www.nature.com/hdy

the geneticssociety

ARTICLE

Genetic variation at innate and adaptive immune genes – contrasting patterns of differentiation and local adaptation in a wild gull

Piotr Minias[1]✉, Patrycja Podlaszczuk[1], Piotr Indykiewicz[2], Mateusz Ledwoń[3], Jacek Nowakowski[4], Amelia Chyb[1] and Tomasz Janiszewski[1]

Immunogenetic variation in natural vertebrate populations is expected to respond to spatial and temporal fluctuations in pathogen assemblages. While spatial heterogeneity in pathogen-driven selection enhances local immunogenetic adaptations and population divergence, different immune genes may yield contrasting responses to the environment. Here, we investigated population differentiation at the key pathogen recognition genes of the innate and adaptive immune system in a colonial bird species, the black-headed gull *Chroicocephalus ridibundus*. We assessed genetic variation at three toll-like receptor (TLR) genes (innate immunity) and the major histocompatibility complex (MHC) class I and II genes (adaptive immunity) in gulls from seven colonies scattered across Poland. As expected, we found much greater polymorphism at the MHC than TLRs. Population differentiation at the MHC class II, but not MHC-I, was significantly stronger than at neutral microsatellite loci, suggesting local adaptation. This could reflect spatial variation in the composition of extracellular parasite communities (e.g., helminths), possibly driven by sharp differences in habitat structure between colonies. Despite contrasting patterns of population differentiation, both MHC classes showed similar regimes of diversifying selection. Some significant population differentiation was also observed at TLRs, suggesting that innate immune receptors may respond to fine-scale spatial variation in pathogen pressure, although this pattern could have been enhanced by drift. Our results suggested that local adaptation at the pathogen recognition immune genes can be maintained at relatively small or moderate spatial scales in species with high dispersal potential and they highlighted the complexity of immunogenetic responses of animals to heterogeneous environments.

Heredity (2023) 131:282–291; https://doi.org/10.1038/s41437-023-00645-2

INTRODUCTION

Host-pathogen interactions and their consequences for population dynamics and evolution have been among the central interests of biologists for decades. Pathogen-driven selection is now considered to constitute one of the major selective pressures on their hosts and it often disfavours genetic variants which are common in the populations, while favouring novel variants which improve resistance (Radwan et al. 2020). These novel variants can increase in frequency until fixation (positive selection) or can be maintained at fluctuating frequencies (balancing selection) (Lighten et al. 2017). In general, different forms of balancing selection (negative frequency-dependent, overdominant, and fluctuating selection) maintain polymorphism within and across populations, but they may also produce contrasting patterns of between-population divergence. Under the spatially homogeneous regime of balancing selection, population differentiation should be reduced compared to neutral expectations, as similar pools of alleles should be selected for and maintained in different populations (Herdegen-Radwan et al. 2021). In contrast, spatial heterogeneity in pathogen communities should promote increased population differentiation and local adaptations, as different pools of alleles should be favoured by selection at different points in space (Spurgin and Richardson 2010). Similar patterns of local adaptation may be observed when directional positive selection varies in space, leading to the fixation of different alleles in different populations (diversifying selection) (Herdegen-Radwan et al. 2021). However, population differentiation may also be driven by neutral processes, such as genetic drift, producing the patterns of isolation by distance (IBD) (Spurgin et al. 2014).

In general, immune genes are among the ones that show the strongest signature of diversifying and balancing selection, as revealed by genomic scans in humans (Andrés et al. 2009) and other vertebrates (Shultz and Sackton 2019). Among a variety of immune and immune-related genes, receptor genes and genes encoding recognition proteins responsible for direct interactions with foreign antigens are most often identified as targets of diversifying or balancing selection (Waterhouse et al. 2007;

[1]Department of Biodiversity Studies and Bioeducation, Faculty of Biology and Environmental Protection, University of Łódź, Banacha 1/3, 90-237, Łódź, Poland. [2]Department of Biology and Animal Environment, Faculty of Animal Breeding and Biology, Bydgoszcz University of Science and Technology, Mazowiecka 28, 85-084, Bydgoszcz, Poland. [3]Institute of Systematics and Evolution of Animals, Polish Academy of Sciences, Sławkowska 17, 31-016, Kraków, Poland. [4]Department of Ecology and Environmental Protection, Faculty of Biology and Biotechnology, University of Warmia and Mazury in Olsztyn, Plac Łódzki 3, 10-727, Olsztyn, Poland. Associate editor: Gerald Heckel. ✉email: pminias@op.pl

Received: 4 May 2023 Revised: 15 July 2023 Accepted: 30 July 2023
Published online: 8 August 2023

Shultz and Sackton 2019). Although different immunological pathways may rely on different pathogen recognition proteins, toll-like receptors (TLRs) and the major histocompatibility complex (MHC) form the key pathogen recognition components of the innate and adaptive vertebrate immune system, respectively. TLRs represent a family of evolutionarily conserved receptors responsible for the recognition of the pathogen-associated molecular patterns (PAMPs), which are structurally similar across a broad spectrum of pathogens (e.g., lipopolysaccharide or flagellin). As such, they show a relatively weak signature of balancing selection, although some degree of polymorphism may be selectively maintained within populations under specific conditions (e.g., Quéméré et al. 2015). There is, however, more evidence for diversifying selection acting at TLRs at the inter-specific level, leading to the fixation of different allelic variants in different taxa (Velová et al. 2018). The signature of this diversifying selective pressure accumulates in the extracellular TLR domains with leucine-rich repeats (LRRs), being directly responsible for PAMP recognition (Yang et al. 2021).

In contrast to TLRs, recognition functions of MHC proteins may be targeted at specific antigens produced intracellularly (MHC class I, MHC-I) or derived from phagocytosis in the extracellular space (MHC class II, MHC-II) (but see Joffre et al. 2012 for antigen cross-presentation). Antigens directly bind to a short peptide-binding region (PBR) of MHC molecules and interactions of MHC-peptide-complexes with T-cell receptors allow for an effective recognition of self and non-self (pathogen-driven) antigens, which, in the latter case, initiates a cascade of immune responses (Janeway et al. 2001). Despite the presence of generalist promiscuous MHC alleles at some loci, MHC molecules often show high specificity in antigen recognition and, thus, are hyper-polymorphic (Kaufman 2020). For example, tens of thousands MHC allelic variants were described globally in humans (Barker et al. 2022), but extreme MHC polymorphism has also been reported for other vertebrate species (Alcaide et al. 2014; Gerdol et al. 2019). This extraordinary variation is maintained by synergistic action of different forms of balancing selection (Spurgin and Richardson 2010), being facilitated by an extensive gene duplication which increases the number of alleles expressed within an individual (Bentkowski and Radwan 2019) and gene conversion occurring within and across duplicated loci (Spurgin et al. 2011). Although MHC and TLR genes share general functional similarities in terms of antigen recognition, their evolutionary trajectories are strikingly different and, thus, they may differently respond to pathogen-driven selection and show different patterns of population differentiation. While population differentiation at the MHC and TLRs has already been investigated in non-model bird species (e.g., Bateson et al. 2015; Kohyama et al. 2015; Levy et al. 2020), research comparing these patterns between the two gene families is generally lacking.

