



**HAL**  
open science

## Diversity of major histocompatibility complex of II B gene and mate choice in a monogamous and long-lived seabird, the Little Auk (*Alle alle*)

Katarzyna Wojczulanis-Jakubas, Brian Hoover, Dariusz Jakubas, Jérôme Fort, David Grémillet, Maria Gavriilo, Sylwia Zielińska, Magdalena Zagalska-Neubauer

### ► To cite this version:

Katarzyna Wojczulanis-Jakubas, Brian Hoover, Dariusz Jakubas, Jérôme Fort, David Grémillet, et al.. Diversity of major histocompatibility complex of II B gene and mate choice in a monogamous and long-lived seabird, the Little Auk (*Alle alle*). PLoS ONE, 2024, 19 (6), pp.e0304275. 10.1371/journal.pone.0304275 . hal-04723989

**HAL Id: hal-04723989**

**<https://hal.science/hal-04723989v1>**

Submitted on 7 Oct 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

## RESEARCH ARTICLE

# Diversity of major histocompatibility complex of II B gene and mate choice in a monogamous and long-lived seabird, the Little Auk (*Alle alle*)

Katarzyna Wojczulanis-Jakubas<sup>1\*</sup>, Brian Hoover<sup>2</sup>, Dariusz Jakubas<sup>1</sup>, Jérôme Fort<sup>3</sup>, David Grémillet<sup>4</sup>, Maria Gavrilov<sup>5†</sup>, Sylwia Zielińska<sup>6</sup>, Magdalena Zagalska-Neubauer<sup>7</sup>

**1** Department of Vertebrate Ecology and Zoology, Faculty of Biology, University of Gdansk, Gdansk, Poland, **2** Farallon Institute, Petaluma, California, United States of America, **3** Littoral, Environnement et Sociétés (LIENSs), UMR 7266 CNRS – La Rochelle University, 17000 La Rochelle, France, **4** Excellence Chair Nouvelle Aquitaine - CEBC UMR 7372 CNRS, La Rochelle Université, Villiers-en-Bois, France & FitzPatrick Institute of African Ornithology, University of Cape Town, Rondebosch, South Africa, **5** National Park Russian Arctic, Archangelsk, Russia, **6** Department of Molecular Biology, Faculty of Biology, University of Gdansk, Gdansk, Poland, **7** Department of Behavioural Ecology, Faculty of Biological Sciences, University of Wrocław, Wrocław, Poland

✉ Current address: Arctic and Antarctic Research Institute, Saint-Petersburg, Russia

\* [katarzyna.wojczulanis-jakubas@ug.edu.pl](mailto:katarzyna.wojczulanis-jakubas@ug.edu.pl)



## OPEN ACCESS

**Citation:** Wojczulanis-Jakubas K, Hoover B, Jakubas D, Fort J, Grémillet D, Gavrilov M, et al. (2024) Diversity of major histocompatibility complex of II B gene and mate choice in a monogamous and long-lived seabird, the Little Auk (*Alle alle*). PLoS ONE 19(6): e0304275. <https://doi.org/10.1371/journal.pone.0304275>

**Editor:** Tzen-Yuh Chiang, National Cheng Kung University, TAIWAN

**Received:** September 8, 2023

**Accepted:** May 9, 2024

**Published:** June 12, 2024

**Peer Review History:** PLOS recognizes the benefits of transparency in the peer review process; therefore, we enable the publication of all of the content of peer review and author responses alongside final, published articles. The editorial history of this article is available here: <https://doi.org/10.1371/journal.pone.0304275>

**Copyright:** © 2024 Wojczulanis-Jakubas et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All sequences obtained, which are the core of the study, are

## Abstract

The major histocompatibility complex (MHC) plays a key role in the adaptive immune system of vertebrates, and is known to influence mate choice in many species. In birds, the MHC has been extensively examined but mainly in galliforms and passerines while other taxa that represent specific ecological and evolutionary life-histories, like seabirds, are underexamined. Here, we characterized diversity of MHC Class II B exon 2 in a colonial pelagic seabird, the Little Auk (or Dovekie *Alle alle*). We further examined whether MHC variation could be maintained through balancing selection and disassortative mating. We found high polymorphism at the genotyped MHC fragment, characterizing 99 distinct alleles across 140 individuals from three populations. The alleles frequencies exhibited a similar skewed distribution in both sexes, with the four most commonly occurring alleles representing approximately 35% of allelic variation. The results of a Bayesian site-by-site selection analysis suggest evidence of balancing selection and no direct evidence for MHC-dependent disassortative mating preferences in the Little Auk. The latter result might be attributed to the high overall polymorphism of the examined fragment, which itself may be maintained by the large population size of the species.

## Introduction

Producing two major classes of cell surface receptors (class I and class II), the polymorphic genes of the major histocompatibility complex (MHC) play an essential role in the adaptive

available in GenBank (Accession numbers: GenBank OR909382-OR909480).

**Funding:** Our study was supported by grants from Poland through the Polish Ministry of Science and Education (Juventus Plus 0470/P01/2010/70), samples on Greenland were collected based on the ADACLIM project (No 388), funded by the French Polar Institute. (IPEV) and LIAK&CC project (project Marie Curie IEF number 273061) funded by the European Commission. Sampling conducted in the specially protected area of the Franz-Josef Land State Federal Refuge, within the framework of the joint international expedition of National Park Russian Arctic and a National Geographic 'Pristine Seas Expedition FJL 2013', was supported by National Geographic, Blancpain and Davidoff Cool Water. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

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

**Abbreviations:** cDNA, complementary deoxyribonucleic acid; DNA, deoxyribonucleic acid; dNTPs, deoxynucleotide triphosphates; MHC, major histocompatibility complex; PBS, peptide-binding sites; PCR, polymerase chain reaction; RNA, ribonucleic acid.

immune function of vertebrates [1–3]. Since MHC receptors are encoded by individual alleles that respond only to a fixed number of antigenic peptides, more diverse MHC genotype provide better protection from a broad range of pathogens and is advantageous for disease resistance [4, 5]. Therefore, high levels of MHC polymorphism in a population often reflect selection pressure from pathogens [6–10].

Several genetic mechanisms have been hypothesized to maintain MHC diversity, with three most frequently evoked: heterozygote advantage [11], rare-allele advantage [12, 13] and fluctuating selection [14–16]. These mechanisms represent forms of balancing selection that favour the existence of multiple alleles within a population [17]. Since MHC genotypes may be phenotypically detectable [18–22], MHC polymorphism, especially MHC class II [23], may be behaviourally maintained through strategies, such as mate choice, wherein the genotypes of both parents influence the immunocompetence of the offspring [24]. In this context, the genetic compatibility hypothesis assumes that individuals may produce MHC-heterozygous offspring if they mate disassortatively in respect to their own MHC profile (i.e. partner's genes are supposed to be sufficiently different or complementary to own genotype) [20, 25].

While the avian MHC complex has been extensively examined, the majority of these studies has been performed on specific ecological/taxonomical groups such as galliforms or passerines. In the context of mate choice and heterogenous distribution of pathogens in the environment, however, other avian taxa that represent distinct ecological and evolutionary life-histories, may experience different selection pressure on MHC polymorphism [26, 27]. Therefore, any general interpretation of the mechanisms and phylogenetic origins of MHC diversity in wild bird populations should ideally be based on a wide array of avian taxa [28]. Consequently, there is a strong need to characterize MHC patterns across different avian species.

Due to habitat preferences and some specific life-history traits, seabirds represent an interesting group to examine the functioning of the immune system, including MHC (e.g. [23, 29]). Seabirds spend most of the year in the marine environment where they are exposed to an array of specific pathogens and endo- and ectoparasites [30–32], and that may differentiate them from other avian groups. In addition, many seabirds breed colonially in dense aggregations that can promote elevated transmission rates of pathogens, and the strength of selection on MHC diversity (especially in class II genes) has been shown to increase with increased colonality and migratory behavior in birds [33]. Recent outbreaks of avian influenza have had particularly strong effects on (colonial) seabirds [34, 35], highlighting the need for a deeper understanding of the functioning immune system in this group. Finally, seabirds form long-term pair bonds that are critical for the successful incubation and provisioning of the brood, and thus represent a promising avenue for MHC-dependent mate-choice studies [36–39]. In this study, we focus on the Little Auk, also known as the Dovekie (*Alle alle*), to characterize their MHC diversity and to test whether MHC variation could be maintained in this species through MHC-dependent mating preferences.

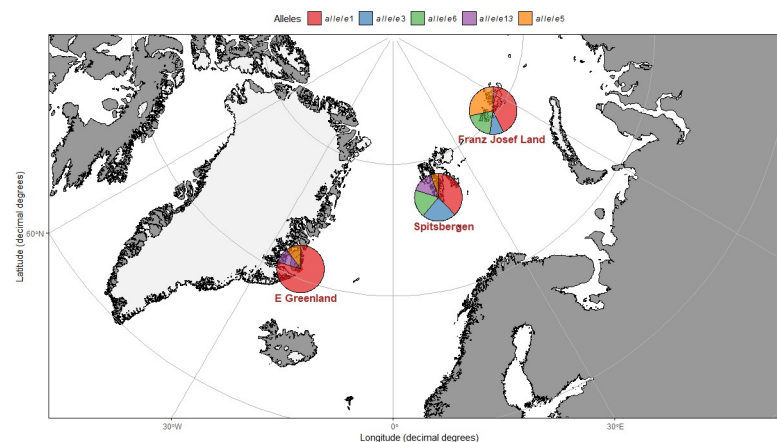
The Little Auk is a small (130–180 g) pelagic seabird that breeds exclusively in the High Arctic [40], with wintering areas also located mainly in sub-Arctic and Arctic zones [41]. This makes the species particularly intriguing in the context of MHC studies as polar marine ecosystems are characterized by relatively low levels of pathogen abundance [42–46]. Such environmental conditions could be expected to relax pathogenic selection on the MHC diversity in Little Auks. However, Little Auks also breed colonially in high densities that may facilitate pathogen transmission [27] (for example, via direct contact during sexual or aggressive interactions [47]), which in turn could increase selection pressure on MHC alleles, therefore increasing MHC diversity. These contrasting scenarios predict opposing patterns of MHC diversity in the Little Auk and raise the need to understand the role of MHC in mating preferences and population biology of this species. In this study, we therefore analyzed MHC alleles

across three main breeding locations (comprising two morphologically distinguishable subspecies) to recognize MHC diversity of the species. Since selecting a mate is a critical decision with long-term fitness consequences in Little Auks (long-term pair bonds, social and genetic monogamy [48], preference of partners with particular phenotypic traits [49]), and MHC II genotypes are known to mediate mate-choice decisions [24], we also tested the hypothesis that Little Auks exhibit MHC-dependent mating preferences.

