detect population bottlenecks and selection, using *pegas* (Paradis 2010). The estimators are Theta Pi (θ_π), based on nucleotide diversity (Tajima 1996); Theta K (θ_K), using the expected number of alleles (Ewens 1972); and Theta S (θ_S), based on the number of segregating sites (Watterson 1975, Tajima 1989).

To summarize the relations between the sequences and their geographic origin, a median-joining network was drawn for all individuals using *popART* with standard settings (Bandelt et al. 1999, Leigh and Bryant 2015). A maximum likelihood (ML) tree was calculated using PhyML 3.3.3 (Guindon and Gascuel 2003), including Red Kite (*Milvus milvus*) as outgroup using the GTR+I+G model, based on the likelihood estimates from jModelTest 2.1.10 (Darriba et al. 2012). As the ML approach was not able to completely resolve the topology, a Bayesian approach was conducted using BEAST 2.6.2 (Bouckaert et al. 2014). The dataset was partitioned into 17 sets (1 for each

gene and rRNA, 1 containing all tRNAs, and 1 containing noncoding sequences) and bModelTest 1.2.1 (Bouckaert and Drummond 2017) to average over substitution models applying the default transition-transversion split option, and uncorrelated lognormal relaxed clock and constant populations size. However, this analysis did not converge for several partitions mainly due to the tree likelihood. To investigate this, we reduced the complexity of our analysis and ran BEAST again, this time on the unpartitioned dataset, applying a strict clock and a HKY substitution model with estimated frequencies and constant population size. We ran the analysis 10 times, each time for 20 million generations, and visualized the output using FigTree 1.4.2 (http://tree.bio.ed.ac.uk/software/figtree/). Convergence was visually inspected using Tracer 1.7.1 (Rambaut et al. 2018) with all effective sample size (ESS) values >200. This resulted in the recovery of six possible topologies (Supplementary Material Figure S3).

Estimation of divergence times were again conducted in BEAST 2.6.2. All analyzes were run using the Bayesian skyline model (Drummond et al. 2005), bModelTest 1.2.1 (Bouckaert and Drummond 2017) applying the default transition-transversion split option, and a strict clock in accordance with the low diversity of our intraspecies data set. The most frequent tree from our initial Bayesian analysis (Supplementary Material Figure S3A), described above, was used as a starting tree, fixing the topology throughout the runs. The MCMC chain was run each time for 20 million generations and the first 10% of each run were discarded as burn-in. Again, convergence was visually inspected in Tracer 1.7.1 (Rambaut et al. 2018) with all ESS values >200.

Because no fossil calibrations are available and substitution rates in birds are subject to discussion, 4 different approaches (giving 6 scenarios) were used to get a set of differently calibrated analyzes (also see Supplementary Material Table S2). First (scenario 1), we used the widely adopted standard rate of 2.1% divergence in birds for cytochrome *b* (cyt *b*) (Shields and Wilson 1987, Paxinos et al. 2002, Weir and Schluter 2008). The rate was applied for cyt *b* and for the rest of the mtDNA it was estimated for 17 partitions: 1 for each gene and 1 for each rRNA, as well as 1 each for all tRNAs and one for all noncoding regions. The rate for each partition was inferred by its relative nucleotide diversity, calculated in *pegas* (Paradis 2010) in comparison to cyt *b*. Secondly, we used body mass related rate (Nabholz et al. 2016), assuming a body mass of 5 kg (Helander et al. 2007, Isomursu et al. 2018). The analysis was run using third codon positions only and fossil calibration sets 2 (scenario 2; see Supplementary Material Table S2 for an overview over all 6 scenarios) and 4 (scenario 3), respectively (Nabholz et al. 2016). Nabholz et al. (2016) use 4 different sets of fossil calibrations to derive a body mass related substitution rate for birds. They forgo 2 calibration sets (1 and 3) since known divergence dates within the avian phylogeny could not be estimated correctly when using these calibration sets. Calibrations sets 2 and 4 both contain the following fossil calibrations: split between Neognathae and Paleognathae (86.5–66 mya); Anseriformes–Galliformes (free [no upper boundary] to 66 mya); Sphenisciformes–Procellariiformes (free to 60.5 mya); Coraciidae–Alcedinidae (free to 51.57 mya); Apodidiae–Trochilidae (free to 51 mya) and Psittaciformes–Passeriformes (65.5–53.5 mya). In addition to that, calibration set 4 contains the split between Oscines and Suboscines (34–28 mya). Although body mass seems to be a more accurate predictor for substitution rates when using third codon positions (Nabholz et al. 2016), the analysis was additionally run with all coding positions included, again using fossil calibration sets 2 (Scenario 4) and 4 (hardly different from scenario 4, so not added as different scenario), respectively. Third, we used rates calculated for Accipitriformes from Arcones et al. (2019) as substitution rates for the different mitochondrial genes (scenario 5). Finally (scenarios 6), we inferred rates using the divergence time between Common Buzzard (*Buteo buteo*) and White-tailed Eagle of 12.25 mya (Mindell et al. 2018). The Common Buzzard was used as it is the closest available relative with an associated divergence date to the White-tailed Eagle. The mean pairwise Tamura-Nei distances were calculated for each of the 17 partitions, between the sequences and Common Buzzard using MEGA X (Kumar et al. 2018).

Partitions of the molecular variance among and within populations (AMOVA) were summarized based on haplotype frequencies (F_{ST}) and evolutionary distances between sequences (Φ_{ST}), as well as between any pairs of samples (Excoffier et al. 1992, Tamura and Nei 1993) using the *hierfstat* (Goudet 2005) and *pegas* (Paradis 2010) packages in R. The significance of the F_{ST} and Φ_{ST} values were estimated by permutation of sequences among samples 1,000 times. A Kruskal's nonmetric multidimensional scaling (MDS) plot was calculated based on the pairwise Φ_{ST} distances, using the *MASS* package (Ripley 1996, Cox and Cox 2008).