Here, we hypothesised that innate and adaptive pathogen recognition genes may respond differently to spatial variation in pathogen-driven selection, producing contrasting patterns of population divergence. To test this hypothesis, we assessed genetic variation at both TLR (TLR1LB, TLR3, and TLR4) and MHC (MHC-I and MHC-II) genes in a colonial bird species, the black-headed gull *Chroicocephalus ridibundus*. Black-headed gulls breed across Europe and Asia, using a wide range of inland habitats for nesting, i.e., flooded wetlands, lakes, marshes, or river deltas. Despite high potential for long-distance dispersal in gulls, their optimal habitats have patchy distribution and their colonies are often scattered in space, which favours site fidelity, constrains dispersal, and reduces gene flow (Péron et al. 2010). Thus, we expected that strong habitat variation between colonies may cause spatial variation in pathogen-driven selection landscapes, facilitating local immunogenetic adaptations in gulls. Specifically, we expected stronger population differentiation at both MHC and TLRs compared to neutral microsatellite markers, reflecting the key role of spatially varying selective regimes in shaping immunogenetic variation. However, because of weaker evolutionary conservedness and greater polymorphism enhancing responsiveness to local environmental or ecological conditions, we also expected stronger population differentiation at the MHC than TLRs.

MATERIALS AND METHODS

Study populations and general field procedures

Fieldwork was conducted in 2018 in seven black-headed gull colonies distributed from north-eastern to southern Poland (Fig. 1). All the colonies were located on natural lakes or artificial reservoirs (including dam reservoirs, fish ponds, and open pit lakes), but the structure and composition of surrounding landscape considerably varied between

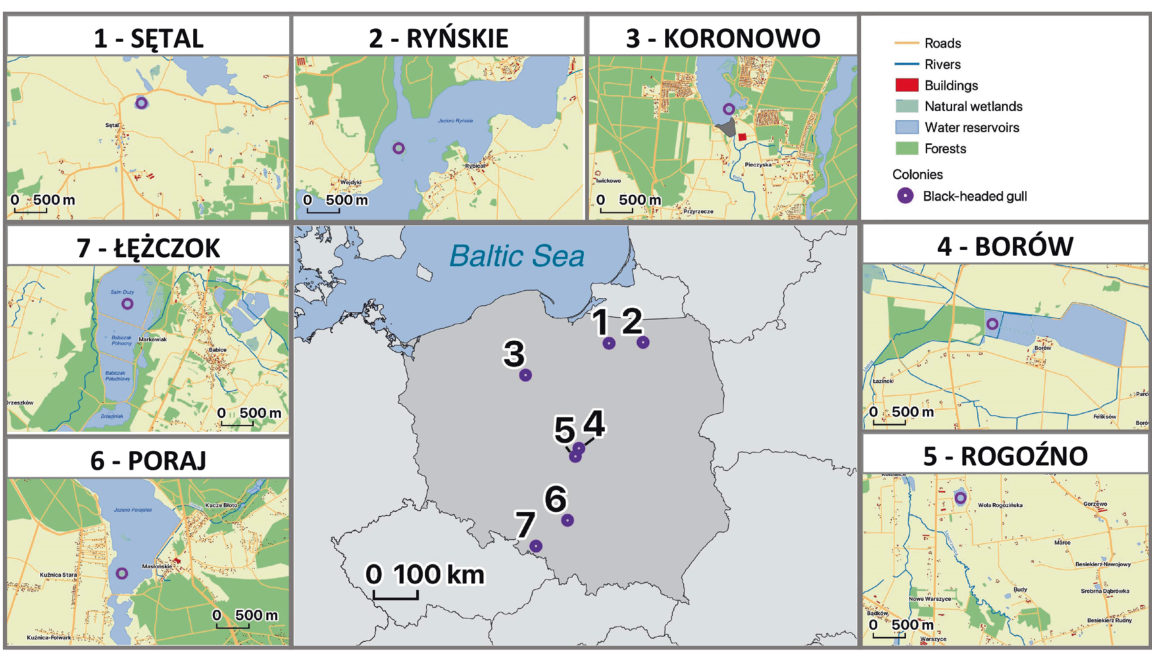

Fig. 1 Study area. Location of black-headed gull populations sampled in this study.

locations. A comparison of land cover categories within a radius of 12 km from each colony (i.e., the maximum average distance of foraging flights recorded in black-headed gulls in Poland; Jakubas et al. 2020) revealed large between-colony variation in the share of artificial areas (2–12%), agricultural areas (42–91%), forests (6–41%), and water bodies (1–12%) (Supplementary Table S1). For example, Koronowo and Poraj showed the highest share of forests and the lowest share of agricultural areas (the opposite pattern found in Borów), while Łężczok and Poraj had the highest share of artificial surfaces and the lowest share of water bodies (the opposite pattern found in Koronowo and Ryńskie) (Supplementary Table S1). The estimated size of our study colonies ranged from 100 to 3000 breeding pairs. Within each colony, we captured 20 adult individuals using nest traps or mist nets during the incubation or chick-rearing period (25 April and 15 June). Each bird was marked with a metal and plastic ring to avoid recaptures.

Blood sampling, DNA extraction, and sex identification

At capture, we collected blood samples from each bird by puncturing the ulnar vein with a disposable needle. Ca. 20 µL of blood was collected into heparinized capillary tubes and transferred to 96% ethanol. All the samples were stored at 4 °C until analysis. Genomic DNA was extracted using Bio-Trace DNA Purification Kit (EURx, Gdańsk, Poland) according to the kit protocol. To identify the sex of captured birds, we used a molecular approach developed by Fridolfsson and Ellegren (1999). Since male and female black-headed gulls participate in incubation and parental care, we expected a roughly equal sex ratio within our sample. Consistent with this expectation, we did not find any significant deviations of sex ratio from parity ($G = 2.39$, df $= 1$, $P = 0.122$), indicating no sex-related biases in our data.

MHC sequencing

Sequencing of MHC-I and MHC-II genes was specifically targeted at the PBR, which is directly responsible for the recognition of foreign antigens. The PBR comprises two different domains: α1 and α2 domains coded by exons 2 and 3 of the same MHC-I gene or α1 and β1 domains coded by exons 2 of MHC-IIA and MHC-IIB genes. Here, we genotyped a single PBR exon per MHC class (MHC-I exon 3 and MHC-IIB exon 2), as these exons are primarily targeted in the avian MHC research and their polymorphism was shown to be well representative for the polymorphism of the entire PBR (Minias et al. 2018; Dearborn et al. 2022). Because MHC genes show high duplication rates (Bentkowski and Radwan 2019), the development of locus-specific primers is usually unfeasible in non-model organisms (Babik 2010). Thus, to amplify MHC-I exon 3 and MHC-IIB exon 2 we took advantage of multi-locus primers originally developed for two other larid species, black-legged kittiwake *Rissa tridactyla* (Leclaire et al. 2014) and common tern *Sterna hirundo* (Minias et al. 2021a). Since no genomic resources were available for the black-headed gull, we examined primer specificity against sequences available for closely related congeneric taxa, i.e., the red-billed gull *Chroicocephalus novaehollandiae scopulinus* for MHC-I (GenBank no. HM008713-16) and the brown-hooded gull *Chroicocephalus maculipennis* for MHC-II (GenBank assembly no. GCA_013397655.1). Based on these comparisons, we slightly modified original MHC-I primers (SHMHC1_ex3_F and SHMHC1_ex3_R from Minias et al. 2021a), which bind conserved flanking intronic regions of exon 3 and amplify the entire exon sequence (276 bp). The final primer sequences used for MHC-I exon 3 amplifications were BHG_MHC1_ex3_F (5′-CCGGCTCCTGRTYGYGTT-3′) and BHG_MHC1_ex3_R (5′-CGTCTCGCCCTCACCTTT-3′). To amplify MHC-IIB exon 2 we used an original forward primer KWMHC2_ex2_F (5′-GCACGAGCAGGGTATTTCCA-3′) from Leclaire et al. (2014), as it showed no mismatches with brown-hooded gull sequences, and a modified reverse primer BHG_MHC2_ex2_R (5′-AAACGTTCTGCCACGCACTC-3′). The primers bind intron 1 – exon 2 junction and a conserved intron 2 region, amplifying most of the MHC-IIB exon 2 sequence (258 out of 270 bp).