## Materials and methods

### Samples collection

To characterize MHC diversity in Little Auks, we collected small blood samples (ca. 20  $\mu$ L of whole blood collected from underwing vein and preserved in 90% ethanol) from 140 birds originating from three breeding colonies: Hornsund, Spitsbergen (77° 00'N, 15° 33'E, n = 99 in 2014); Ukaleqarteq (Kap Höegh; in 2013), East Greenland (70° 43'N, 22° 38'W, n = 12 in 2013); and Tikhaya Bay, Hooker Island, Franz Josef Land (80° 20'N; 52° 49'E, n = 29) (Fig 1). Spitsbergen and Greenland colonies are inhabited by the nominative subspecies *Alle a. alle*, while Franz Josef Land is inhabited by the subspecies *Alle a. polaris*. The two subspecies are distinguished based on body size differences (the nominative subspecies is smaller) but are genetically very similar when assessed using neutral genetic markers [50]. The three colony locations sampled represent the most important breeding aggregations of the species (after the Thule area in Northwest Greenland, [51]). The birds sampled in Spitsbergen were breeding birds (44 pairs plus single male-sexed molecularly), captured in the nest while incubating. The birds from Greenland and Franz Josef Land were also breeding adults but represented random individuals captured in mist-net or foot-loop traps during the chick rearing period (breeding status was established based on the gular pouch containing food for the chick). The sex of the birds sampled from Greenland and Franz Josef Land was not established but given the randomized capture protocols and similar colony attendance pattern for males and females, the sex ratio of these samples should be balanced. In summary, samples collected from all three colonies were used to characterize the general pattern of MHC diversity in the species, and the samples collected from the Spitsbergen colony were used to identify sex specific patterns of



**Fig 1. The relative frequency (bars) of five the most frequent alleles (overall observed in >10 individuals) across the three sampling locations representing three little auk breeding colonies.** The map has been prepared in *ggOceanMaps* [101] and *scatterpie* [102] in packages R software 3.6.3 (R Core Team 2020). Map source: Natural Earth (public domain): <http://www.naturalearthdata.com/>.

<https://doi.org/10.1371/journal.pone.0304275.g001>

MHC diversity and potential MHC-dependent mating preferences within known breeding pairs.

All birds were individually marked with ornithological metal rings to avoid double sampling of the same individual. Eight birds from Greenland population were sampled *post-mortem* as subjects of another project (ADACLIM no: 388 of JF and DG). From those individuals, the spleen was collected immediately (<5 min) after euthanasia. The spleen was then preserved in RNAlater solution (Qiagen) and stored at -20°C until the laboratory analysis. We used the spleen samples to amplify and sequence complementary DNA (cDNA), to control and validate the expression status of MHC class II B sequences obtained from DNA samples (see the details below).

### Ethics statement

All the samples were collected following rules for the use of animals, as specified in the guidelines of the Association for the Study of Animal Behaviour. All blood sampling and bird euthanasia was performed under permission issued by the Norwegian Animal Research Authority, the Governor of Svalbard (the Spitsbergen population) and Government of Greenland (Ministry of Fisheries, Hunting and Agriculture, the Greenland population): 2007/00150–9, 2007/00150–11, 17/00663–2, 17/00663–7; 2011–047447, and permits from Russian Ministry of Higher Education (the Franz Josef Land population): 71 from 03.06.2013.

### DNA and RNA extraction

DNA was extracted from the blood using the kit Blood Mini (A&A Biotechnology, Poland). RNA was extracted from the spleen tissue with the use of MasterPure™ Complete DNA and RNA Purification Kit (Epicentre), according to manufacturer's protocol. For tissue homogenization PRO DPS-20 Multi-Sample Dual Processing Homogenizing System (mechanical homogenization– 16000rpm for 60 sec—2 cycles, ultrasonic homogenization– 75% input amplitude for 30 sec) was applied. Extracted RNA was portioned and stored at -80°C for further use.

### Molecular sexing

For sexing (Spitsbergen population only), we performed PCR using the avian sexing primer pair P2 and P8, following the PCR procedure described in [48]. To evaluate PCR products, we applied standard electrophoresis procedure by dyeing the fragments with Midori Green (Nippon, Europe) and examining bands under UV light using a 2.5% agarose gel.

### Primer development

To characterize MHC diversity in the Little Auk, we identified and sequenced the second exon of MHC class II B genes, which is the most polymorphic segment of the peptide-binding region (PBR) [52]. First, we designed and tested a set of primers (Table 1 and Supplementary materials: S1 Table), based on available sequences for closely related alcid species: the Razorbill (*Alca torda*) (Accession Number: EU326270.1, EU326269.1, EU326268.1), the Common Guillemot (*Uria aalge*) (Accession Number: EU326278.1, EU326276.1, EU326275.1, EU326274.1) and the Atlantic Puffin (*Fratercula arctica*) (Accession Number: EU326267.1, EU326267.1, HQ822507.1, HQ822494.1). In the next step, the designed primers along with universal ones (C20, B21, [53]) were used in vectorette PCR, to obtain longer exon sequences covering the region, where we intended to locate the final primers for genotyping (Table 1 and Supplementary materials: S1 Table). In the vectorette PCR, we used total genomic DNA from the eight

**Table 1. Sequences of primers that identify exon 2 MHC in the Little Auk, used for final MHC genotyping.**

Primer name	Sequence 5'–3'
AukLz1	ATGTCTGCMCGAGCAGGGWA
AukRz3	CCRGGGCTTGGCTTGTGCCTG
AukL0709	GCCCTTGGGCACTGGTGCA
AukR0709	TCCGTTCTGCATCACCTCCGT
AukL0809	GCAGGAGGATGCTGTGCAGGGA
AukR0809	TCACCAGCACCTGGTAGGTCCAGTC

<https://doi.org/10.1371/journal.pone.0304275.t001>

individuals digested with four restriction enzymes to generate vectorette-ligated DNA templates.

For vectorette PCR, we adopted the modified vectorette PCR protocol from [53] as it allowed us to obtain information on poorly known DNA fragments of interest, and has been applied successfully in MHC research (e.g. [54–56]). In this technique, one primer (so called “internal”) is designed based on the available sequence, and the second primer is universal (C20 and B21 used as forward or reverse). This approach allows the ‘unknown’ part of the sequence to be extended in one direction and amplified (contrary to standard PCR, which requires two specific primers for amplification). Therefore, the technique provides an opportunity to design more precise, final primers for focal fragments.

First, to construct vectorette libraries, we used ca. 5 µg portion of genomic DNA which was digested with 10 U of enzymes: MunI, XapI, EcoRI, Bsu15i REs (Fermentas) at 37°C for 5 min (other steps, including ligation, were performed according [53]). Then, we have performed vectorette PCRs, where 15 µl of reaction mixture contained 7.5 µl QIAGEN Multiplex PCR Master Mix, 0.15 µM of the specific (AukLz1 or AukLz2/AukRz1, Table 1 and Supplementary materials: S1 Table) and C20 vectorette [53] primers, and 0.5 µl of vectorette-ligated digested DNA. The PCR touchdown was as follows: 95°C/15 min, 5×(94°C/30 s, 64°C/30 s, 72°C/60 s), 5×(94°C/30 s, 60°C/30 s, 72°C/60 s), 20×(94°C/30 s, 58°C/30 s, 72°C/60 s), 72°C/3min. In the second nested vectorette we used PCR 0.5 µl of product of the first PCR as template. 30 µl of PCR mixture contained 2 µL of 10× PCR buffer with (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 2 mM MgCl<sub>2</sub>, 1 µM of each primer, 0.2 mM of each dNTP and 1 U of Taq polymerase (Fermentas). In this PCR we used one specific primer (Alal1L/Alal2R/Alal3R, Supplementary materials: S1 Table) and B21 vectorette primer [53], PCR scheme was as follows: 94°C/3 min, 30×(94°C/30 s, 58°C/30 s, 72°C/60 s), 72°C/3 min. The second PCR products were run on a 1.5% agarose gel and visible fragments (ca. 300 bp) were cut off from the gel and purified with MiniElute Gel Purification Kit (Qiagen). These fragments were Sanger-sequenced with specific vectorette primers. The obtained sequences were then compared to the NCBI nucleotide resources with nblast (Nucleotide BLAST, NCBI). Based on the sequencing results, we selected fragments that aligned to MHC II B exon 2 and corresponded to sequences known from closely related auk species. Ultimately, we were able to select final primers: AukLz1 and AukRz3 for MHC II B second exon genotyping (Table 1). The PCR product of additional pairs of primers (AukR0709, AukL0709 and AukL0809, AukR0809; Supplementary materials: S1 Table) was also identified as a region of interest, i.e. the second and partial third exon of the MHC class II gene. However, generated fragments were too long to use freely with Ion Torrent technology.