Signs of selection on the mitochondrial variation were investigated by calculating the NI (Rand and Kann 1996) and with comparisons based on different estimates of the population parameter θ (see below). NI was calculated using the standard and generalized MacDonald and Kreitman test (http://mkt.uab.es) (Egea et al. 2008) using the vertebrate mitochondrial code for all 13 mitochondrial protein coding genes concatenated (with ND6 reversed), as well as for each gene for our complete data set and separately for the different countries. The ratio of segregating synonymous and non-synonymous variants within the populations was compared with the number of corresponding differentiations from Black Kite (*Milvus migrans*, NC_038195) as closest related species with an

TABLE 1. Molecular diversity per country and overall for White-tailed Eagle sampled between 1990 and 2018. Sample size (*n*), number of haplotypes (*nh*), haplotype richness (*HR*), haplotype diversity (*h*), nucleotide diversity (π), genomic diversity Theta Pi (θ_π) = nucleotide diversity (π) *17,461 bp (17,461 is the length of the analyzed mtDNA; Tajima 1996), theta K (θ_K) (Ewens 1972), theta S (θ_S) (Tajima 1989, Watterson 1975), census size as reported and referenced in Introduction (overall just report the estimate for Europe), numbers in parentheses are standard deviations. Standard deviation for θ_K were obtained with bootstrap. The individual from Greenland displaying signs of being a chimera is excluded here

Country	*n*	*nh*	*HR*	*h*	π	θ_π	θ_K	θ_S	Census
Greenland	11	8	5.68	0.93 (0.05)	3.61e-4 (2.0e-4)	6.31 (3.66)	11.69 (1.99)	6.83 (3.01)	150–200
Iceland	42	14	5.29	0.73 (0.06)	4.82e-4 (2.5e-4)	8.54 (4.47)	6.94 (1.04)	7.43 (2.45)	80
Norway	21	13	8.76	0.95 (0.01)	5.79e-4 (3.0e-4)	10.1 (5.37)	13.57 (1.98)	8.33 (3.11)	2,000
Denmark	11	7	7.00	0.87 (0.08)	1.83e-3 (9.7e-4)	31.97 (17.07)	6.15 (1.58)	24.58 (9.97)	61
Estonia	3	3	3.00	1.00 (0.19)	2.6e-4 (2.2e-4)	4.67 (3.90)	NA	4.66 (3.12)	290–330
Overall	88	43	8.4	0.92 (0.02)	1.08e-3 (5.3e-4)	18.97 (9.39)	32.57 (2.21)	24.5 (6.39)	17,900–24,500

available complete mitochondrial sequence. Results were obtained for both uncorrected NI as well as corrected for distance by Jukes and Cantor (Jukes and Cantor 1969) as implemented on the website. When at least one of the cells of the McDonald-Kreitman 2 ×2 contingency table contained a count of zero, a count of 1 was added to each cell. Differences in the estimates of the population parameter theta were tested with Tajima's *D* (Tajima 1989), Fay and Wu's *H* (Fay and Wu 2000), and the E-test (Zeng et al. 2006) which may also reveal changes in population sizes, were calculated using Zeng's DH-software (http://zeng-lab.group.shef.ac.uk) and tested by coalescence simulations using the ms-program assuming constant population sizes (Hudson 2002).

RESULTS

A total of 124 variable sites defining 43 mitogenomic haplotypes were found in the 88 individuals from Iceland, Norway, Denmark, Estonia, and Greenland (1 individual from Greenland was excluded from statistical analysis, see below). 48,495,446 reads were mapped to the individuals with a length of 257 bases each, after filtering 45,303,601 were left (93,4%). Only 5 individuals had missing data, with a maximum of 49 missing base pairs (bp) (i.e. the breadth (coverage) was minimum 0.997, and 84 individuals had a breadth of 1). The mean depth ranged from 12.7 to 7894.84 (mean = 857.3). The control region was separately analyzed and had a mean depth between 9.9 and 6774,34 (mean = 800), and thus slightly but not substantially lower. Overall quality, depth, ratio of the common allele, and binomial probability of segregating sites in each individual within each country, summarized from the filtered VCF file, show a good support for the segregating sites and that they behave as haploid markers (Supplementary Material Figure S1), with no evidence of sample contamination or index hopping. Inspection of highly variable sites within countries was no exception, as found for 11 sites separating the 2 main mtDNA lineages in Iceland (Supplementary Material Table S3). Nucleotide diversity along the mitochondrial genome was well distributed (Supplementary Material Table S4). The small population in Iceland has the lowest haplotype richness and displays lowest haplotype diversity (0.7), but the population from Greenland show the lowest nucleotide diversity (3.61×10^{-4}) (Table 1). The large population in Norway has the highest haplotype richness and diversity, whereas the recently established population in Denmark has the highest nucleotide diversity. The population parameters θ_S based on segregating sites and nucleotide diversity are also higher in Denmark than for the other countries. θ_K, which unlike nucleotide diversity is sensitive to changes in number of alleles and thus bottlenecks, was lowest in Iceland and the newly established population in Denmark.

Both the median-joining network (Figure 2) and the Bayesian tree (Figure 3) revealed 2 major haplogroups in our data, one with all Estonian and about half the Danish individuals (HG-B), the other with Norway, Iceland, Greenland, and the rest of the Danish individuals (HG-A). The haplotype detected in the 3 Danish individuals clustering in HG-A were identical to the haplotype from Germany (MN356434.1). The Korean reference specimen forms a third and a distinct branch (Figure 2) that shows greater similarity to the branch with Estonian and half of the Danish individuals (HG-B) than to the others (HG-A) (Supplementary Material Figure S2). One sequence from Greenland (GL-3) clustered far from the rest of the sequences from Greenland in the network; but in the Bayesian analyzes, this sequence clustered within the group IG-1. An inspection of the sequence showed signs of it being a chimera, as 17 of 25 variable sites in the sequence covering most of the sequence—that is from position 1 to 14597 were shared by other individuals from Greenland (in group HG-A), whereas 8 variable sites from the control region (CR) were shared with the distinct lineage (HG-B) observed in individuals from Denmark, Estonia, and Korea. The individual was marked with an asterisk (Figures 2 and 3), but removed from statistical analysis, except for the neutrality index that was based on the coding sequence. Six topologies were found with the Bayesian