In order to prepare multi-locus MHC amplicons for Illumina MiSeq sequencing, we fused our forward and reverse MHC primers with unique 7-bp barcodes to identify samples and Illumina Nextera Transposase adaptor sequences (Illumina Corp., San Diego, CA, USA). All PCR amplifications followed protocols described for MHC-I and MHC-II in Minias et al. (2021a). We successfully obtained MHC-I and MHC-II amplicons for all captured black-headed gulls ($n = 140$ individuals), as confirmed with electrophoresis on agarose gels. We also prepared independent replicate amplicons from 15 randomly selected samples to quantify the reproducibility of allele calling. All amplicons were purified and we used NEBNext DNA Library Prep Master Mix Set for Illumina (New England Biolabs, Ipswich, MA, USA) to prepare two separate (MHC-I and MHC-II) Illumina MiSeq libraries by pooling samples in equimolar quantities. Both libraries were sequenced using the 2 × 250-bp paired-end Illumina MiSeq platform.

MHC allele calling

We used Amplicon Sequencing Analysis Tools (AmpliSAT) webserver (Sebastian et al. 2016) to process raw Illumina data and identify true MHC alleles based on the relative variant frequencies within an amplicon. Our analyses followed four key steps, i.e., merging, demultiplexing, clustering, and filtering. In the first step, we merged pair-ended Illumina reads using the FLASH algorithm implemented in AmpliMERGE tool (default settings). In the second step, AmpliSAS tool was used to demultiplex merged reads, i.e., to assign them to the samples of origin based on unique barcodes. Next, we used AmpliSAS to identify artefact reads resulting from genotyping errors, identify putative true alleles from which these artefacts were derived, and cluster them together to obtain reliable estimates of allele frequencies within each amplicon (Sebastian et al. 2016). For this purpose, all read variants within each amplicon were ordered by depth and putative artefact variants (representing high-throughput sequencing errors) were identified based on their similarity to the dominant (highest depth) read variants. In this step we used default criteria recommended by Biedrzycka et al. (2017) for Illumina sequencing data, i.e., artefact read variants were identified based on the 1% substitution error and 0.001% indel error thresholds, unless they showed >25% depth of the dominant variant (in which case they were identified as putative true alleles and formed new clusters). In the last stage (filtering), AmpliSAS was used to identify and remove putative chimeric sequences (based on modified algorithms developed by Edgar et al. 2011), as well as to discard remaining low-frequency variants (<3% per amplicon). The minimum amplicon depth in our dataset was 1181 reads, while the maximum amplicon depth was limited to 5000 reads by AmpliSAS performance constraints.

Prior to the processing, the average amplicon depth was 4873.8 ± 31.8 [SE] and 3275.1 ± 45.3 [SE] reads per sample for MHC-I and MHC-II, respectively. The average amplicon depth after filtering (across all alleles) was 4018.3 ± 52.5 [SE] reads per sample for MHC-I and 2656.6 ± 39.5 [SE] reads per sample for MHC-II. There was no significant correlation between amplicon depth and the number of alleles at either MHC-I ($P = 0.82$) or MHC-II ($P = 0.27$), indicating that sequencing coverage was sufficient for reliable allele calling. Processing of raw sequencing data from 15 technical replicates (independent amplicons) revealed high reproducibility of MHC-I (92.7%) and MHC-II (92.6%) allele calling, which met the expectation for multi-locus MHC data obtained with Illumina approach (Biedrzycka et al. 2017). All putative true alleles were aligned in Geneious v10.0.5 (Biomatters Ltd., Auckland, New Zealand) and intron codons were removed, retaining only exonic regions.

TLR sequencing

We used conserved primers developed by Alcaide and Edwards (2011) to test amplifications of all ten TLR genes in the black-headed gull, as these primers specifically target TLR extracellular domains with LRRs, being directly responsible for PAMP recognition. While most of genes showed successful amplifications, we ultimately focused on three representative loci targeting a broad range of pathogen-derived antigens, including triacyl lipopeptides and other ligands from mycobacteria, gram-negative bacteria, and protozoans (TLR1LB and TLR4), as well as viral double-stranded RNA (TLR3) (Quesniaux et al. 2004; Gazzinelli and Denkers 2006; Chang 2010). All PCR amplifications followed original protocols from Alcaide and Edwards (2011) and PCR products were sequenced in forward and reverse directions (Sanger method). Short-read next-generation approaches (e.g., Illumina MiSeq) could not be used for TLR genotyping, because of excessive sequence length (846–1128 bp per locus, Table 1). The application of Sanger sequencing for TLR genotyping was also facilitated by the availability of single-locus primers (in contrast to multi-locus primers used for MHC genotyping). To reduce the costs of bidirectional Sanger sequencing, we randomly selected 10 samples per colony for genotyping. Our analyses indicated that this subsampling approach allowed us to characterise most of the standing variation across all three TLR genes and was unlikely to introduce any major bias in the analyses of population differentiation (details in Supplementary Methods in Appendix 1). All TLR sequences were aligned and trimmed to equal lengths using Geneious v10.0.5. In total, 72% codons from our alignments (across all three loci) were recognised as forming extracellular LRR domains and we genotyped the entire (TLR1LB) or most (ca. 67 and 52% at TLR3

Table 1. Polymorphism of MHC and TLR genes in the black-headed gull, as measured with the number of allelic variants (N_A; number of unique amino acid variants provided in the parentheses), number of segregating sites (S), total number of mutations (η), and nucleotide diversity (π).

Gene	Exon	Length (bp)	N_{IND}	N_A	Sequence polymorphism		
					S	η	π
MHC-I	3	273	140	224 (166)	102	127	0.0797
MHC-II	2	258	140	77 (76)	106	149	0.1188
TLR1LB	1	1044	70	25 (5)	20	20	0.0025
TLR3	4	1128	70	24 (7)	17	17	0.0022
TLR4	3	846	70	8 (3)	8	8	0.0004

and TLR4, respectively) of the LRR region (as determined based on the domestic chicken and feral pigeon *Columba livia* TLR structure; Temperley et al. 2008; Shapiro et al. 2013). The most conserved TLR motifs (e.g., the cytoplasmic toll/interleukin 1 resistance domain) were not covered by sequencing. Individual sequences of each TLR gene were phased into haplotypes using the PHASE algorithm (Stephens and Donnelly 2003) with a burn-in of 1000 iterations, 1000 final iterations and a thinning interval of ten, as implemented in DnaSP v6.10.03 (Rozas et al. 2017). Phasing was conducted using the entire alignments (across all populations) and its reliability was enhanced by considerable frequencies of homozygous genotypes and single-SNP genotypes recorded at each locus (TLR1LB: 30%, TLR3: 36%, TLR4: 91%). Overall, 53% of phased haplotypes matched TLR haplotypes previously described for the black-headed gull (Genbank accession nos: MN995080-MN995216), further supporting the reliability of our approach.