### Primers validation

The complementary DNA (cDNA) samples were used to generate sequences which were regarded as a control of the reliability of functionality of MHC fragments amplified with

designed specific primers. Fragments obtained from cDNA samples represent putative, functional alleles, not non-functional pseudogene. Therefore, in the following step we confronted MHC sequences for DNA and cDNA samples generated with use of final primers: AukLz1 and AukRz3 (Table 1). Finally, before performing the next-generation sequencing of all samples, to confirm target sequence (MHC II, exon 2 fragment), we also checked the PCR product (of DNA and cDNA samples) by Sanger sequencing in two directions with both final PCR primers. For validation, the sequenced fragments were aligned with use of the Basic Local Alignment Search Tool (BLAST).

To obtain cDNA we used RNA template ( $n = 8$ ) and performed reverse transcription to generate cDNA. In the first step (pre-processing) we added 2  $\mu$ l of RNA template (in total around 50ng) with 0,5  $\mu$ l of 10 mM primer AukRz1 (degenerate primer, Table 1) and 10,5  $\mu$ l of nuclease-free treated water and incubated in 65°C for 5 min, according to manufacturer's protocol. Then, we cooled the sample to room temperature. In the next step, we added 4  $\mu$ l of 5x Reaction Buffer (part of RevertAid™ M-MuLV) to the preprocessed sample with 2  $\mu$ l of mixed dNTP (10 mM, Thermo Scientific) and 1  $\mu$ l of reverse transcriptase RevertAid™ M-MuLV, Thermo Scientific. We incubated the mixture for 1h at 42°C according to manufacturer's protocol and the reverse transcriptase was inactivated for 10 min at 70°C.

Subset of DNA samples and all cDNA samples were used as a matrices in an amplification reaction in 21  $\mu$ l volume with the use of QIAGEN Multiplex PCR Kit (10  $\mu$ l of 2x QIAGEN Multiplex PCR Kit, 2  $\mu$ l of each 10 mM final primer and 6  $\mu$ l of water). For each reaction we used a 1  $\mu$ l of template. We performed all reactions using a thermal cycler (Eppendorf) under the following conditions settings: 94°C for 15 min for initial denaturation followed by 34 cycles of 94°C for 30 sec, 71°C for 90 sec and 72°C for 90 sec, with additional 72°C for 10 min at the end. Obtained products were visualized by gel electrophoresis, Sanger sequenced and aligned.

### Next generation sequencing

We genotyped 298-bp long amplicons of the Little Auk MHC class II B second exon with Ion Torrent Personal Genome Machine (PGM, Life Technologies, Carlsbad, CA, USA) in the Laboratory of Molecular Biology Technics at Adam Mickiewicz University in Poznań (Poland). First, we performed PCR with so-called fusion primers comprised of Ion Torrent adaptors (A adaptor and P1 adaptor), a 10-bp barcode, and our designed pair of specific final primer (for PCR conditions see the section above). The unique barcodes were used to assign amplicons to individuals (Supplementary materials: S2 Table). In total, a combination of 16 tagged forward primers and 16 tagged reverse primers was used to amplify individual samples (with different combinations of tags and primers). Two independent PCRs were performed on two sets of samples: 140 (all) and 126 individuals (randomly selected of the 140). Following amplification we prepared two sets of individuals pooled equimolarly and sequenced independently to estimate genotyping error. To check the quality of both sequencing data sets we used *Rqc Bioconductor* package [57].

### Data filtering

The processing of raw data was conducted with AmpliSAT, which provides a wide range of online tools for amplicon analysis, and is also appropriate for technology that generates single-ended reads such as Ion Torrent ([58]; <http://evobiolab.biol.amu.edu.pl/amplisat/index.php>). Following the pipeline AmpliSAT manual, we applied the AmpliCLEAN, AmpliCHECK, AmpliSAS, AmpliCOMPARE, and AmpliCOMBINE tools. These tools provided a standardised approach for identifying and filtering reads, variants and amplicons according to their similarity, coverage and frequency. We removed reads with a lower average Phred quality

score ( $< 30$ ) and reads with anomalous length (either too short or too long, as identified by AmpliCLEAN). To filter out the putative alleles from artifacts we used default Ion Torrent parameters: 0.5% substitution error rate, 1% indel error rate and minimum per amplicon frequency of 1%. We also removed sequencing errors [59] and chimeras (the most abundant ones) [58, 60] setting 5% frequency threshold of substitution (AmpliCHECK). To compare replicated genotyping results first, we combined multiple genotyping results (AmpliCOMBINE) and then checked for variants present in one or both compared files (AmpliCOMPARE). For identified alleles, variant depth and other statistics were summed together from both runs of sequencing, and sequencing errors were removed based on their low frequency and/or absence in one set of individuals [54, 60].

In total, we successfully sequenced 140 individuals (including those for which we obtained cDNA samples). Of these, 126 individuals were sequenced twice. Individuals that were sequenced only once were included in analyses if they provided at least 200 reads per allele. All sequences exhibited the correct size (298 bp) with no stop codons and no frame-shift mutations which allow exclude putative pseudogenes. They also correspond to sequences obtained for cDNA samples. Therefore, obtained MHC class II exon 2 fragments were considered putative alleles and submitted to GenBank (Accession numbers: OR909382-OR909480).

### Genetic analysis

We used the DNAsp software [61] to calculate basic polymorphism statistics, including the overall number of segregating (variable) sites, nucleotide diversity ( $\pi$ ), Watterson's estimate of the population mutation rate ( $\theta\omega$ ), and the average number of nucleotide differences among alleles ( $k$ ). We used Tajima's D test [62] within DNAsp to analyse the frequency distribution of variable sites, based on a window size of 25 nucleotides and a step size of 5 nucleotides. Departure from neutral selection at MHC Class IIB in Little Auks was tested within putative PBS and non-PBS regions, which were defined using [63] and previous seabird genetic studies, i.e. [29]. We tested for evidence of positive selection across the alignment of exon 2 sequences using the HyPhy package [64] in the webserver [www.datamonkey.org](http://www.datamonkey.org), and applied the Random Effects Likelihood method to investigate site-by-site evidence of positive selection. As recombination can overestimate the degree of positive selection [65], we ran the Genetic Algorithm for Recombination Detection module within the HyPhy package to test for significant evidence of recombination breakpoints. Finally, we used the software MEGA 4.0 [66] to test synonymous vs. non-synonymous substitution rates within codons presumably representing the protein-binding region within each amplicon (PBR) [67], using the Modified Nei-Gojorobi method with Jukes-Cantor correction.

### Mate-choice analysis

We used four distinct genetic metrics to test whether Little Auks select mates on the basis of MHC compatibility: i) differences in heterozygosity; ii) MHC-band sharing coefficients [68]; iii) additive amino acid differences; iv) maximum amino acid difference. Heterozygosity differences were calculated as the overall difference in MHC heterozygosity between each breeding pair, and can only be used to test whether disassortative mating occurs within a population (e.g. whether males or females are on average significantly more heterozygous at MHC II B than their partners). Band-sharing represents the proportion of unique MHC alleles shared within a breeding pair, and is calculated as twice the number of alleles shared within a pair, divided by the sum of unique alleles occurring within both individuals [ $(2F_{ab}/(F_a+F_b))$ , as per [67]]. Amino acid differences represent the number of amino acid substitutions occurring between pairs of different MHC alleles, and were quantified here as either the sum of all the



differences between a pair at both loci, or the maximum difference observed between loci. We assessed these amino acid substitutions within two different contexts; i) using only the substitutions occurring within PBR regions of the MHC II B sequence, and ii) using substitutions occurring throughout the entire fragment of the exon sequence. For each of the four criteria listed above, we used Monte Carlo randomization tests [69] to determine whether average observed mating patterns for each metric were significantly different from simulated expectations of random mating. To obtain random mating distributions, we resampled males and females from the study population without replacement and formed new random pairs, performing 10,000 iterations of this randomization procedure. We then compared the mean observed values for each metric (i.e. the values we calculated for naturally occurring pairs) with the distribution of randomly calculated means. We determined statistical significance by comparing the observed mean of the breeding pairs with the 2.5% and 97.5% tail of the randomized distribution of means, and Bonferroni-adjusting the alpha value by the number of randomization tests ( $0.05/6 = 0.008$ ). For each metric, we assessed the likelihood of error within Monte Carlo randomization tests by iteratively decreasing the sample size and examining the effect of sample size on the consistency of obtaining a given result (for further explanation, see [70]). We performed all statistical analyses in R 3.6.3 software [71].