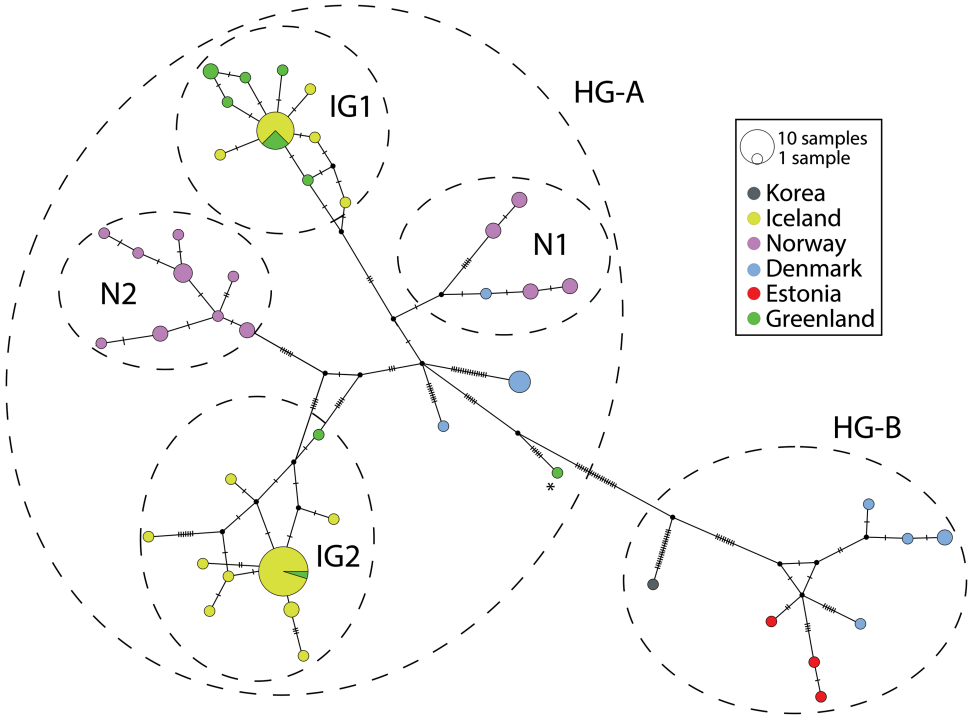

FIGURE 2. Median-joining network presenting the divergence between mitochondrial genomes in White-tailed Eagles from the five different countries sampled between 1990 and 2018 compared to the Korean reference specimen. The size of the pies represents the number of individuals that share the same haplotype. Marks on the lines refer to number of sites that differ among the connected haplotypes. HG-A and HG-B corresponds to the haplogroups A and B from Hailer et al (2007) (Supplementary Material Figure S4). IG1, IG2, N1, and N2 are Iceland-Greenland group 1, Iceland-Greenland group 2, Norway group 1, and Norway group 2, respectively. The individual from Greenland displaying signs of being a chimera is marked with an asterisk.

analysis; they all show identical grouping of individuals but differ in the sequence of splits that were poorly supported (Supplementary Material Figure S3). The most frequent topology (Figure 3) was observed 4 times. One topology was observed twice and 4 topologies only once. Two distinct clusters were found within each country in HG-A. These clusters were separated by several mutations (Figures 2 and 3) and contain highly similar haplotypes (N1 and IG1 in one, and N2 and IG2 in the other). Iceland and Greenland share closely related or similar haplotypes that differ in frequency at the different clusters (IG1 and IG2). The topology indicates no batch effect even though the data was obtained from various sources and sequenced at two different facilities. A network based on the part of the control region analyzed by Hailer et al. (Hailer et al. 2007), and for the whole region excluding the control region, both present the same overall pattern (except for the individual from Greenland displaying signs of being a chimera), although the CR-fragment does not provide as high a resolution as the entire mitogenome (Supplementary Material Figures S4 and S5 and Table S1).

The earliest divergence of haplogroup HG-B from haplogroup HG-A (Figure 3) occurred on average 130 kya. A second major split distinguishes the Danish individuals (except one) from the Norwegian, Icelandic, and Greenlandic eagles (57 kya). The third major split divides one-third of the Norwegian individuals, one-third of the Icelandic individuals, and all but 2 Greenlandic individuals into one clade (N1 and IG1), and the remaining in the other clade (N2 and IG2) (51kya). In both of these two last clades, the Norwegian individuals make up separate monophyletic clades (with 1 Danish individual), splitting from the Icelandic and Greenlandic individuals around 30 and 37 kya. All these referred times are, as in Figure 3, assuming standard mutation rate.

Split times for all the 6 methods used (3 methods were used for body mass), and for the 3 major splits mentioned above are displayed in Supplementary Material Figures S6–S8. The most recent splits obtained with the different methods range from around 336 kya (237–437 kyr) to 51 kya (36–68 kyr). For the root node, the mean split times range from 847 kyra (621 kyr to 1.086 Myr) to 130 kya (94–169 kyr) for the oldest and youngest, respectively.

The AMOVA (Excoffier et al. 1992) analysis reveals that a large proportion of the variation in genetic distances is due to differences between populations from the different countries ($\Phi_{ST} = 0.56$, $P < 0.01$). This proportion was considerably larger than the one obtained with the partition based on haplotypes only ($F_{ST} = 0.079$, $P < 0.01$) and reflects the unique lineages found in most samples (Figure 2).