Neutral genetic variation

To quantify neutral genetic variation, we used a panel of ten microsatellite loci. All the markers were originally developed for two other larid species, the black-legged kittiwake (K6, K16, K31, K32; Tirard et al. 2002) and red-billed gull (RBG13, RBG18, RBG20, RBG27, RBG28, RBG29; Given et al. 2002), but they were previously shown to successfully cross-amplify in the black-headed gull (Indykiewicz et al. 2018). Amplification protocols followed Indykiewicz et al. (2018). All forward primers were labelled with 6-FAM fluorescent dye. Fragment size analysis was conducted with ABI 3730XL capillary sequencer (Applied Biosystems, Foster City, CA, USA) and allele sizes were scored against GeneScan TM 600 LIZ Standard (Applied Biosystems) in Geneious v10.0.5. Repeatability of allele scoring was very high (97.7%), as assessed based on genotyping of 65 independent amplicons. We found no evidence of genotyping errors due to null alleles, stutter peaks or large allele dropout, as assessed with MicroChecker v2.2.3 (Van Oosterhout et al. 2004). Likewise, we found no evidence for linkage disequilibrium between any pairs of markers (tested in Fstat v2.9.3; Goudet 1995) and for departures from Hardy-Weinberg equilibrium at any locus in any population (tested in Arlequin v1.3.5.2; Excoffier and Lischer 2010). Polymorphism of markers ranged from 7 to 25 alleles per locus with 0.45–0.91 observed heterozygosity across all populations.

Immune gene polymorphism

Polymorphism of immune genes (MHC and TLRs) was quantified as allelic richness (total number of nucleotide and amino acid allelic variants), total number of segregating sites, total number of mutations, and nucleotide diversity. All measures of polymorphism were calculated using DnaSP software. Although sequence polymorphism estimates for the MHC and TLRs provide information on the level of functional variation within each gene family, they are not directly comparable, as based on multi- and single-locus data, respectively.

Population divergence

Population divergence at both immune genes (MHC and TLRs) and neutral microsatellites was calculated based on allele frequencies and molecular distances between immune gene alleles were not incorporated in the analyses, so that the results were more comparable between different types of markers. We used two complementary unbiased estimators of differentiation, D_{est} developed by Jost (2008) and D_{est_Chao} based on the generalisation of Morisita-Horn similarity measure (Chao et al. 2008, Jost 2008). Both estimators are based on the effective number of alleles instead of heterozygosity (Jost 2008) and, thus, may be effectively calculated for multi-locus MHC data, where true heterozygosity remains unknown. Pairwise differentiation estimators for TLRs (both across and within loci) were calculated with *pairwiseTest* function in the *strataG* package (Archer et al. 2017) developed for R statistical environment (R Foundation for Statistical Computing, Vienna, Austria). Estimator significance was tested with 1000 bootstrap replicates. The *strataG* package was also used to estimate D_{est} for microsatellite markers. For the multi-locus MHC data, D_{est} values were computed using *polysat* R package (Clark and Jasieniuk 2011) originally designed to handle genotype data, where the determination of allele copy number in partial heterozygotes is ambiguous. D_{est_Chao} values for the MHC were computed using *SpadeR* R package (Chao et al. 2008), where each allele was coded as in a dominant marker (present vs. absent), which is a traditional way of dealing with allelic ambiguity in the multi-locus MHC data (Herdegen et al. 2014). The significance of D_{est_Chao} was tested with 1000 bootstrap replicates (not available for D_{est} in the *polysat* package). Differentiation at immune genes was assessed at the level of nucleotide and amino acid variants. To address the problem of multiple comparisons, raw P values for all pairwise estimators of population differentiation were corrected for the false discovery rate (FDR) (Benjamini and Hochberg 1995).

To test for differences in D_{est} between immune genes and microsatellites we used nonparametric Friedman ANOVA test for dependent samples and pairwise nonparametric comparisons based on Wilcoxon method, as implemented in the JMP v12.1 software (SAS Institute Inc., Cary, NC, USA). Correlations between immune gene (nucleotide variants) and microsatellite D_{est} values were tested using Spearman correlation coefficient. We also used Mantel test to assess IBD at immune genes and microsatellite markers. Correlations between matrices of pairwise D_{est} values and the corresponding pairwise geographical distances were quantified with 10,000 permutations in the *ade4* R package (Dray and Dufour 2007).

Population differentiation at the MHC-II (highest pairwise D_{est}, see results for details) was visualised using discriminant functions from the discriminant analysis of principal components (DAPC), as implemented in the *adegenet* R package (Jombart 2008). Population identity was used to define prior groups. The number of principal components (PCs) in DAPC analysis was selected based on the maximum α-score inferred with the *optim.a.score* function ($n = 38$ PCs retained). We retained all available discriminant functions ($n = 6$) in the analysis. Discriminant functions were also used to reassign all individuals to their prior groups (populations) and to derive their posterior membership probabilities.

Selection on MHC genes

Since the analyses revealed contrasting patterns of population differentiation between the MHC-I and MHC-II (see results for details), we aimed to test if these differences could be attributed to varying selection regimes at these genes. The signature of selection at the MHC was quantified using the relative rate of nonsynonymous (amino acid altering) nucleotide substitutions per nonsynonymous site to synonymous (silent) nucleotide substitutions per synonymous site (dN/dS ratio). In general, high dN/dS values (>1) provide evidence for positive (diversifying) selection (nonsynonymous substitutions are favoured and accumulate faster than synonymous substitutions), while low dN/dS values (<1) are characteristic for negative (purifying) selection (nonsynonymous substitutions are removed by selection and accumulate slower than synonymous substitutions). Similar rates of nonsynonymous and synonymous substitutions ($dN/dS = 1$) are consistent with neutral evolution (no apparent signature of either positive or negative selection). The dN/dS ratio was originally developed to infer selection acting on allelic variants that represent fixation events across independent lineages and this approach may yield reduced power to infer positive selection across polymorphisms segregating within a single population (Kryazhimskiy and Plotkin 2008). Despite this limitation, due to the extraordinary diversity of the MHC, *dN/dS* ratio has been widely implemented to infer intra-specific signature of selection at these genes (e.g., Promerová et al. 2013; Alcaide et al. 2014). Because of low sequence polymorphism at all three TLR loci (see results for details), we did not aim to analyse nucleotide substitution rate at these genes, as this was unlikely to yield meaningful results.