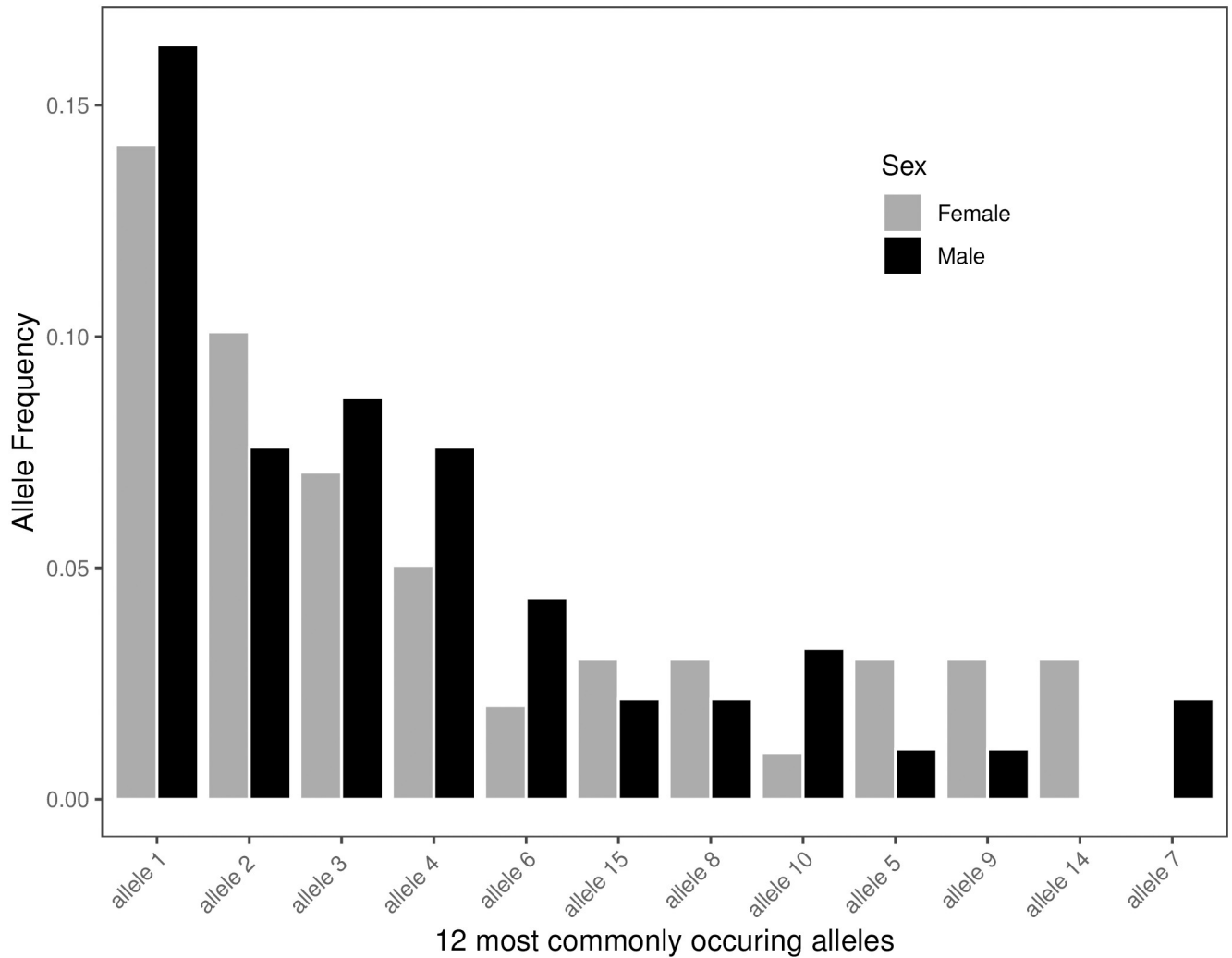
## Results

### MHC diversity and selection

We characterized a high level of genetic diversity at MHC II B exon 2. We sequenced 298 bp fragments that contained 99 distinct alleles across 140 individuals. Within this set of individuals, eight samples were also used for cDNA sequencing (spleen samples). We did not perform any formal analysis but examining alleles from the eight individuals, we could confirm the expression status of identified MHC fragments (putative alleles), generated by primers, thus MHC diversity we quantified in the Little Auk is reliable.

A maximum of four alleles was documented within an individual, indicating at least two loci, i.e. the presence of a duplicated MHC II gene fragment, which is consistent with findings in other alcid species [72, 73]. On average two alleles per individual were recorded, both within the total sample pool and separately for each population. Overall, four individuals had 4 alleles (3%), 39 individuals had 3 alleles (28%), 56 individuals had two alleles (40%), and 41 individuals had one allele (29%).

Out of 99 alleles identified from nucleotide sequences 54 were present only in a single individual, 19 other alleles were present in two individuals, and remaining 26 alleles were observed in three or more individuals (see Supplementary materials: S3 Table). The number of alleles identified per population was: in Spitsbergen—75 alleles (64% of these were unique to the population, of  $n = 99$ ), in Greenland—20 alleles (35% unique, of  $n = 12$ ) and in Franz Josef Land—36 alleles (44% unique, of  $n = 29$ ). The frequency of the five most frequent alleles, i.e. observed in  $>10$  individuals, in each of the sampled colonies is presented on Fig 1. Obtained alleles translated into 89 amino acid sequences. Genotyped alleles exhibited a similar skewed distribution in both sexes, with the four most commonly occurring alleles representing approximately 35% of allelic variation (Fig 2). Basic polymorphism statistics are reported in Table 2. We found a significant excess of non-synonymous over synonymous substitutions for those codons presumably comprising the peptide binding site (PBS)-coding region ( $\omega = 2.404$ ; Z-test,  $p = 0.004$ ), but not for the non-PBS regions ( $\omega = 1.257$ ; Z-test,  $p = 0.132$ ), suggesting evidence of a positive selection. A Bayesian analysis of site-by-site selection across an alignment of exon 2 sequences indicated three sites undergoing a positive selection and one site



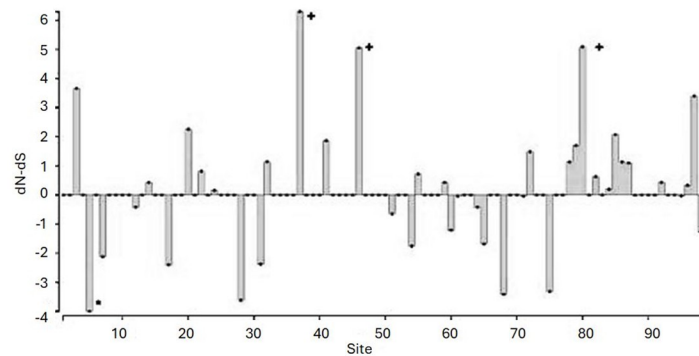
**Fig 2. Distribution of allele frequencies at genotyped MHC IIB loci in the little auk in respect to sex (Spitsbergen colony only).** Only alleles found at frequencies > 2%, in either sex, are depicted.

<https://doi.org/10.1371/journal.pone.0304275.g002>

**Table 2. Polymorphism statistics at genotyped MHC II B loci in Little auks.** Diversity statistics: the overall number of nucleotide sequences found ( $h$ ), the overall number of segregating sites ( $S$ ), the total number of mutations ( $Eta$ ), the nucleotide diversity estimator ( $\pi$ ), the Watterson’s estimator of the population mutation rate ( $\theta_w$ ) and the average number of nucleotide differences among alleles  $K$ .

Diversity statistic	Value
$h$	99 (69)
$S$ ( $Eta$ )	87 (89)
$\pi$	0.063
$\theta_w$	0.069
$K$	16.67

<https://doi.org/10.1371/journal.pone.0304275.t002>



**Fig 3. Bayesian analysis of site-by-site positive selection across an alignment of exon 2 sequences from two MHC class II B gene duplicates in the little auk.** Codons displaying decisive evidence of positive selection (Bayes factors >100) are indicated by crosses. Sites showing decisive evidence of negative selection are indicated by asterisks.

<https://doi.org/10.1371/journal.pone.0304275.g003>

undergoing a negative selection (Fig 3, Bayes Factor 100), while Tajimas D analysis revealed sites with a significant excess of high- frequency segregating sites (S4 Table).

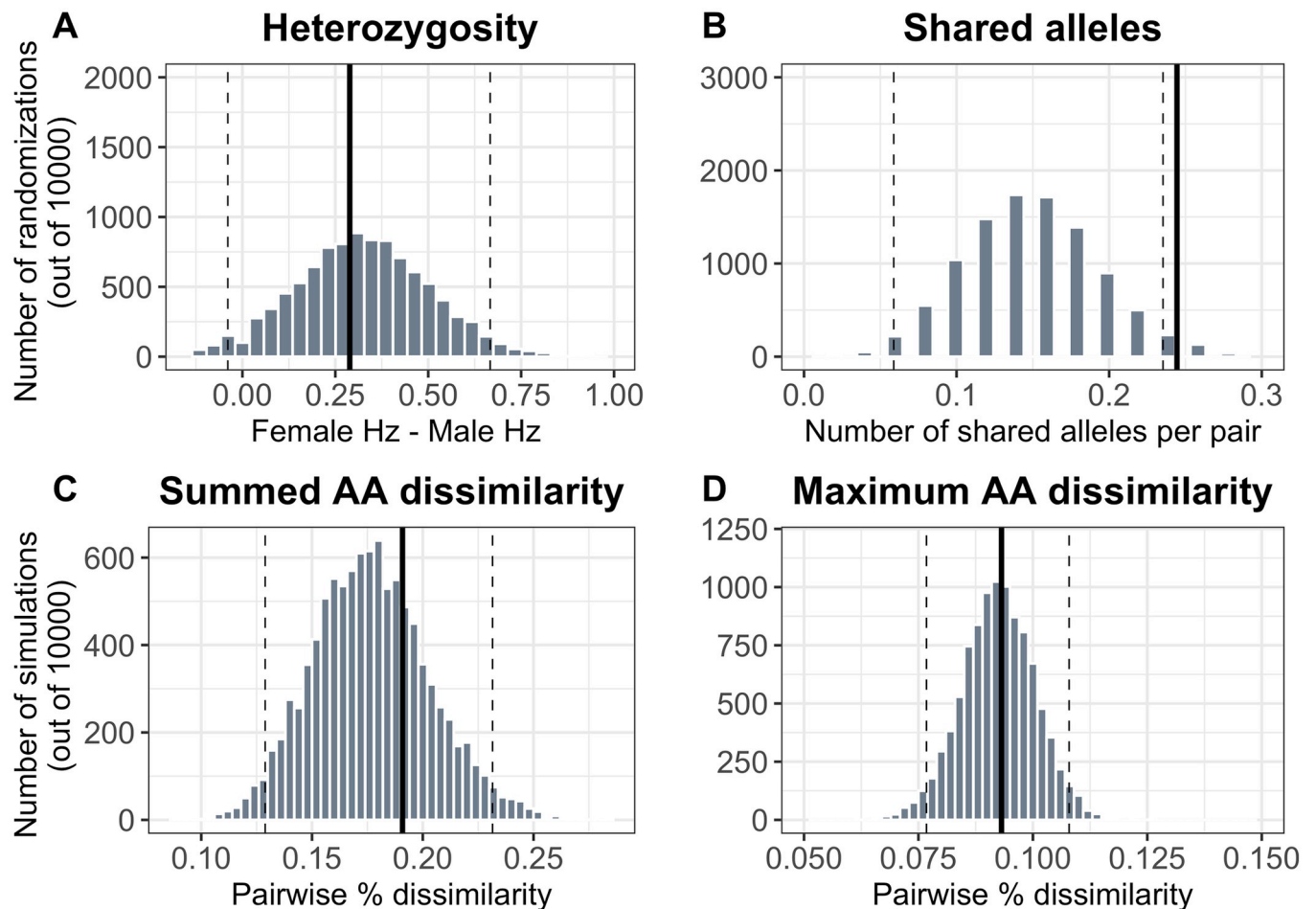
### Disassortative mating

We found no evidence of non-random mating based on overall differences in MHC II heterozygosity, average amino acid dissimilarity, or maximum amino acid dissimilarity using the whole fragment region (Fig 4) or PBS regions (S1 Fig). We found a significant effect of MHC-based assortative mating on shared alleles, suggesting that breeding pairs shared alleles in common more often than would be randomly expected ( $p = 0.016$ ). However, this effect may be biased by the inability to distinguish between duplicated loci. Overall, insufficient evidence is available to conclude the presence of MHC-based mating effects.