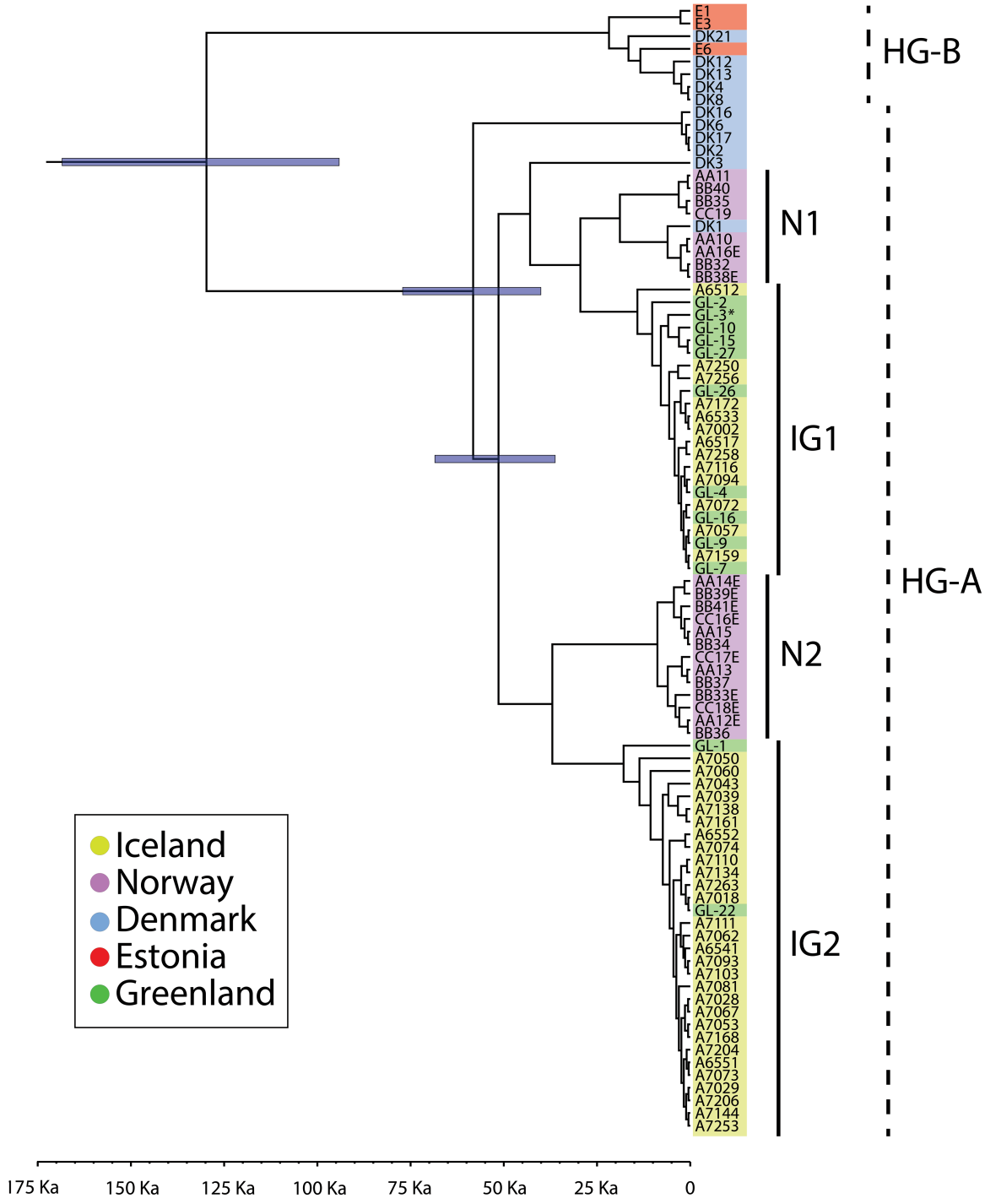

FIGURE 3. Phylogenetic tree based on the mitochondrial genomes in White-tailed Eagles from 5 countries collected between 1990 and 2018. The tree was reconstructed using BEAST. The dates (Ka) are shown on the *x*-axis from today and back in time. The dates are based on the standard rate (2.1%), and error bars for the 3 major splits are marked in blue. HG-A and HG-B corresponds to the haplogroups A and B from Hailer et al (2007) (Figure 2 and Supplementary Figure S4). IG1, IG2, N1, and N2 are Iceland-Greenland group 1, Iceland-Greenland group 2, Norway group 1 and Norway group 2, respectively. The individual from Greenland displaying signs of being a chimera is marked with asterisk.

Accordingly, pairwise Φ_{ST} were larger than F_{ST} values (Table 2) and showed significant differences between all countries, excluding the small sample of the Estonian population. The positions based on the pairwise Φ_{ST} distances in the MDS-plot reflect well the original genetic distances (with a stress of 2.5%) and show relation

TABLE 2. Pairwise comparison among 5 populations of White-tailed Eagle sampled between 1990 and 2018. Φ_{ST} below the diagonal, and pairwise F_{ST} above the diagonal. Asterisks *, ** and *** indicates *P*-values: $0.01 \leq P < 0.05$, $0.001 \leq P < 0.01$, and $P < 0.001$, respectively. IS: Iceland, NO: Norway, DK: Denmark, EE: Estonia, and GL: Greenland

	IS	NO	DK	EE	GL
Iceland		0.083***	0.109***	0.085	0.041*
Norway	0.408***		0.043***	0.012	0.026**
Denmark	0.560***	0.453***		0.038	0.049*
Estonia	0.940***	0.935***	0.273		0.018
Greenland	0.397***	0.459***	0.308*	0.961**	

to geographic maps (Supplementary Material Figure S9), although the distance between Greenland and Iceland is large despite close similarity of the sequences from the two clades IG1 and IG2 in Figure 2. The overall distance between the 2 samples can be explained by the different proportions of Icelandic and Greenland haplotypes in the 2 clusters (IG1 has 13 Icelandic and 10 Greenlandic individuals, IG2 has 29 Icelandic and 2 Greenlandic individuals).

We found an excess of segregating amino acids in comparison with the fixed differences between the White-tailed Eagle and the Black Kite, indicating negative selection (NI > 1) (Table 3), with no stop codons detected. This excess of non-synonymous polymorphisms is especially driven by high NI values for the COX complex (COX 1 and 2) and for cyt *b* when all individuals are pooled. High significant NI values in the COX complex for all populations are found, primarily in COX2 in Iceland, Greenland, and Norway, but COX1 in DK. Iceland and Greenland show significant negative selection in ND1, Norway and Denmark display significant negative selection in cyt *b* (Table 3). The Fu and Way's *H* test statistic and especially Zeng's *E*-test support this (Table 4); there are high frequency variants segregating within each population which are located at the internal branches within the two distinct lineages in each country, each with closely related haplotypes (i.e. the tree has a greater height than expected based on its overall length; Figure 3). An evaluation of the impact of sample size in Iceland revealed similar results (Supplementary Material Table S5). None of the Tajima's *D* values were significant, but interestingly, the test statistic is positive for all countries except Greenland, indicating there are less rare alleles than expected (Table 4).

DISCUSSION

Population Structure

The mtDNA genomes of White-tailed Eagles sampled from Greenland, Iceland, Norway, Denmark, and Estonia show difference in haplotype frequencies, occurrence of different haplotypes in west and east Europe, and thus differentiation between the countries (Φ_{ST}), which are all in line with earlier findings on the mitochondrial DNA in White-tailed Eagles based on a fragment of the control region (Hailer et al. 2007, Honnen et al. 2010). The study by Hailer et al. (2007) revealed a recent genetic differentiation between the populations from Iceland and Greenland, where the frequency of 2 haplotypes differed between the 2 islands. One of the haplotypes was endemic to the islands and the other common in northern Europe. Even though Hailer et al. (2007) observed large spatial overlap of CR haplotypes on the mainland, when constructing a neighbor-joining tree they detected a split between western and eastern Europe. Here we also observe the recent divergence of the two island populations and a split between western Europe and the samples from Estonia in eastern Europe, yet the recently established population from Denmark shows an admixture of the 2 groups. The sample from Denmark shares 1 haplotype with the Norwegian sample, suggesting migration from Norway; another haplotype is identical to 1 from Germany and 5 others share a similar haplotype to the Estonian samples, suggesting a settlement from the east. However, further sampling, and comparison with the surrounding countries is needed to confirm geneflow, and the full population structure.

Divergence of Populations

Despite the similarities between the main geographic patterns in our study and the previous studies, a more complex pattern is observed based on the entire mitochondrial genomes than the control region. Within each country, except in the small sample set from Estonia, 2 distinct mitogenomic lineages are observed and are being upheld, potentially through selection. The sequences within the two separate lineages in Iceland and Greenland are similar to each other but have diverged from the Norwegian haplotypes in the same clade for about 30,000–37,000 yr ago, as estimated by the standard mutation rate for cyt *b* or around the last glacial maximum, ~25,000 yr ago, when the three countries were covered by glaciers (Clark and Mix 2002). It is interesting that the split happened during or before the last glacial maximum (i.e. prior to the colonization of the 3 countries). The time to the most recent common ancestor (TMRCA) for the 2 distinct lineages within Greenland, Iceland, and Norway was older or about ~50,000 yr (CI: 115,000–11,700 yr ago) or during the last glaciation, suggesting that they originated from two refugia during the last glacial period. The dating to the TMRCA for all countries occurred close to onset of the last glaciation assuming the standard rate (Supplementary Material Figure S6), 130,000 yr (mean standard rate).