Selection inference may be biased by the presence of recombination, which alters tree topologies used to estimate substitution rates (i.e., no unique tree topology can describe the evolutionary history of the entire sequence; Anisimova et al. 2003). Since codon substitution models cannot effectively deal with the effects of recombination, we identified and removed recombinant sequences from our dataset prior to selection analyses. The recombination signal at the MHC was quantified with RDP v.4.97 software (Martin et al. 2015). Briefly, RDP4 implements several different algorithms to detect recombinant sequences and here we used seven independent approaches (BootScan, Chimaera, Genconv, Maxchi, RDP, SiScan, and 3Seq). All analyses were run using default settings and Bonferroni correction for multiple comparisons was used to infer statistical significance. A recombination event was recognised when supported by at least two independent algorithms and different recombination events were identified based on the different locations of recombination breakpoints. The recombination signal was also assessed as the number of breakpoints within 150 nucleotide (nt) window and the presence of recombination hot-spots and cold-spots was tested with the local hot/cold-spot test (1000 permutations), as implemented in RDP software. G test was used to compare the proportion of recombinant sequences between MHC-I and MHC-II.

After the removal of recombinant sequences, we used Fast Unconstrained Bayesian AppRoximation (FUBAR) approach (Murrell et al. 2013) to estimate codon-specific *dN/dS* ratios at both MHC-I and MHC-II and to identify codons under pervasive positive and negative selection (detectable across the entire allele tree). In addition, we used Mixed Effects Model of Evolution (MEME) approach (Murrell et al. 2012) to identify codons under episodic positive selection (detectable at a proportion of allele tree branches). Both FUBAR and MEME analyses were run at the Datamonkey webserver (Weaver et al. 2018) using default settings and input trees inferred from the sequence alignments. Sites with Bayesian posterior probabilities >0.90 (FUBAR) or statistical significance $P < 0.05$ (MEME) were considered to have enough support to infer the signature of selection. We also re-ran FUBAR analyses across full MHC-I and MHC-II alignments (including recombinant sequences). The estimates of codon-specific substitution rates were highly consistent between both approaches (Pearson correlation: $r = 0.96$, $P < 0.001$ for MHC-I; $r = 0.90$, $P < 0.001$ for MHC-II) indicating that our results were robust to recombination (and to the removal of recombinant sequences).

Codon-specific *dN/dS* estimates were averaged either across all sites or across 20 most positively selected sites (PSS) (as identified based on Bayes factors from FUBAR), allowing us to quantify the strength of positive selection with no interfering effects of negative selection. The location of PSS was compared to the location of sites putatively involved in ligand binding (PBR sites), as identified based on the crystallographic structure of human MHC molecules (Saper et al. 1991; Brown et al. 1993) and general distribution of selection signature in non-passerine birds (Minias et al. 2018). All values are presented as means ± SE.

RESULTS

MHC and TLR diversity

We identified a total of 224 MHC-I and 77 MHC-II allelic variants across all the populations, translating into 166 and 76 unique amino acid variants, respectively. We recorded 5–8 MHC-I and 1–8 MHC-II alleles per individual, providing evidence for the presence of at least four loci at each MHC class in the black-headed gull. The mean number of alleles per individual was 6.82 ± 0.08 (MHC-I) and 3.64 ± 0.10 (MHC-II), but there was no significant correlation between allele numbers at both MHC classes ($r = 0.01$, $P = 0.89$). The distribution of allele frequencies was similar between MHC-I and MHC-II, except for the presence of a single clearly dominant MHC-I allele (77.1% frequency across all populations) and a lack of dominant alleles at the MHC-II (max. 25% frequency across all populations) (Supplementary Fig. S1). Despite the lower allelic richness, we recorded higher nucleotide diversity at MHC-II than MHC-I (Table 1). The diversity of TLR genes, in terms of allelic richness and sequence polymorphism, was much lower compared to the MHC. We recorded 8–25 allelic variants per TLR gene, but they only translated into 3–7 unique amino acid variants (Table 1). All measures of sequence polymorphism were clearly lower for all TLR genes, when compared to the MHC (Table 1). The lowest allelic richness and sequence polymorphism was recorded at TLR4 (Table 1).

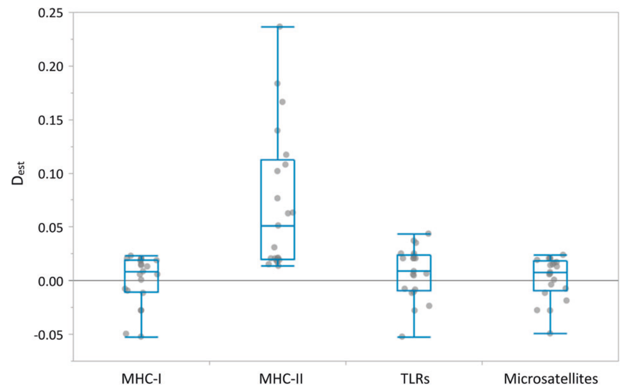

Fig. 2 Genetic population differentiation. Pairwise D_{est} between seven black-headed gull populations were estimated for different immune genes (amino acid variants) and neutral microsatellite markers. D_{est} values were block-centred for each pairwise comparison across all the markers. Median (central line), Q1 and Q3 quantiles (box) and outlier range (whiskers) are shown for each marker.

Immunogenetic divergence between populations

We found significant differences in the pairwise population differentiation (D_{est}) among different immune genes (Kruskal–Wallis: $\chi^2 = 22.49$, df = 3, $P < 0.001$ for nucleotide variants; $\chi^2 = 27.94$, df = 3, $P < 0.001$ for amino acid variants; Fig. 2). At both levels (nucleotide and amino acid), the MHC-II showed significantly greater population differentiation than other immune genes (MHC-I and TLRs) and microsatellites (all $P < 0.05$; Table 2 and Fig. 2). Population differentiation at MHC-I and TLRs did not differ from differentiation at neutral microsatellite markers (Table 2 and Fig. 2). We also found significant differences in D_{est} estimates among TLR genes, but only at the level of nucleotide variants (Kruskal–Wallis: $\chi^2 = 12.68$, df = 2, $P = 0.002$), where TLR1LB and TLR3 showed stronger differentiation than TLR4 (Supplementary Fig. S2A and Supplementary Table S2). There was no significant variation in D_{est} between TLR genes at the amino acid level (Kruskal–Wallis: $\chi^2 = 4.71$, df = 2, $P = 0.095$; Supplementary Fig. S2B). There was also no significant correlation between pairwise immune gene and microsatellite D_{est} values (all $P > 0.05$).

Although the strongest average population differentiation was found for the MHC-II, we found no evidence for significant (FDR-corrected) pairwise differentiation at either MHC-I or MHC-II (Supplementary Tables S3 and S4). There was, however, an indication for relatively strong differentiation at the MHC-II between Poraj vs. Koronowo and Rogoźno colonies (Fig. 3) and this pattern was apparent at both nucleotide and amino acid levels (Supplementary Table S4). Significant pairwise differentiation was only found for multi-locus TLR data (amino acid variants). Specifically, we found that Koronowo colony showed significant TLR differentiation from four other colonies (Supplementary Table S5) and this pattern was primarily driven by differentiation at TLR1LB gene (Supplementary Table S6). No significant differentiation was found at neutral microsatellite markers (Supplementary Table S8). No significant IBD was found at either immune genes or microsatellites, as assessed with Mantel tests (all $P > 0.05$).

Table 2. Pairwise nonparametric comparisons of D_{est} estimates between different genetic markers in the black-headed gulls.