### Discussion

Our study revealed a high level of MHC class II B diversity within Little Auks, one of the highest described in non-passerine birds. In 140 screened individuals, we detected 99 different alleles in MHC II B exon 2, with an average 16.2 nucleotide differences between alleles. Four alleles were dominant in the allele distribution (Fig 1) but most genotypes represented rare alleles. We are confident the high number of detected alleles is accurate as we performed PCR in several independent sets, samples were sequenced independently twice, obtained sequences did not resemble chimeras, and were regarded as reliable alleles if present in a high number of reads (minimum 200, in two runs). Additionally, we have applied restricted criteria to sort out sequences. We also used cDNA to exclude putative alleles from the analyses (i.e., a with no stop codons, deletion or frameshift mutations).

The high number of MHC alleles could be generated by a number of different processes such as point mutation, recombination and gene conversion [74–76]. Generated alleles might be then preserved and maintained by adaptive forces such as balancing selection [74–76]. A high genetic diversity is an inherent feature of MHC genes (including class II) that has been reported in numerous avian species. Especially for passerines, the reported MHC loci and alleles can reach very high numbers, e.g. Sedge Warbler (*Acrocephalus schoenobaenus*) [77], Great Tit (*Parus major*) [78], Collared Flycatcher (*Ficedula albicollis*) [54]. In non-passerine birds the highest allelic MHC diversity has been reported in the Eurasian Coot (*Fulica atra*; 3 loci, 265 alleles) [79], the Greater Flamingo (*Phoenicopterus roseus*; 2 loci, 109 alleles) [80], and



**Fig 4.** A-D. Results of randomization tests testing for MHC-based non-random mating. No significant effects of MHC genotype were found when assessing differences in heterozygosity (4A) or dissimilarities in amino acid composition (4C–4D). Little Auks exhibited a significant assortative mating effect based on shared alleles (4B), which may be attributed to the frequency of the most commonly genotyped allele in the population and the inability to discriminate between locus-specific alleles. Dashed vertical lines denote 95% of random distribution, whereas the solid vertical lines indicate the observed mean value. An observed value located within the 95% random distribution indicates a random mating pattern, whereas an observed mean located outside the 95% range on the right side indicates assortative mating.

<https://doi.org/10.1371/journal.pone.0304275.g004>

the Lesser Kestrel (*Falco naumanni*; 1 locus, 103 alleles) [81]. With 2 loci per individual and a total of 99 identified MHC alleles, the Little Auk polymorphism is lower than these examples, however it remains one of the highest values reported among non-passerines and represents the highest polymorphism value detected so far among seabirds.

This high MHC allele diversity in the Little Auk is quite surprising given the low pathogen-abundance in the Arctic environment [42–46]. For example, no blood parasites have been detected in the Little Auk [82]. However, many MHC loci class I (but not class II) have also been detected in another polar (Antarctic) seabird, the Blue Petrel (*Halobaena caerulea*) [29]. It has been then suggested that the reason for this is that blue petrels are a colonial species that reuse underground nests, which might favor the horizontal transmission of pathogens, even if the pathogens are not that frequent in the polar environment [29]. The Little Auk is also a colonial species, yet Little Auks nest above ground on cliffs or slopes, and the horizontal transmission of pathogens in this Arctic environment has not been investigated. Thus, the high degree of MHC polymorphism that we identified in Little Auks may represent an ancestral

polymorphism that is enhanced by a large effective population size, both in the past [51] and at present [83], and subsequently maintained by high gene flow among the breeding colonies [50, 84].

MHC diversity in the Little Auk might be greater than we identified in our study, as we sequenced samples from three breeding locations only. Although the sampled colonies represent a large portion of the breeding area for this species, they certainly do not cover the whole range and some other colonies may be genetically distinct (due to their isolation and/or habitat). High frequency of unique alleles in each population and inter-population differences in the frequencies of the most common shared alleles (Fig 1) suggest there is some divergence between breeding populations and/or adaptation to local conditions (pathogen pressure, population size, etc). Although we are not able to fully examine the inter-population differences in MHC diversity (due to limited population coverage and the small number of sampled individuals), our results suggest biogeographical patterns of MHC diversity within the Little Auk population that are worthy of further examination.

In some seabird populations, MHC polymorphism has also been suggested to result from ancient gene duplication events that led to a functional genetic divergence between duplicated loci [37]. Gene duplication at MHC II B appears to be a conserved trait within many seabird clades, including Procellariiformes [38, 85] and alcids [72, 73]. However, in most studies it remains unclear whether such duplicates are functionally identical or could be duplicated loci that have evolved and are expressed independently. While locus-specific MHC II B genotyping was outside the focus of this study, gene duplicates that have diverged and function independently can conceivably maintain high MHC polymorphism [86].

We also found some evidence of positive and balancing selection within Little Auk MHC. As reported in a previous study [33], the strength of balancing selection on MHC class II genes in birds is higher in colonially breeding and migratory species. It may be linked to elevated transmission rates of pathogens in species breeding in dense aggregations, or/and exposure to a more diverse array of pathogens and parasites in migratory species [33]. While we did not investigate gene expression in this study, our evidence supports the idea that the identified alleles are functional and expressed (e.g., [27, 87, 88]). It is further unlikely that the genotyped region represents pseudogenes as those are rarely reported for non-passerines [89–91]. Thus, although we did not examine fitness consequences (e.g. life-time reproductive success) associated with MHC II B polymorphism, in the absence of strong mate choice effects, the high MHC II B diversity we found may be maintained by a combination of pathogen-mediated selection heterozygote advantage, and gene flow, as reviewed in [10, 92, 93].

Selective evolutionary pressure from mate choice behaviours may be particularly strong in species with long-term pair bonds [94, 95]. Indeed, previous studies have reported some evidence of disassortative mating in regard to MHC II class in other seabirds with similar mate-choice strategies [29, 38] but see [37]. For this reason, and given Little Auk life-history traits (partners mate for multiple breeding attempts if not for life [40]), we expected to find evidence of disassortative mating but failed to do so. There could be several reasons for the absence of such a pattern. First, disassortative mating could be hindered by the overall high MHC diversity existing in Little Auks (i.e. 99 alleles across two loci), meaning that most prospective mates are already quite dissimilar at MHC II B. Second, mate choice could be operating on only one of the two characterized loci, and our inability to distinguish loci is masking an existing disassortative mating trend. Third, specific MHC genes may be only partially associated with a separate phenotypic trait valued by Little Auks, and thus mate choice may not operate on MHC disassortative principles at all. Finally, random mating occurring on two separate yet functionally expressed MHC II B loci may be sufficient to maintain the high degree of MHC

polymorphism that we find in the population, yet Little Auks may select mates on the basis of other criteria that may only loosely be correlated with MHC, such as size or experience [49].

Interestingly, we found some evidence of assortative mating. Despite the high number of MHC alleles, breeding pairs shared alleles in common more often than would be expected by chance, yet this result may be biased by the inability to distinguish between duplicated loci. However, there is a supporting rationale for non-random mating in the Little Auk that would support this result, as phenotypic mate choice decisions in the Little Auk have previously been reported to be non-random, with pairs mating assortatively in regard to some morphological traits (wing length, the size of white spot on the eye-lid [49]). Exhibiting a mate choice preference for similar MHC alleles could operate on a similar principle, or MHC alleles could covary with morphological traits (i.e. MHC genotype correlates with waddle size in pheasants [96]), which might be associated with assortative mating. While speculative, this perspective is supported in a broader ecological context: some studies suggest that assortative mating can mitigate sexual conflict over parental care [97]. If so, this could be particularly important dynamic in the case of the Little Auk, which require an extensive parental effort with a high degree of parental coordination during the incubation and provisioning periods, in order to successfully fledge their chicks [98–100].

## Conclusions

Our study found a high level of genetic diversity at the MHC II B gene in a colonial Arctic seabird, the Little Auk, and the 99 alleles we identified is one of the highest examples of MHC polymorphism described in non-passerine birds. The results agree with evolutionary theory that suggest balancing selection helps to create and maintain the high levels of polymorphism within the examined MHC region. Contrary to our expectations, we did not find evidence of disassortative mating regarding the examined MHC region, and more studies are needed to fully understand the mechanism of the mate choice in the Little Auk.