The other methods of timing based on more ancient divergence times between species—that is, on body size (Nabholz et al. 2016) and the one from Arcones et al.

TABLE 3. Neutrality index (NI) (Rand and Kann 1996). Bold = significant, italic = to all values in the 2 ×2 contingency table a count of 1 was added, because at least 1 cell contained a 0 and a calculation of the neutrality index would not have been possible otherwise. All: all samples, JC: JC corrected. IS: Iceland, GL: Greenland, NO: Norway, DK: Denmark. Bases per gene; ATP6: 683; ATP8: 167; COX1: 1,550; COX2: 683; COX3: 783; CYTB: 1,142; ND1: 977; ND2: 1,038; ND3: 351; ND4: 1,377; ND4L: 296; ND5: 1,817; ND6: 518

	All	All–JC	IS	IS–JC	GL	GL–JC	NO	NO–JC	DK	DK–JC
All genes	**2.3**	**3.0**	1.8	2.4	2.2	2.9	2.3	**3.1**	1.9	**2.5**
ATP6	*1.1*	*1.3*	*4.3*	*5.3*	*4.3*	*5.3*	*4.3*	*5.3*	*1.1*	*1.3*
ATP8	*0.5*	*0.5*	*0.9*	*1.1*	*0.9*	*1.1*	*0.5*	*0.5*	*0.9*	*1.1*
COX1	**8.4**	**11.0**	*7.1*	***9.3***	*7.1*	***9.3***	*7.1*	***9.3***	**12.6**	**16.4**
COX2	**45.6**	**60.8**	***29.0***	***38.6***	***29.0***	***38.6***	**22.8**	**30.4**	*4.8*	*6.4*
COX3	*3.8*	*4.8*	*7.6*	***9.7***	*7.6*	***9.7***	*5.0*	*6.4*	*7.8*	***10.0***
CYTB	**14.8**	**22.2**	*3.4*	*4.6*	*3.4*	*4.6*	**15.4**	**21.0**	**21.8**	**29.7**
ND1	4.0	5.4	***10.9***	***14.8***	***10.9***	***14.8***	*5.4*	*7.4*	6.0	8.1
ND2	0.8	1.1	*3.2*	*4.2*	*3.2*	*4.2*	*3.2*	*4.2*	0.8	1.1
ND3	1.2	1.7	*1.9*	*2.4*	*1.9*	*2.4*	*1.9*	*2.4*	3.7	5.2
ND4	3.2	4.3	*6.4*	*8.3*	*6.4*	*8.3*	***12.6***	***16.4***	*2.1*	*2.8*
ND4L	*0.4*	*0.6*	*3.5*	*4.6*	*3.5*	*4.6*	*3.5*	*4.6*	*0.4*	*0.5*
ND5	1.7	2.3	1.1	1.6	1.7	2.3	0.7	0.9	2.8	**3.7**
ND6	1.0	1.3	*3.8*	*5.3*	*3.8*	*5.3*	*3.8*	*5.3*	1.0	1.3
Complex I (ND)	1.4	1.9	1.3	1.7	1.8	2.4	1.3	1.7	1.4	1.8
Complex IV (COX)	**11.2**	**14.6**	**12.5**	**16.3**	**12.5**	**16.3**	**7.5**	**9.8**	**6.3**	**8.2**
Complex V (ATPase)	*0.5*	*0.6*	*2.7*	*3.3*	*2.7*	*3.3*	*1.3*	*1.6*	*0.7*	*0.8*

TABLE 4. Tests of selection for White-tailed Eagle samples between 1990 and 2018. Tajima's *D*, the normalized Fay and Wu's *H*, and the *E*-test. 10.000 coalescence simulations with the number of segregating sites fixed. *P*-values as proportions of number of simulations resulting in equal or lower values than the observed statistic

Region	Tajima's *D*	Fay and Wu's *H*	*E*-test
Greenland	–0.46 (0.343)	–1.99 (0.035)	1.68 (0.968)
Iceland	0.92 (0.863)	–2.66 (0.017)	3.22 (0.999)
Norway	0.85 (0.849)	–1.19 (0.091)	1.92 (0.979)
Denmark	1.47 (0.961)	–0.19 (0.268)	1.59 (0.966)

(2019) gave more ancient splits than based on the standard rate for cyt *b* (Weir and Schluter 2008). This difference could result from molecular clock rates obtained from comparisons between species, which may overestimate the divergence times within species as they do not consider the influence of population dynamics such as selection (Ho et al. 2005, 2007; Zink and Barrowclough 2008). The standard rate may be less affected, and our estimates come close to the earlier estimates (Hailer et al. 2007); however, the large divergence within populations, which seems to be maintained by selection, as discussed below, will also affect the time of divergence.

Different refugia during the last glacial period have been suggested as the main factor driving the formation of the western and eastern clades in the White-tailed Eagle (Hailer et al. 2007, Langguth et al. 2013), and the estimates here from the standard rate are close to these earlier suggested split times. The exact location of the refugia is unknown, but such refugia driving population splits are also seen in other birds with a large east/west range on the northern hemisphere (Ruokonen et al. 2004, Lovette 2005). More extensive geographic sampling could further refine theories of the location of the refugia and the colonization history of Greenland and northern Europe.

Selection

As the mitochondria contain haploid, non-recombining genomes, selection is predicted to reduce the molecular variation within populations due to background selection against deleterious mutations (Charlesworth et al. 1993) or due to hitchhiking with positively favored mutations (Smith and Haigh 1974). These effects are expected to be even stronger in birds due to linkage with the W-chromosome as they are inherited together (Berlin et al. 2007).

Here we observe a different outcome, high variation within populations with distinct lineages and an excess of high frequency variants as revealed by the negative Fay and Wu's *H* in Iceland and Greenland, which is in line with Zeng's *E*, displaying long internal branches (i.e. the deep split within the countries). The significantly low *H* and high *E* values could indicate balancing selecting for the two lineages, with IG1 and NO1 in one, and IG2 and NO2 in the other (Zeng et al. 2006). However, balancing selection is unlikely to be directed toward the mitochondrial genome, but could rather act on the W-chromosome, due to the linkage, and be driven by its interaction with the Z-chromosome (e.g., in the pseudoautosomal region [PAR] or non-recombining homologs; Xu and Zhou 2020) and then indirectly maintain different lineages in

the mitochondria. And as balancing selection will result in larger sequence variation within populations, our estimates of TMRCA within populations are overestimated as well as the time since the populations diverged from each other (Zink and Barrowclough 2008, Cooper et al. 2015). Further evidence for the effect of selection was seen with the NI > 1 (i.e. higher ratio of polymorphic amino acids within the White-tailed Eagles in comparison to the ratio of fixed amino acid replacements relative to a related species), which suggests that deleterious mutations are segregating within the mitochondrial genome of the eagles so the variation may not reflect neutral processes.