Type of data	Genetic marker	TLR	MHC-I	MHC-II	Microsatellites
Nucleotide variants	TLR	–	0.980	**<0.001**	0.314
	MHC-I	0.03	–	**<0.001**	0.378
	MHC-II	**3.50**	**3.42**	–	**<0.001**
	Microsatellites	1.01	0.88	**4.29**	–
Amino acid variants	TLR	–	0.227	**<0.001**	0.302
	MHC-I	1.21	–	**<0.001**	0.910
	MHC-II	**3.41**	**4.48**	–	**<0.001**
	Microsatellites	1.03	0.11	**4.53**	–

Z values are shown below diagonal, while *P* values are shown above diagonal. Significant comparisons are marked in bold.

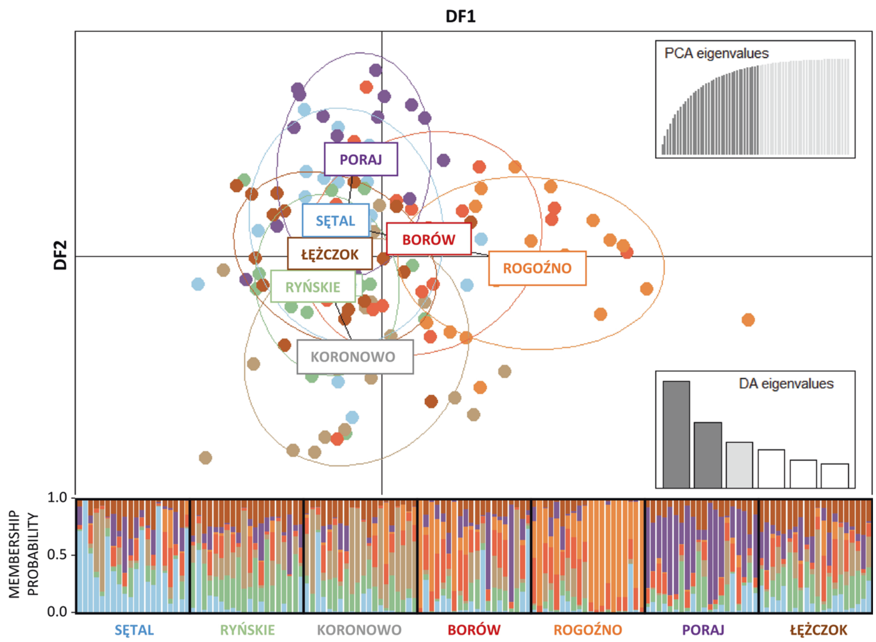

Fig. 3 Population differentiation at the MHC-II. Discriminant functions (DF1 and DF2) were calculated with the discriminant analysis of principal components (DAPC) for the MHC-II (amino acid variants) across seven black-headed gull populations (each population marked with a different colour). The reassignment of individuals into their prior groups (populations) is shown below the DAPC scatterplot. Each bar represents proportional posterior membership probability of each individual into each population.

Selection on MHC-I and MHC-II

Prior to selection analysis, we examined recombination signal at MHC-I exon 3 and MHC-II exon 2. We recorded four (MHC-I) and five (MHC-II) recombination events per exon and the number of breakpoints per 150 nt window was estimated at 3–7 for both MHC classes (Supplementary Fig. S3). We found two recombination hot-spots at MHC-I and one recombination hot-spot at MHC-II; no recombination cold-spots were recorded (Supplementary Fig. S3). The proportion of recombinant sequences was significantly larger at MHC-II (54.5%) than MHC-I (35.3%) ($G = 4.39$, df $= 1$, $P = 0.037$).

After removing recombinant sequences, we identified sites under pervasive (positive and negative) and episodic (positive) selection at the MHC. We found a similar number of sites under pervasive positive selection at both MHC-I ($n = 11$) and MHC-II ($n = 13$) (Table 1, Fig. 4 and Supplementary Fig. S4). We also found additional three (MHC-I) and two (MHC-II) sites under episodic positive selection (Fig. 4 and Supplementary Fig. S4). Despite lower allelic richness, we found slightly higher excess of nonsynonymous substitutions at the MHC-II (all codons: $dN/dS = 1.05$, PSS: $dN/dS = 8.26$) than MHC-I (all codons: $dN/dS = 0.65$, PSS: $dN/dS = 6.75$).

Compared to the MHC-I, the location of PSS at MHC-II in black-headed gulls showed better agreement with human PBR residues and PSS previously recognised in non-passerine birds. At the MHC-I we observed only moderate agreement, as 27.3% of PSS were located directly at residues corresponding to human PBR and additional 18.1% of PSS were located at residues neighbouring to human PBR (Fig. 4). At the same time, 63.6% of MHC-I PSS were located at residues previously recognised as under positive selection in non-passerine birds (Fig. 4). In contrast, all MHC-II PSS were located directly at residues corresponding or neighbouring to human PBR (69.2% and 30.8% of PSS, respectively). Most of MHC-II PSS in the black-headed gull (76.9%) were also located at residues previously recognised as PSS in non-passerine birds (Fig. 4).

DISCUSSION

In this study, we compared polymorphism and population differentiation at two key pathogen recognition gene families of the innate and adaptive immune system in the colonial black-headed gull. As expected, we found much greater polymorphism

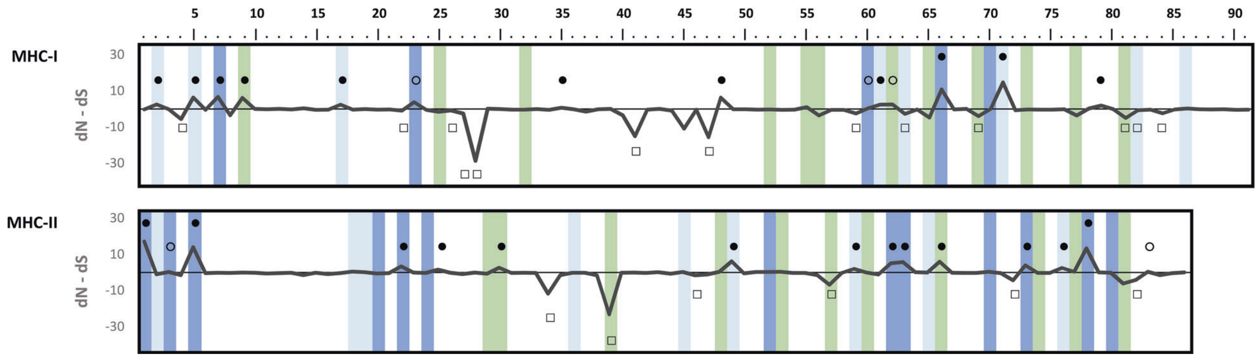

Fig. 4 Selection on MHC genes. Sites under pervasive and episodic diversifying (positive) selection are marked above the *dS-dN* curve with filled and open dots, respectively. Sites under purifying (negative) selection are marked below the *dS-dN* curve with open squares. Human peptide-binding region (PBR) sites (according to Saper et al. 1991 for MHC-I and Brown et al. 1993 for MHC-II) are marked in light green, sites previously identified as under positive selection in non-passerine birds (according to Minias et al. 2018) are marked in light blue, while overlapping sites (human PBR sites under positive selection in non-passerines) are marked in dark blue.

at the MHC (adaptive immunity) than TLRs (innate immunity), which are evolutionarily and functionally more conserved. Population divergence at MHC-II was significantly stronger than at neutral microsatellite loci, suggesting local adaptation. This could reflect spatial variation in the composition of extracellular parasite communities, possibly driven by differences in habitat structure. The significant pairwise divergence between the populations was also observed at TLRs, suggesting that innate pathogen recognition receptors may also respond to fine-scale spatial variation in pathogen pressure.