## Supporting information

**S1 Table. Sequences of primers that identify exon 2 MHC in the Little Auk not used for further genotyping as provided longer products, not appropriate for Ion Torrent sequencing.**

(DOCX)

**S2 Table. List of barcodes used in bidirectional tagging for Ion Torrent PGM sequencing.**

(DOCX)

**S3 Table. Full list of alleles with their depth and frequency in the sampled individuals.**

(DOCX)

**S4 Table. Tajima's D test statistics.**

(DOCX)

**S1 Fig. Results of randomization tests testing for MHC-based non-random mating, using average and maximum amino acid dissimilarity of PBS regions.**

(DOCX)

## Acknowledgments

We are grateful to Monika Prełowska, Mirosława Dabert and Marcin Łoś for lab assistance, and Yuki Brooknievskaya for constructive discussion, encouragement and overall support

during the preparing and submission of the manuscript. We also thank Magdalena Migalska for her help and assistance with AmpliSAT data filtering. We are also grateful to Enric Sala for his support in accessing field site and collecting samples in Franz Josef Land. We also warmly thank all fieldworkers for their help in the field. NanuTravel is acknowledged for their logistical support in East Greenland. Finally, we thank four anonymous reviewers for their suggestions that helped to improve the manuscript.

## Author Contributions

**Conceptualization:** Katarzyna Wojczulanis-Jakubas, Brian Hoover, Dariusz Jakubas, Magdalena Zagalska-Neubauer.

**Data curation:** Katarzyna Wojczulanis-Jakubas.

**Formal analysis:** Brian Hoover, Magdalena Zagalska-Neubauer.

**Funding acquisition:** Katarzyna Wojczulanis-Jakubas.

**Investigation:** Katarzyna Wojczulanis-Jakubas, Dariusz Jakubas, Jérôme Fort, David Grémillet, Maria Gavriilo, Sylwia Zielińska, Magdalena Zagalska-Neubauer.

**Project administration:** Katarzyna Wojczulanis-Jakubas.

**Visualization:** Brian Hoover, Dariusz Jakubas.

**Writing – original draft:** Katarzyna Wojczulanis-Jakubas.

**Writing – review & editing:** Brian Hoover, Dariusz Jakubas, Jérôme Fort, David Grémillet, Maria Gavriilo, Sylwia Zielińska, Magdalena Zagalska-Neubauer.

## References

1. Klein J, Figueroa F. Evolution of the major histocompatibility complex. *Critical reviews in immunology*. 1986. pp. 295–386. [https://doi.org/10.1016/0145-305x\(86\)90049-2](https://doi.org/10.1016/0145-305x(86)90049-2) PMID: 3536303
2. Kelley J, Walter L, Trowsdale J. Comparative genomics of major histocompatibility complexes. *Immunogenetics*. 2005. pp. 683–695. <https://doi.org/10.1007/s00251-004-0717-7> PMID: 15605248
3. Murphy K., Travers P, Walport M. *Janeway's immunobiology*. 7th ed. London: Tylor & Francis; 2008.
4. Apanius V, Penn D, Slev PR, Ruff L Ramelle, Potts WK. The nature of selection on the major histocompatibility complex. *Crit Rev Immunol*. 2017; 37: 75–120. <https://doi.org/10.1615/CritRevImmuno.v37.i2-6.10> PMID: 29773018
5. Hawley DM, Fleischer RC. Contrasting epidemic histories reveal pathogen-mediated balancing selection on class II MHC diversity in a wild songbird. Williams BL, editor. *PLoS One*. 2012; 7: e30222. <https://doi.org/10.1371/journal.pone.0030222> PMID: 22291920
6. Hedrick PW. Pathogen resistance and genetic variation at MHC loci. *Evolution (N Y)*. 2002; 56: 1902–1908. <https://doi.org/10.1111/j.0014-3820.2002.tb00116.x> PMID: 12449477
7. Minias P, Gutiérrez JS, Dunn PO. Avian major histocompatibility complex copy number variation is associated with helminth richness. *Biol Lett*. 2020; 16: 20200194. <https://doi.org/10.1098/rsbl.2020.0194> PMID: 32634375
8. Wills C. Maintenance of multiallelic polymorphism at the MHC region. *Immunol Rev*. 1991; 124: 165–220. <https://doi.org/10.1111/j.1600-065x.1991.tb00621.x> PMID: 1804778
9. Silva MC, Edwards S V. Structure and evolution of a new avian MHC class II B gene in a sub-antarctic seabird, the thin-billed prion (*Procellariiformes: Pachyptila belcheri*). *J Mol Evol*. 2009; 68: 279–291. <https://doi.org/10.1007/s00239-009-9200-2>
10. Spurgin LG, Richardson DS. How pathogens drive genetic diversity: MHC, mechanisms and misunderstandings. *Proceedings of the Royal Society B: Biological Sciences*. Royal Society; 2010; 277: pp. 979–988. <https://doi.org/10.1098/rspb.2009.2084> PMID: 20071384
11. Hughes AL, Nei M. Models of host-parasite interaction and MHC polymorphism. *Genetics*. 1992; 132: 863–864. <https://doi.org/10.1093/genetics/132.3.863> PMID: 1468636

12. Slade RW, McCallum HI. Overdominant vs. frequency-dependent selection at MHC loci. *Genetics*. 1992; 132: 861–862. <https://doi.org/10.1093/genetics/132.3.861> PMID: 1468635
13. Takahata N, Nei M. Allelic genealogy under overdominant and frequency-dependent selection and polymorphism of major histocompatibility complex loci. *Genetics*. 1990; 124: 967–978. <https://doi.org/10.1093/genetics/124.4.967> PMID: 2323559
14. Hedrick PW, Ginevan ME, Ewing EP. Genetic polymorphism in heterogeneous environments. *Annu Rev Ecol Syst*. 1976; 7: 1–32. <https://doi.org/10.1146/annurev.es.07.110176.000245>
15. Gillespie JH, Turelli M. Genotype-environment interactions and the maintenance of polygenic variation. *Genetics*. 1989; 121: 129–138. <https://doi.org/10.1093/genetics/121.1.129> PMID: 17246488
16. Hedrick PW. Genetic polymorphism in heterogeneous environments: The age of genomics. *Annual Review of Ecology, Evolution, and Systematics*. 2006. pp. 67–93. <https://doi.org/10.1146/annurev.ecolsys.37.091305.110132>
17. Hughes AL, Yeager M. Natural selection at major histocompatibility complex loci of vertebrates. *Annual Review of Genetics*. *Annu Rev Genet*; 1998. pp. 415–435. <https://doi.org/10.1146/annurev.genet.32.1.415> PMID: 9928486
18. Penn DJ. The scent of genetic compatibility: Sexual selection and the major histocompatibility complex. *Ethology*. John Wiley & Sons, Ltd; 2002. pp. 1–21.
19. Dunn PO, Bollmer JL, Freeman-Gallant CR, Whittingham LA. MHC variation is related to a sexually selected ornament, survival, and parasite resistance in common yellowthroats. *Evolution (N Y)*. 2013; 67: 679–687. <https://doi.org/10.1111/j.1558-5646.2012.01799.x> PMID: 23461319
20. Brennan PA, Zufall F. Pheromonal communication in vertebrates. *Nature*. Nature Publishing Group; 2006. pp. 308–315.
21. Capittini C, Martinetti M, Cuccia M. MHC variation, mate choice and natural selection: The scent of evolution. *Riv di Biol—Biol Forum*. 2008; 101: 463–480. <https://doi.org/10.1400/111894> PMID: 19322759
22. Boehm T, Zufall F. MHC peptides and the sensory evaluation of genotype. *Trends in Neurosciences*. *Trends Neurosci*; 2006. pp. 100–107. <https://doi.org/10.1016/j.tins.2005.11.006> PMID: 16337283
23. Strandh M, Westerdahl H, Pontarp M, Canbäck B, Dubois MP, Miquel C, et al. Major histocompatibility complex class II compatibility, but not class I, predicts mate choice in a bird with highly developed olfaction. *Proc R Soc B Biol Sci*. 2012; 279: 4457–4463. <https://doi.org/10.1098/rspb.2012.1562> PMID: 22951737
24. Kamiya T, O'Dwyer K, Westerdahl H, Senior A, Nakagawa S. A quantitative review of MHC-based mating preference: The role of diversity and dissimilarity. *Mol Ecol*. 2014; 23: 5151–5163. <https://doi.org/10.1111/mec.12934> PMID: 25251264
25. Zelano B, Edwards S V. An MHC component to kin recognition and mate choice in birds: predictions, progress, and prospects. *Am Nat*. 2002; 160. <https://doi.org/10.1086/342897> PMID: 18707479
26. Minias P, Pikus E, Whittingham LA, Dunn PO. A global analysis of selection at the avian MHC. *Evolution (N Y)*. 2018; 72: 1278–1293. <https://doi.org/10.1111/evo.13490> PMID: 29665025
27. Bernatchez L, Landry C. MHC studies in nonmodel vertebrates: What have we learned about natural selection in 15 years? *J Evol Biol*. 2003; 16: 363–377. <https://doi.org/10.1046/j.1420-9101.2003.00531.x> PMID: 14635837
28. Minias P, Pikus E, Whittingham LA, Dunn PO. Evolution of copy number at the MHC varies across the avian tree of life. *Genome Biol Evol*. 2019; 11: 17–28. <https://doi.org/10.1093/gbe/evy253> PMID: 30476037
29. Strandh M, Lannefors M, Bonadonna F, Westerdahl H. Characterization of MHC class I and II genes in a subantarctic seabird, the blue petrel, *Halobaena caerulea* (Procellariiformes). *Immunogenetics*. 2011; 63: 653–666. <https://doi.org/10.1007/s00251-011-0534-8> PMID: 21607694
30. Landsberg J, Van Dolah F, Doucette G. Marine and estuarine harmful algal blooms: Impacts on human and animal health. *Oceans and Health: Pathogens in the Marine Environment*. Springer US; 2005. pp. 165–215. [https://doi.org/10.1007/0-387-23709-7\\_8](https://doi.org/10.1007/0-387-23709-7_8)
31. Munn CB. Pathogens in the sea: An overview. *Oceans and Health: Pathogens in the Marine Environment*. Springer US; 2005. pp. 1–28. [https://doi.org/10.1007/0-387-23709-7\\_1](https://doi.org/10.1007/0-387-23709-7_1)
32. Quillfeldt P, Arriero E, Martínez J, Masello JF, Merino S. Prevalence of blood parasites in seabirds—a review. *Front Zool*. 2011; 8. <https://doi.org/10.1186/1742-9994-8-26>
33. Minias P, Whittingham LA, Dunn PO. Coloniality and migration are related to selection on MHC genes in birds. *Evolution (N Y)*. 2017; 71: 432–441. <https://doi.org/10.1111/evo.13142> PMID: 27921293
34. Hartby CM, Krog JS, Merkel F, Holm E, Larsen LE, Hjulsgager CK. First characterization of avian influenza viruses from Greenland 2014. *Avian Dis*. 2016; 60: 302–310. <https://doi.org/10.1637/11119-050515-RegR> PMID: 27309071