A sympatric occurrence of distinct genetic lineages from a deep gene tree as seen here had been suggested earlier for the White-tailed Eagle (Honnen et al. 2010); and although rarely seen, it has been suggested for other species (Adelie Penguin [*Pygoscelis adeliae*], Mallard [*Anas platyrhynchos*], Snow Goose [*Anser caerulescens*], Eurasian Blue Tit [*Cyanistes caeruleus*], and orcas [*Orcinus orca*] (Avise et al. 1987, Dizon et al. 1992, Avise and Walker 1998). Although this dichotomy may represent an admixture of different clades, 2 different colonization events, high ancestral variation or incomplete linage sorting, the maintenance of these distinct clades within the small populations of White-tailed Eagles in Greenland and Iceland, and in the larger population in Norway, indicate that natural selection has played a role. Nuclear inserts of mitochondrial DNA (NUMTs) could have led to such a pattern, but inspection of the variable sites separating the lineages within populations did not reveal any such evidence.

Conclusion

Here we confirm the suggested western and eastern clades of mitochondrial variation in White-tailed Eagle and the main phylogeographic pattern in Greenland and Northern Europe. Surprisingly, 2 distinct lineages are found within each country. These 2 lineages in Greenland, Iceland, and Norway seem to have been upheld by natural selection and may reflect divergence in different refugia during the last glacial period. In Denmark, the 2 lineages derive from the eastern and western clades and reflect a recent recolonization from both areas. Wider geographic sampling is warranted for a complete phylogeography of Europe and to rebuild the natural history of the species. Though useful for conservation efforts an addition of other genomic markers is needed to assess the diversity within populations, including assessments of demography, effective population size, and divergence. Furthermore, to explore the signal of selection an analysis of the sex-chromosomes Z and W could reveal whether balancing selection acting on Z and W can extend over to the mitogenome and explain the coexistence of the two lineages within countries.

SUPPLEMENTARY MATERIAL

Supplementary material is available at *Ornithology* online.

ACKNOWLEDGMENTS

We thank BirdLife Denmark (DOF BirdLife) for letting us use the samples they have gathered and also the Natural History Museum of Denmark for providing samples.
Funding statement: The study was supported by Research Grant nr 185280-052 from The Icelandic Research Council, the Doctoral student fund of the University of Iceland and The University of Iceland Research fund, and an NTNU Onsager Fellowship award to MDM. For the Norwegian samples; NovaSeq 6000 sequencing services were provided by the Norwegian Sequencing Centre, a national technology platform hosted by the University of Oslo and supported by the "Functional Genomics" and "Infrastructure" programs of the Research Council of Norway and the Southeastern Regional Health Authorities. Illumina HiSeq 4000 sequencing was performed by the NTNU Genomics Core Facility (GCF), which is funded by the Faculty of Medicine and Health Sciences, Norwegian University of Science and Technology (NTNU), and the Central Norway Regional Health Authority. Some analyzes were performed on resources provided by the National Infrastructure for High Performance Computing and Data Storage in Norway (UNINETT Sigma2).
Ethics statement: This research was conducted in compliance with the Guidelines to the Use of Wild Birds in Research.
Conflict of interest statement: The authors declare no competing interests.
Author contributions: Conceived the idea: CCRH and SP; Performed the experiments: CCRH, SB, SMG, JAR, JACB, MDM and SP; Analyzed the data: CCRH, SB, SMG and SP; wrote the paper: CCRH, SB, MDM and SP; Corrected the paper: SMG, JAR, JACB, GTH, RAS, MvS, KHS, ALL, ML, CS, RD, KS, DB and IE; Contributed substantial materials, resources or funding: CCRH, GTH, RAS, MvS, KHS, ALL, ML, CS, RD, KS, DB, IE, MDM and SP.
Data availability: The DNA sequences mapped to the Korean specimen is available at https://doi.org/10.5061/dryad.stqjq2c32.

LITERATURE CITED

Arcones, A., R. Ponti, and D. R. Vieites (2019). Mitochondrial substitution rates estimation for molecular clock analyzes in modern birds based on full mitochondrial genomes. bioRxiv 11:1–11.

Avise, J. C., and D. E. Walker (1998). Pleistocene phylogeographic effects on avian populations and the speciation process. Proceedings of the Royal Society of London, Series B: Biological Sciences 265:457–463.

Avise, J. C., J. Arnold, R. M. Ball, E. Bermingham, T. Lamb, J. E. Neigel, C. A. Reeb, and N. C. Saunders (1987). Intraspecific Phylogeography: The mitochondrial DNA bridge between population genetics and systematics. Annual Review of Ecology and Systematics 18:489–522.

Babraham Bioinformatics (2010). FastQC. http://www.bioinformatics.babraham.ac.uk/projects/fastqc/

Bandelt, H. J., P. Forster, and A. Rohl (1999). Median-joining networks for inferring intraspecific phylogenies. Molecular Biology and Evolution 16:37–48.

Berlin, S., D. Tomaras, and B. Charlesworth (2007). Low mitochondrial variability in birds may indicate Hill–Robertson effects on the W chromosome. Heredity 99:389–396.

BirdLife International (2016). Haliaeetus albicilla. The IUCN Red List of Threatened Species 2016 e.T2269513.

Bouckaert, R. R., and A. J. Drummond (2017). bModelTest: Bayesian phylogenetic site model averaging and model comparison. BMC Evolutionary Biology 17:1–11.

Bouckaert, R., J. Heled, D. Kühnert, T. Vaughan, C. H. Wu, D. Xie, M. A. Suchard, A. Rambaut, and A. J. Drummond (2014). BEAST 2: A software platform for Bayesian evolutionary analysis. PLoS Computational Biology 10:1–6.

Bragin, E. A., S. A. Poessel, M. J. Lanzone, and T. E. Katzner (2018). Post-fledging movements and habitat associations of White-tailed Sea Eagles (*Haliaeetus albicilla*) in Central Asia. The Wilson Journal of Ornithology 130:784–788.

Carøe, C., S. Gopalakrishnan, L. Vinner, S. S. T. Mak, M. H. S. Sinding, J. A. Samaniego, N. Wales, T. Sicheritz-Pontén, and M. T. P. Gilbert (2018). Single-tube library preparation for degraded DNA. Methods in Ecology and Evolution 9:410–419.

Charlesworth, B., M. T. Morgan, and D. Charlesworth (1993). The effect of deleterious mutations on neutral molecular variation. Genetics 134:1289–1303.