Pathogen and parasite communities may vary in space and this variation is likely driven by a multitude of environmental and ecological factors, including habitat structure and host availability (Kozakiewicz et al. 2018). Consequently, different host populations are subject to varying regimes of pathogen-driven selection and pathogen recognition genes are likely to primarily respond to these processes by local adaptation. However, any local adaptation at immune genes can emerge and be maintained under restricted gene flow (Kawecki and Ebert 2004). Otherwise, diversity of immune genes can be maintained across populations (at the meta-population level) by other forms of balancing selection, including overdominant selection, negative frequency-dependent selection, and selection fluctuating in time (Spurgin and Richardson 2010). Here, we found evidence for strong population differentiation at the MHC-II genes in the black-headed gull, suggesting local adaptation. The MHC-II codes for molecules responsible for presenting antigens derived from the processing of extracellular parasites to T cells, thus, providing a key mechanism of immune defence against macroparasites, such as helminths. Helminth infections are known to negatively impact condition and fitness in gulls and other waterbirds (Bosch et al. 2000; Amundson and Arnold 2010), although to the best of our knowledge black-headed gulls were not subject to this kind of research. However, gulls seem to be exposed to exceptionally strong selection from helminth parasites, showing one of the highest helminth species richness among all bird lineages (Gutiérrez et al. 2019). For example, black-headed gulls have been reported to host nearly two hundred helminth species with trematodes being recognised as the most species-rich group ($n = 120$ species) among the hosted helminths (Gibson et al. 2005). Furthermore, only one other larid species, the herring gull *Larus argentatus*, was reported to host a similar richness of helminth parasites, while all the other taxa from Laridae family were reported to host much fewer helminth species (mean 19.0 ± 3.8 helminth parasites reported per gull species) (Gibson et al. 2005). This exceptional macroparasite (helminth) richness in black-headed gulls may likely enhance local adaptations at MHC-II genes, especially if spatial variation in helminth communities creates a complex selective landscape across habitats. Such a situation has been reported for the red-legged partridge *Alectoris rufa*, which showed low spatial repeatability and predictability of composition and structure of parasite helminth communities at the local scale, being primarily attributed to habitat variation (Calvete et al. 2004). Similar patterns have been reported for waterbirds in Poland, e.g., the helminth fauna of great cormorants *Phalacrocorax carbo* from the brackish water habitat was far richer and strongly diverged from helminths detected in cormorants inhabiting freshwater lakes (Kanarek and Zaleśny 2014). It is possible that in such conditions, local adaptations at the MHC-II can be maintained despite the presence of gene flow between populations. Here, we found no significant population differentiation at the neutral microsatellite loci, suggesting ongoing dispersal of individuals and gene flow between our study colonies. This finding was consistent with our previous analyses of population structure in black-headed gulls, providing no support for neutral genetic differentiation between landscapes characterised by different urbanisation level (Indykiewicz et al. 2018).

In contrast to the MHC-II, we found no significant population differentiation at the MHC-I genes, which code for molecules targeting a different group of pathogens (mostly intra-cellular viruses, bacteria, and protists). In fact, pairwise estimators of population divergence were similar between MHC-I and neutral markers, suggesting no local adaptation. These contrasting patterns in population differentiation between MHC-II and MHC-I could reflect varying degrees of local spatial variation in pathogens targeted by MHC class I and II molecules. Although we lack robust quantitative data on the composition of pathogen and parasite communities across our study populations, our previous research on haemoparasite infections (primarily targeted by MHC-I) in the black-headed gull revealed a relatively homogenous infection rates (40–70%) of *Haemoproteus* and *Plasmodium* and no skewed distribution of haemoparasite molecular lineages between the colonies (Włodarczyk et al. 2022). In general, MHC-I governs immune defence against many groups of pathogens being more "internal" to the host (i.e., lacking external stages), and thus less directly affected by environmental variability (Guernier et al. 2004). Also, these taxa may more readily spread over longer distances via their hosts (Guernier et al. 2004), thus increasing spatial homogeneity in pathogen community structure. Although spatial homogeneity in pathogen pressure may preclude local adaptations, high MHC-I allelic polymorphism can still be maintained at the meta-population level. In fact, the total allelic richness across our study populations was over twice higher for MHC-I (166 amino acid variants) compared to MHC-II genes (76 amino acid variants) and the signature of positive diversifying selection was similar at both MHC classes. As this signal concentrated in the PBR of MHC molecules, it was likely to generate a huge variation in antigen recognition properties between different allelic variants. However, we also found evidence for similar recombination rates at the MHC-I and MHC-II, indicating similar contribution of gene conversion processes to the maintenance of allelic diversity. This suggests that despite contrasting spatial patterns, general mechanisms (i.e., selection and gene conversion) may act with similar strength on MHC-I and MHC-II in the black-headed gull.

Previous research has already suggested that MHC-I and MHC-II genes may differently respond to spatial heterogeneity in pathogen-driven selection. Local adaptation at the MHC-I was generally limited by large population sizes, high gene flow, and/or similar selection pressures in the Magellanic penguins *Spheniscus magellanicus* and Humboldt penguins *S. humboldti*, but IBD at the MHC-II suggested local adaptation in the latter species (Sallaberry-Pincheira et al. 2016). Similarly, MHC-II allelic repertoire was strongly associated with fine-scale habitat structure in a recently bottlenecked population of the white-tailed eagle *Haliaeetus albicilla*, whereas MHC-I showed negligible habitat-related variation in space (Minias et al. 2021b). There is also ample evidence for local adaptation at either MHC-I or MHC-II in birds, but these patterns were usually detected at relatively broad geographical scales. One of the early studies on the great snipe *Gallinago media* revealed genetic differentiation at the MHC-II between two geographical regions (Scandinavia vs. East Europe), most likely representing local adaptation to different environments (mountain habitats vs. lowland flood plains) (Ekblom et al. 2007). Local adaptations at the MHC were also directly linked to spatial distribution of pathogens, e.g., population-specific alleles provided resistance to local malaria strains in the house sparrow *Passer domesticus* (Bonneaud et al. 2006; Loiseau et al. 2011). In contrast, no population structure at the MHC-II was found within the Mediterranean region in greater flamingos *Phoenicopterus roseus*, suggesting that local adaptation was constrained by high gene flow, although allelic diversity was maintained by range-wide pathogen-driven selection (Gillingham et al. 2017), which appeared similar to the patterns we described here for the MHC-I in gulls.