35. Adlhoch C, Fusaro A, Gonzales JL, Kuiken T, Marangon S, Niqueux É, et al. Avian influenza overview June–September 2022. *EFSA J*. 2022; 20. <https://doi.org/10.2903/j.efsa.2022.7597> PMID: 36247870
36. Juola FA, Dearborn DC. Sequence-based evidence for major histocompatibility complex-disassortative mating in a colonial seabird. *Proc R Soc B Biol Sci*. 2011; 279: 153–162. <https://doi.org/10.1098/rspb.2011.0562>
37. Dearborn DC, Gager AB, McArthur AG, Gilmour ME, Mandzhukova E, Mauck RA. Gene duplication and divergence produce divergent MHC genotypes without disassortative mating. *Mol Ecol*. 2016; 25: 4355–4367. <https://doi.org/10.1111/mec.13747> PMID: 27376487
38. Hoover B, Alcaide M, Jennings S, Sin SYW, Edwards S V., Nevitt GA. Ecology can inform genetics: Disassortative mating contributes to MHC polymorphism in Leach's storm-petrels (*Oceanodroma leucorhoa*). *Mol Ecol*. 2018; 27: 3371–3385. <https://doi.org/10.1111/mec.14801>
39. Hunt GL. Mate Selection and Mating Systems in Seabirds. *Behavior of Marine Animals*. Springer US; 1980. pp. 113–151.
40. Stempniewicz L. BWP update. *Little Auk (Alle alle)*. Oxford Uni. BWP Update The Journal of the Birds of the Western Palearctic. Oxford Uni. Oxford: Oxford University Press; 2001. pp. 175–201.
41. Fort J, Moe B, Strøm H, Grémillet D, Welcker J, Schultner J, et al. Multicolony tracking reveals potential threats to little auks wintering in the North Atlantic from marine pollution and shrinking sea ice cover. *Divers Distrib*. 2013; 19: 1322–1332. <https://doi.org/10.1111/ddi.12105>
42. Ghiglione JF, Galand PE, Pommier T, Pedrós-Alió C, Maas EW, Bakker K, et al. Pole-to-pole biogeography of surface and deep marine bacterial communities. *Proc Natl Acad Sci U S A*. 2012; 109: 17633–17638. <https://doi.org/10.1073/pnas.1208160109> PMID: 23045668
43. Rampelotto PH. Polar microbiology: Recent advances and future perspectives. *Biology*. MDPI AG; 2014. pp. 81–84.
44. Cavicchioli R. Microbial ecology of Antarctic aquatic systems. *Nature Reviews Microbiology*. Nature Publishing Group; 2015. pp. 691–706.
45. Boetius A, Anesio AM, Deming JW, Mikucki JA, Rapp JZ. Microbial ecology of the cryosphere: Sea ice and glacial habitats. *Nat Rev Microbiol*. 2015; 13: 677–690. <https://doi.org/10.1038/nrmicro3522> PMID: 26344407
46. Yau S, Seth-Pasricha M. Viruses of polar aquatic environments. *Viruses*. MDPI AG; 2019.
47. Wojczulanis-Jakubas K, Dynowska M, Jakubas D. Fungi prevalence in breeding pairs of a monogamous seabird—Little auk, *Alle alle*. *Ethol Ecol Evol*. 2011; 23: 240–247. <https://doi.org/10.1080/03949370.2011.566582>
48. Wojczulanis-Jakubas K, Jakubas D, Øigarden T, Lifjeld JT. Extrapair copulations are frequent but unsuccessful in a highly colonial seabird, the little auk, *Alle alle*. *Anim Behav*. 2009; 77: 433–438. <https://doi.org/10.1016/j.anbehav.2008.10.019>
49. Wojczulanis-Jakubas K, Drobniak SM, Jakubas D, Kulpińska-Chamera M, Chastel O. Assortative mating patterns of multiple phenotypic traits in a long-lived seabird. *Ibis (Lond 1859)*. 2018; 160: 464–469. <https://doi.org/10.1111/ibi.12568>
50. Wojczulanis-Jakubas K, Kilikowska A, Fort J, Gavrilov M, Jakubas D, Friesen V. No evidence of divergence at neutral genetic markers between the two morphologically different subspecies of the most numerous Arctic seabird. *Ibis (Lond 1859)*. 2015; 157: 787–797. <https://doi.org/10.1111/ibi.12294>
51. Keslinka LK, Wojczulanis-Jakubas K, Jakubas D, Neubauer G. Determinants of the little auk (*Alle alle*) breeding colony location and size in W and NW coast of Spitsbergen. *PLoS One*. 2019; 14: e0212668. <https://doi.org/10.1371/journal.pone.0212668> PMID: 30840697
52. Trowsdale J. Natural history of the major histocompatibility complex. *Trends in Genetics*. New York: John Wiley & Sons, Ltd; 1987.
53. Ko WY, David RM, Akashi H. Molecular phylogeny of the *Drosophila melanogaster* species subgroup. *J Mol Evol*. 2003; 57: 562–573. <https://doi.org/10.1007/s00239-003-2510-x>
54. Zagalska-Neubauer M, Babik W, Stuglik M, Gustafsson L, Cichoń M, Radwan J. 454 sequencing reveals extreme complexity of the class II Major Histocompatibility Complex in the collared flycatcher. *BMC Evol Biol*. 2010; 10: 395. <https://doi.org/10.1186/1471-2148-10-395> PMID: 21194449
55. Kuduk K, Babik W, Bojarska K, Śliwińska EB, Kindberg J, Taberlet P, et al. Evolution of major histocompatibility complex class I and class II genes in the brown bear. *BMC Evol Biol*. 2012; 12. <https://doi.org/10.1186/1471-2148-12-197> PMID: 23031405
56. Schad J, Dechmann DKN, Voigt CC, Sommer S. MHC class II DRB diversity, selection pattern and population structure in a neotropical bat species, *Noctilio albiventris*. *Heredity (Edinb)*. 2011; 107: 115–126. <https://doi.org/10.1038/hdy.2010.173> PMID: 21245894

57. de Souza W, Carvalho de B S, Lopes-Cendes I. Rqc: A bioconductor package for quality control of high-throughput sequencing data. *J Stat Softw.* 2018; 87: 1–14. <https://doi.org/10.18637/jss.v087.c02>
58. Lenz TL, Becker S. Simple approach to reduce PCR artefact formation leads to reliable genotyping of MHC and other highly polymorphic loci—Implications for evolutionary analysis. *Gene.* 2008; 427: 117–123. <https://doi.org/10.1016/j.gene.2008.09.013> PMID: 18848974
59. Shendure J, Ji H. Next-generation DNA sequencing. *Nat Biotechnol.* 2008; 26: 1135–1145. <https://doi.org/10.1038/nbt1486> PMID: 18846087
60. Galan M, Guivier E, Caraux G, Charbonnel N, Cosson JF. A 454 multiplex sequencing method for rapid and reliable genotyping of highly polymorphic genes in large-scale studies. *BMC Genomics.* 2010; 11: 296. <https://doi.org/10.1186/1471-2164-11-296> PMID: 20459828
61. Librado P, Rozas J. DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. *Bioinformatics.* 2009; 25: 1451–1452. <https://doi.org/10.1093/bioinformatics/btp187> PMID: 19346325
62. Tajima F. Statistical analysis of DNA polymorphism. *Japanese J Genet.* 1993; 68: 567–595. <https://doi.org/10.1266/jjg.68.567>
63. Brown JH, Jardetzky T, Saper MA, Samraoui B, Bjorkman PJ, Wiley DC. A hypothetical model of the foreign antigen binding site of class II histocompatibility molecules. *Nature.* 1988; 332: 845–850. <https://doi.org/10.1038/332845a0> PMID: 3258651
64. Pond SLK, Muse S V. HyPhy: Hypothesis Testing Using Phylogenies. *Statistical Methods in Molecular Evolution.* Springer-Verlag; 2005. pp. 125–181.
65. Frankham R, Ballou JD, Briscoe DA, McInnes KH. *Introduction to Conservation Genetics.* Introduction to Conservation Genetics. 2002. <https://doi.org/10.1017/cbo9780511808999>
66. Kumar S, Tamura K, Jakobsen IB, Nei M. MEGA2: Molecular evolutionary genetics analysis software. *Bioinformatics.* 2002; 17: 1244–1245. <https://doi.org/10.1093/bioinformatics/17.12.1244>
67. Brown JH, Jardetzky TS, Gorga JC, Stern LJ, Urban RG, Strominger JL, et al. Three-dimensional structure of the human class II histocompatibility antigen HLA-DR1. *Nature.* 1993; 364: 33–39. <https://doi.org/10.1038/364033a0> PMID: 8316295
68. Wetton JH, Carter RE, Parkin DT, Walters D. Demographic study of a wild house sparrow population by DNA fingerprinting. *Nature.* 1987; 327: 147–149. <https://doi.org/10.1038/327147a0> PMID: 3574474
69. Manly BFJ. *Randomization, Bootstrap and Monte Carlo Methods in Biology, Third Edition.* Randomization, Bootstrap and Monte Carlo Methods in Biology, Third Edition. 2018.
70. Hoover B, Nevitt G. Modeling the importance of sample size in relation to error in MHC-based mate-choice studies on natural populations. *Integrative and Comparative Biology.* Oxford University Press; 2016. pp. 925–933. <https://doi.org/10.1093/icb/icw105>
71. 3.6.3 RDCT. A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. Vienna, Austria: Foundation for Statistical Computing; 2020. p. <http://www.r-project.org>
72. Walsh HE, Friesen VL. A comparison of intraspecific patterns of DNA sequence variation in mitochondrial DNA,  $\alpha$ -Enolase, and MHC Class II B Loci in Auklets (Charadriiformes: Alcidae). *J Mol Evol.* 2003; 57: 681–693. <https://doi.org/10.1007/s00239-003-2518-2>
73. Vásquez-Carrillo C, Friesen V, Hall L, Peery MZ. Variation in MHC class II B genes in marbled murrelets: Implications for delineating conservation units. *Anim Conserv.* 2014; 17: 244–255. <https://doi.org/10.1111/acv.12089>
74. Miller HC, Lambert DM. Genetic drift outweighs balancing selection in shaping post-bottleneck major histocompatibility complex variation in New Zealand robins (Petroicidae). *Mol Ecol.* 2004; 13: 3709–3721. <https://doi.org/10.1111/j.1365-294X.2004.02368.x> PMID: 15548285
75. Fijarczyk A, Babik W. Detecting balancing selection in genomes: Limits and prospects. *Mol Ecol.* 2015; 24: 3529–3545. <https://doi.org/10.1111/mec.13226> PMID: 25943689
76. Hess CM, Edwards S V. The evolution of the major histocompatibility complex in birds. *Bioscience.* 2002; 52: 423–431. [https://doi.org/10.1641/0006-3568\(2002\)052\[0423:TEOTMH\]2.0.CO;2](https://doi.org/10.1641/0006-3568(2002)052[0423:TEOTMH]2.0.CO;2)
77. Biedrzycka A, O'Connor E, Sebastian A, Migalska M, Radwan J, Zając T, et al. Extreme MHC class I diversity in the sedge warbler (*Acrocephalus schoenobaenus*): Selection patterns and allelic divergence suggest that different genes have different functions. *BMC Evol Biol.* 2017; 17: 1–12. <https://doi.org/10.1186/s12862-017-0997-9>
78. Sepil I, Moghadam HK, Huchard E, Sheldon BC. Characterization and 454 pyrosequencing of Major Histocompatibility Complex class I genes in the great tit reveal complexity in a passerine system. *BMC Evol Biol.* 2012; 12: 1–19. <https://doi.org/10.1186/1471-2148-12-68>