Clark, P. U., and A. C. Mix (2002). Ice sheets and sea level of the Last Glacial Maximum. Quaternary Science Reviews 21:1–7.

Cooper, B. S., C. R. Burrus, C. Ji, M. W. Hahn, and K. L. Montooth (2015). Similar efficacies of selection shape mitochondrial and nuclear genes in both *Drosophila melanogaster* and *Homo sapiens*. G3: Genes, Genomes, Genetics 5:2165–2176.

Cox, M. A. A., and T. F. Cox (2008). Multidimensional scaling. In Handbook of Data Visualization (C-H. Chen, W. Härdle, and A. Unwin, Editors). Springer Berlin Heidelberg, Berlin, Heidelberg, Germany. pp. 315–347.

Darriba, D., G. L. Taboada, R. Doallo, and D. Posada (2012). JModelTest 2: More models, new heuristics and parallel computing. Nature Methods 9:772.

Dizon, A. E., C. Lockyer, W. F. Perrin, D. P. Demaster, and J. Sisson (1992). Rethinking the stock concept: A phylogeographic approach. Conservation Biology 6:24–36.

Drummond, A. J., A. Rambaut, B. Shapiro, and O. G. Pybus (2005). Bayesian coalescent inference of past population dynamics from molecular sequences. Molecular Biology and Evolution 22:1185–1192.

Egea, R., S. Casillas, and A. Barbadilla (2008). Standard and generalized McDonald-Kreitman test: a website to detect selection by comparing different classes of DNA sites. Nucleic Acids Research 36:W157–W162.

Elts, J., A. Leito, M. Leivits, L. Luigujõe, R. Nellis, M. Ots, I. Tammekänd, and Ü. Väli (2019). Eesti lindude staatus, pesitsusaegne ja talvine arvukus 2013-2017. Hirunda 32:1–29.

Ewens, W. J. (1972). The sampling theory of selectively neutral alleles. Theoretical Population Biology 3:87–112.

Excoffier, L., P. E. Smouse, and J. M. Quattro (1992). Analysis of molecular variance inferred from metric distances among DNA haplotypes: Application to human mitochondrial DNA restriction data. Genetics 131:479–491.

Fay, J. C., and C. I. Wu (2000). Hitchhiking under positive Darwinian selection. Genetics 155:1405–1413.

Feng, S., J. Stiller, Y. Deng, J. Armstrong, Q. Fang, A. H. Reeve, D. Xie, G. Chen, C. Guo, B. C. Faircloth, B. Petersen, et al. (2020). Dense sampling of bird diversity increases power of comparative genomics. Nature 587:252–257.

Goudet, J. (2005). HIERFSTAT, a package for R to compute and test hierarchical F -statistics. Molecular Ecology Notes 5:184–186.

Greenland Institute of Natural Resources (2020). White-tailed Eagle. https://natur.gl/arter/white-tailed-eagle/?lang = en

Guindon, S., and O. Gascuel (2003). A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Systematic Biology 52:696–704.

Hailer, F., B. Helander, A. O. Folkestad, S. A. Ganusevich, S. Garstad, P. Hauff, C. Koren, V. B. Masterov, T. Nygård, J. A. Rudnick, S. Shiraki, et al. (2007). Phylogeography of the White-tailed Eagle, a generalist with large dispersal capacity. Journal of Biogeography 34:1193–1206.

Hansen, K. (1979). Status over bestanden af Havørn *Haliaeetus albicilla* groenlandicus Brehm i Grønland i årene 1972–74. Dansk Ornitologisk Forenings Tidsskrift 73:107–130.

Helander, B., F. Hailer, and C. Vilà (2007). Morphological and genetic sex identification of White-tailed Eagle *Haliaeetus albicilla* nestlings. Journal of Ornithology 148:435–442.

Ho, S. Y. W., B. Shapiro, M. J. Phillips, A. Cooper, and A. J. Drummond (2007). Evidence for time dependency of molecular rate estimates. Systematic Biology 56:515–522.

Ho, S. Y. W., M. J. Phillips, A. Cooper, and A. J. Drummond (2005). Time dependency of molecular rate estimates and systematic overestimation of recent divergence times. Molecular Biology and Evolution 22:1561–1568.

Honnen, A. C. C., F. Hailer, N. Kenntner, I. Literák, L. Dubská, and F. E. Zachos (2010). Mitochondrial DNA and nuclear microsatellites reveal high diversity and genetic structure in an avian top predator, the white-tailed sea eagle, in central Europe. Biological Journal of the Linnean Society 99:727–737.

Hudson, R. R. (2002). Generating samples under a Wright-Fisher neutral model of genetic variation. Bioinformatics 18:337–338.

Isomursu, M., J. Koivusaari, T. Stjernberg, V. Hirvelä-Koski, and E. R. Venäläinen (2018). Lead poisoning and other human-related factors cause significant mortality in White-tailed Eagles. Ambio 47:858–868.

Jais, M. (2020). White-tailed Eagle. http://europeanraptors.org/white-tailed-eagle/

Jukes, T. H., and C. R. Cantor (1969). Evolution of protein molecules. In Mammalian Protein Metabolism (H. N. Munro, Editor). Elsevier, New York, NY, USA. pp. 21–132.

Kim, J. A., S. G. Kang, H. S. Jeon, J. H. Jeon, J. H. Jang, S. Kim, and J. An (2019). Complete mitogenomes of two Accipitridae, *Haliaeetus albicilla*, and *Pernis ptilorhynchus*. Mitochondrial DNA Part B: Resources 4:392–393.

Kircher, M., S. Sawyer, and M. Meyer (2012). Double indexing overcomes inaccuracies in multiplex sequencing on the Illumina platform. Nucleic Acids Research 40:e3–e3.

Kongeaastien (2020). The White-tailed Eagle. http://kongeaastien.dk/en/content/white-tailed-eagle

Kumar, S., G. Stecher, M. Li, C. Knyaz, and K. Tamura (2018). MEGA X: Molecular evolutionary genetics analysis across computing platforms. Molecular Biology and Evolution 35:1547–1549.

Langguth, T., A. C. C. Honnen, F. Hailer, T. Mizera, S. Skoric, Ü. Väli, and F. E. Zachos (2013). Genetic structure and phylogeography of a European flagship species, the white-tailed sea eagle *Haliaeetus albicilla*. Journal of Avian Biology 44:263–271.

Leigh, J. W., and D. Bryant (2015). POPART: full-feature software for haplotype network construction. Methods in Ecology and Evolution 6:1110–1116.

Li, H., and R. Durbin (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25:1754–1760.

Li, H., B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis, and R. Durbin (2009). The sequence alignment/map format and SAMtools. Bioinformatics 25:2078–2079.