79. Alcaide M, Muñoz J, Martínez-de la Puente J, Soriguier R, Figuerola J. Extraordinary MHC class II B diversity in a non-passerine, wild bird: The Eurasian Coot *Fulica atra* (Aves: Rallidae). *Ecol Evol.* 2014; 4: 688–698. <https://doi.org/10.1002/ece3.974> PMID: 24683452
80. Gillingham MAF, Béchet A, Courtiol A, Rendón-Martos M, Amat JA, Samraoui B, et al. Very high MHC Class IIB diversity without spatial differentiation in the mediterranean population of greater Flamingos. *BMC Evol Biol.* 2017; 17: 1–10. <https://doi.org/10.1186/s12862-017-0905-3>
81. Alcaide M, Edwards S V., Negro JJ, Serrano D, Tella JL. Extensive polymorphism and geographical variation at a positively selected MHC class II B gene of the lesser kestrel (*Falco naumanni*). *Mol Ecol.* 2008; 17: 2652–2665. <https://doi.org/10.1111/j.1365-294X.2008.03791.x>
82. Wojczulanis-Jakubas K, Svoboda A, Kruszewicz A, Johnsen A. No evidence of blood parasites in little auks (*Alle alle*) breeding on svalbard. *J Wildl Dis.* 2010; 46: 574–578. <https://doi.org/10.7589/0090-3558-46.2.574>
83. Watanabe J, Koizumi A, Nakagawa R, Takahashi K, Tanaka T, Matsuoka H. Seabirds (Aves) from the Pleistocene Kazusa and Shimosa groups, central Japan. *J Vertebr Paleontol.* 2019; 39: 1–20. <https://doi.org/10.1080/02724634.2019.1697277>
84. Wojczulanis-Jakubas K, Kilikowska A, Harding AMA, Jakubas D, Karnovsky NJ, Steen H, et al. Weak population genetic differentiation in the most numerous Arctic seabird, the little auk. *Polar Biol.* 2014; 37: 621–630. <https://doi.org/10.1007/s00300-014-1462-5>
85. Dearborn DC, Gager AB, Gilmour ME, McArthur AG, Hinerfeld DA, Mauck RA. Non-neutral evolution and reciprocal monophyly of two expressed MHC class II B genes in Leach's storm-petrel. *Immunogenetics.* 2015; 67: 111–123. <https://doi.org/10.1007/s00251-014-0813-2> PMID: 25416539
86. Fawcett JA, Innan H. Neutral and non-neutral evolution of duplicated genes with gene conversion. *Genes (Basel).* 2011; 2: 191–209. <https://doi.org/10.3390/genes2010191> PMID: 24710144
87. Sutton JT, Nakagawa S, Robertson BC, Jamieson IG. Disentangling the roles of natural selection and genetic drift in shaping variation at MHC immunity genes. *Mol Ecol.* 2011; 20: 4408–4420. <https://doi.org/10.1111/j.1365-294X.2011.05292.x> PMID: 21981032
88. Pierrney SB, Oliver MK. The evolutionary ecology of the major histocompatibility complex. *Heredity (Edinb);* 2006. pp. 7–21. <https://doi.org/10.1038/sj.hdy.6800724> PMID: 16094301
89. Gangoso L, Alcaide M, Grande JM, Muñoz J, Talbot SL, Sonsthagen SA, et al. Colonizing the world in spite of reduced MHC variation. *J Evol Biol.* 2012; 25: 1438–1447. <https://doi.org/10.1111/j.1420-9101.2012.02529.x> PMID: 22686489
90. Cloutier A, Mills JA, Baker AJ. Characterization and locus-specific typing of MHC class I genes in the red-billed gull (*Larus scopulinus*) provides evidence for major, minor, and nonclassical loci. *Immunogenetics.* 2011; 63: 377–394. <https://doi.org/10.1007/s00251-011-0516-x>
91. Burri R, Hirzel HN, Salamin N, Roulin A, Fumagalli L. Evolutionary patterns of MHC class II B in owls and their implications for the understanding of avian MHC evolution. *Mol Biol Evol.* 2008; 25: 1180–1191. <https://doi.org/10.1093/molbev/msn065> PMID: 18359775
92. Kubinak J, Nelson A, Ruff J, Potts W. Trade-offs limiting MHC heterozygosity. In: Demas G, Nelson R, editors. *Ecoimmunology.* Oxford University Press; 2012. pp. 225–258.
93. Doherty PC, Zinkernagel RM. A Biological Role for the Major Histocompatibility Antigens. *Lancet.* 1975; 305: 1406–1409. [https://doi.org/10.1016/s0140-6736\(75\)92610-0](https://doi.org/10.1016/s0140-6736(75)92610-0) PMID: 49564
94. Tregenza T, Wedell N. Genetic compatibility, mate choice and patterns of parentage: Invited review. *Molecular Ecology.* *Mol Ecol;* 2000. pp. 1013–1027. <https://doi.org/10.1046/j.1365-294x.2000.00964.x> PMID: 10964221
95. Jennions MD, Petrie M. Why do females mate multiply? A review of the genetic benefits. *Biol Rev.* 2000; 75: 21–64. <https://doi.org/10.1017/s0006323199005423> PMID: 10740892
96. Baratti M, Ammannati M, Magnelli C, Massolo A, Dessi-Fulgheri F. Are large wattles related to particular MHC genotypes in the male pheasant? *Genetica.* 2010; 138: 657–665. <https://doi.org/10.1007/s10709-010-9440-5> PMID: 20145977
97. Wang D, Forstmeier W, Valcu M, Dingemanse NJ, Bulla M, Both C, et al. Scrutinizing assortative mating in birds. *PLoS Biol.* 2019; 17. <https://doi.org/10.1371/journal.pbio.3000156> PMID: 30789896
98. Kidawa D, Jakubas D, Wojczulanis-Jakubas K, Iliszko L, Stempniewicz L. The effects of loggers on the foraging effort and chick-rearing ability of parent little auks. *Polar Biol.* 2012; 35: 909–917. <https://doi.org/10.1007/s00300-011-1136-5>
99. Kidawa D, Jakubas D, Wojczulanis-Jakubas K, Stempniewicz L, Trudnowska E, Boehnke R, et al. Parental efforts of an Arctic seabird, the little auk *Alle alle*, under variable foraging conditions. *Mar Biol Res.* 2015; 11: 349–360. <https://doi.org/10.1080/17451000.2014.940974>

100. Wojczulanis-Jakubas K, Araya-Salas M, Jakubas D. Seabird parents provision their chick in a coordinated manner. PLoS One. 2018; 13: e0189969. <https://doi.org/10.1371/journal.pone.0189969> PMID: [29320525](https://pubmed.ncbi.nlm.nih.gov/29320525/)
101. Vihtakari M. ggOceanMaps: Plot data on oceanographic maps using “ggplot2”. R package version 1.0.9. 2021. <https://cran.r-project.org/package=ggOceanMaps>
102. Yu G. scatterpie: scatter pie plot. R package version 4.0.2. 2020. [cran.r-project.org/package=scatterpie](https://cran.r-project.org/package=scatterpie)