Lõhmus, A. (1998). Eesti kotkaste argipäev. Eesti Loodus 5/6:210–213.

Lovette, I. J. (2005). Glacial cycles and the tempo of avian speciation. Trends in Ecology & Evolution 20:57–59.

Lynch, M. (1996). Mutation accumulation in transfer RNAs: Molecular evidence for Muller's ratchet in mitochondrial genomes. Molecular Biology and Evolution 13:209–220.

Macherey-Nagel (2014). Genomic DNA from Tissue. User manual, NucleoMag 96 Tissue, Takara Bio USA, Inc. San Jose, CA, USA. https://www.takarabio.com/assets/documents/User Manual/NucleoMag 96 Tissue Genomic DNAPurification User Manual(PT5008-1)_Rev_03.pdf

Mindell, D. P., J. J. Fuchs, and J. A. Johnson (2018). Phylogeny, taxonomy, and geographic diversity of diurnal raptors: Falconiformes, Accipitriformes, and Cathartiformes. In Birds of Prey: Biology and conservation in the XXI century (J. H. Sarasola, J. M. Grande and J. J. Negro, Editors). Springer International Publishing, Cham, Denmark. pp. 1–522.

Nabholz, B., R. Lanfear, and J. Fuchs (2016). Body mass-corrected molecular rate for bird mitochondrial DNA. Molecular ecology 25:4438–4449.

Nei, M. (1987). Molecular Evolutionary Genetics. Columbia University Press, New York, NY, USA.

Nei, M., and F. Tajima (1981). DNA polymorphism detectable by restriction endonucleases. Genetics 97:145–163.

Neiman, M., and D. R. Taylor (2009). The causes of mutation accumulation in mitochondrial genomes. Proceedings of the Royal Society B: Biological Sciences 276:1201–1209.

Paradis, E. (2010). PEGAS: An R package for population genetics with an integrated-modular approach. Bioinformatics 26:419–420.

Paxinos, E. E., H. F. James, S. L. Olson, M. D. Sorenson, J. Jackson, and R. C. Fleischer (2002). mtDNA from fossils reveals a radiation of Hawaiian Geese recently derived from the Canada Goose (*Branta canadensis*). Proceedings of the National Academy of Sciences USA 99:1399–1404.

Petersen, Æ. (1998). Haförn. In Íslenskir Fuglar (S. Harðardóttir, Editor). Vaka-Helgafell hf., Reykjavik, Iceland. pp. 120–123.

Qiagen (2006). DNeasy Blood & Tissue Handbook. http://www.qiagen.com/products/genomicdnastabilizationpurification/dneasytissuesystem/dneasybloodtissuekit.aspx#Tabs = t0

R Core Team (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.r-project.org/

Rambaut, A., A. J. Drummond, D. Xie, G. Baele, and M. A. Suchard (2018). Posterior summarization in Bayesian phylogenetics using Tracer 1.7. Systematic Biology 67:901–904.

Rand, D. M., and L. M. Kann (1996). Excess amino acid polymorphism in mitochondrial DNA: Contrasts among genes from Drosophila, mice, and humans. Molecular Biology and Evolution 13:735–748.

Randla, T., and A. Õun (1980). Kaljukotkas ja merikotkas Eestis 1970-ndail aastail. Eesti Loodus 4:512–515.

Ripley, B. D. (1996). Pattern Recognition and Neural Networks. Cambridge University Press, Cambridge, UK.

Rohland, N., and D. Reich (2012). Cost-effective, high-throughput DNA sequencing libraries for multiplexed target capture. Genome Research 22:939–946.

RStudio Team (2019). RStudio: Integrated Development for R. http://www.rstudio.com/

Ruokonen, M., L. Kvist, T. Aarvak, J. Markkola, V. v. Morozov, I. J. Øien, E. E. Syroechkovsky, P. Tolvanen, and J. Lumme (2004). Population genetic structure and conservation of the lesser white-fronted goose Anser erythropus. Conservation Genetics 5:501–512.

Salomonsen, F. (1979). Ornithological and ecological studies in S.W. Greenland (59°46′-62°27′ N. Lat.). Meddelelser om Grønland 204:1–214.

Schubert, M., S. Lindgreen, and L. Orlando (2016). AdapterRemoval v2: rapid adapter trimming, identification, and read merging. BMC Research Notes 9:88.

Shields, G. F., and A. C. Wilson (1987). Calibration of mitochondrial DNA evolution in geese. Journal of Molecular Evolution 24:212–217.

Skarphéðinsson, K. H. (2013). Haförninn. Fuglavernd (Fuglaverndarfélag íslands), Reykjavik.

Smith, J. M., and J. Haigh (1974). The hitch-hiking effect of a favourable gene. Genetical Research 23:23–35.

Tajima, F. (1989). Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics 123:585–595.

Tajima, F. (1996). The amount of DNA polymorphism maintained in a finite population when the neutral mutation rate varies among sites. Genetics 143:1457–1465.

Tamura, K., and M. Nei (1993). Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. Molecular Biology and Evolution 10:512–526.

Thermo Fisher (2016). Thermo Scientific GeneJET Genomic DNA Purification Kit #K0721, #K0722. https://www.thermofisher.com/document-connect/document-connect.html?url = https%3A%2F%2Fassets.thermofisher.com%2FTFS-Assets%2FLSG%2Fmanuals%2FMAN0012667_GeneJET_Whole_Blood_Genomic_DNA_Purification_Mini_Kit_UG.pdf&title = VXNIciBHdWlkZTogR2VuZUpFVCBXaG9sZSBC

Treinys, R., D. Dementavičius, S. Rumbutis, S. Švažas, D. Butkauskas, A. Sruoga, and M. Dagys (2016). Settlement, habitat preference, reproduction, and genetic diversity in recovering the White-tailed Eagle *Haliaeetus albicilla* population. Journal of Ornithology 157:311–323.

Ueta, M., F. Sato, E. G. G. Lobkov, and N. Mita (1998). Migration route of White-tailed Sea Eagles *Haliaeetus albicillain* northeastern Asia. Ibis 140:684–686.

Watterson, G. A. (1975). On the number of segregating sites in genetical models without recombination. Theoretical Population Biology 7:256–276.

Weir, J. T., and D. Schluter (2008). Calibrating the avian molecular clock. Molecular Ecology 17:2321–2328.

Xu, L., and Q. Zhou (2020). The female-specific W chromosomes of birds have conserved gene contents but are not feminized. Genes 11:1–14.

Zeng, K., Y. X. Fu, S. Shi, and C. I. Wu (2006). Statistical tests for detecting positive selection by utilizing high-frequency variants. Genetics 174:1431–1439.

Zink, R. M., and G. F. Barrowclough (2008). Mitochondrial DNA under siege in avian phylogeography. Molecular Ecology 17:2107–2121.