Our study also provided support for population differentiation at the innate immune genes. While the average population differentiation did not differ significantly between TLRs and neutral microsatellites, we found significant multi-locus TLR divergence between specific populations. TLR genes are generally much more conserved than the MHC and, thus, are less prone to local adaptations, although some allelic polymorphism can be maintained at TLRs via balancing selection operating within and across populations (Minias and Vinkler 2022). So far, evidence for local adaptations at TLRs in natural avian populations is scant and research on the song sparrow *Melospiza melodia* provides a notable exception. Comparisons of two Canadian song sparrow populations revealed significant differences in the frequencies of four TLR loci between Ontario and British Columbia (Nelson-Flower et al. 2018), and this differentiation was stronger when compared to other innate immune genes responsible for infection clearance (beta-defensins) (Boyd et al. 2021). Both populations also differed in susceptibility to *Plasmodium* strains (Sarquis-Adamson and MacDougall-Shackleton 2016), but these could not be attributed to local protective MHC alleles, since, surprisingly, no population differentiation was found at the MHC (Slade et al. 2017). Significant differentiation at TLR3 was also found among island populations of endemic Seychelles warbler *Acrocephalus sechellensis* and this differentiation reflected different rates of temporal decline in the frequency of the single minor allele (Davies et al. 2021). Overall, these findings suggested that positive selection operating on infection-sensing innate immune genes among, rather than within populations may facilitate local adaptations to spatially variable parasite assemblages (Boyd et al. 2021). Population differentiation at TLRs was also reported for some other avian species (e.g., Xu et al. 2020), but these patterns were often attributed to drift rather than selection (Grueber et al. 2013; Gonzalez-Quevedo et al. 2015). Here, we found no evidence for IBD at both immune genes and microsatellites, suggesting limited role of drift. Although a combination of both mechanisms (selection and drift) could determine spatial patterns of population differentiation at innate immune genes, more studies are needed to draw any robust conclusions about the importance of these processes in shaping TLR polymorphism.

While our comparisons of population differentiation between different immune genes and microsatellites provided novel information on population differentiation and putative local adaptation in gulls, we recognise some limitations of our study. First, we compared genetic markers that show substantially different evolutionary trajectories and levels of polymorphism. We have also used a set of microsatellite loci as a reference for neutral evolutionary processes. Although this is a common practice in population genetic studies (e.g., Miller et al. 2010; Gonzalez-Quevedo et al. 2015), there is an increasing evidence that the evolution of tandem repeats may not be purely neutral (e.g., due to their location in genic regions or linkage with adaptive markers) (Gemayel et al. 2010; Ranathunge et al. 2022). Taking all this into account, mechanistic interpretations of our results need to be cautious. Second, our analyses were based on only three out of ten TLR genes described in birds. Although all three targeted genes showed consistently low differentiation between gull populations, we cannot exclude that other TLR loci could show contrasting patterns. Finally, population differentiation could possibly vary between TLRs and other pattern recognition receptors, e.g., RIG-I-like receptors or Nod-like receptors, and focusing on a broader spectrum of immune receptors in future research could shed new light on immunogenetic adaptations in birds.

In conclusion, our results suggested that local adaptation at the pathogen recognition immune genes can be maintained at relatively small or moderate spatial scales in species with high dispersal potential, likely due to strong variation in pathogen-driven selection and possibly reflecting sharp differences in habitat structure between breeding locations. The results also suggested that the patterns of population differentiation can markedly differ not only between innate and adaptive immune genes (TLRs and MHC), but also between genes representing the same axis of immunity, especially if they target different types of antigens (MHC- I and MHC-II). Thus, our study highlights the complexity of immunogenetic responses of animals to heterogeneous environments and indicates a need for a broader application of multi-gene approaches in future research on immune adaptations in non-model organisms.

DATA AVAILABILITY

All data used in this study are available in Dryad (https://doi.org/10.5061/dryad.15dv41p32). All novel sequences generated in this study have been deposited in GenBank (OR333996-OR334322).

REFERENCES

Alcaide M, Edwards SV (2011) Molecular evolution of the toll-like receptor multigene family in birds. Mol Biol Evol 28:1703–1715

Alcaide M, Muñoz J, Martínez-de la Puente J, Soriguer R, Figuerola J (2014) Extraordinary MHC class II B diversity in a non-passerine, wild bird: the Eurasian Coot *Fulica atra* (Aves: Rallidae). Ecol Evol 4:688–698

Amundson CL, Arnold TW (2010) Anthelmintics increase survival of American Coot (Fulica americana) chicks. Auk 127:653–659

Andrés AM, Hubisz MJ, Indap A, Torgerson DG, Degenhardt JD, Boyko AR et al. (2009) Targets of balancing selection in the human genome. Mol Biol Evol 26:2755–2764

Anisimova M, Nielsen R, Yang Z (2003) Effect of recombination on the accuracy of the likelihood method for detecting positive selection at amino acid sites. Genetics 164:1229–1236

Archer FI, Adams PE, Schneiders BB (2017) stratag: an r package for manipulating, summarizing and analysing population genetic data. Mol Ecol Res 17:5–11

Babik W (2010) Methods for MHC genotyping in non-model vertebrates. Mol Ecol Res 10:237–251

Barker DJ, Maccari G, Georgiou X, Cooper MA, Flicek P, Robinson J, Marsh SG (2022) The IPD-IMGT/HLA Database. Nucleic Acids Res 2022:gkac1011

Bateson ZW, Whittingham LA, Johnson JA, Dunn PO (2015) Contrasting patterns of selection and drift between two categories of immune genes in prairie-chickens. Mol Ecol 24:6095–6106

Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc B 57:289–300

Bentkowski P, Radwan J (2019) Evolution of major histocompatibility complex gene copy number. PLoS Comput Biol 15:e1007015

Biedrzycka A, Sebastian A, Migalska M, Westerdahl H, Radwan J (2017) Testing genotyping strategies for ultra-deep sequencing of a co-amplifying gene family: MHC class I in a passerine bird. Mol Ecol Res 17:642–655

Bonneaud C, Pérez-Tris J, Federici P, Chastel O, Sorci G (2006) Major histocompatibility alleles associated with local resistance to malaria in a passerine. Evolution 60:383–389

Bosch M, Torres J, Figuerola J (2000) A helminth community in breeding Yellow-legged Gulls (*Larus cachinnans*): pattern of association and its effect on host fitness. Can J Zool 78:777–786

Boyd RJ, Denommé MR, Grieves LA, MacDougall-Shackleton EA (2021) Stronger population differentiation at infection-sensing than infection-clearing innate immune loci in songbirds: different selective regimes for different defenses. Evolution 75:2736–2746

Brown JH, Jardetzky TS, Gorga JC, Stern LJ, Urban RG, Strominger JL, Wiley DC (1993) Three-dimensional structure of the human class II histocompatibility antigen HLA-DR1. Nature 364:33–39

Calvete C, Blanco-Aguiar JA, Virgós E, Cabezas-Díaz S, Villafuerte R (2004) Spatial variation in helminth community structure in the red-legged partridge (*Alectoris rufa* L.): effects of definitive host density. Parasitology 129:101–113

Chang ZL (2010) Important aspects of Toll-like receptors, ligands and their signaling pathways. Inflamm Res 59:791–808

Chao A, Jost L, Chiang SC, Jiang YH, Chazdon RL (2008) A two-stage probabilistic approach to multiple-community similarity indices. Biometrics 64:1178–1186

Clark LV, Jasieniuk M (2011) Polysat: an R package for polyploid microsatellite analysis. Mol Ecol Res 11:562–566

Davies CS, Taylor MI, Hammers M, Burke T, Komdeur J, Dugdale HL, Richardson DS (2021) Contemporary evolution of the innate immune receptor gene TLR3 in an isolated vertebrate population. Mol Ecol 30:2528–2542

Dearborn DC, Warren S, Hailer F (2022) Meta-analysis of major histocompatibility complex (MHC) class IIA reveals polymorphism and positive selection in many vertebrate species. Mol Ecol 31:6390–6406

Dray S, Dufour AB (2007) The ade4 package: implementing the duality diagram for ecologists. J Stat Softw 22:1–20

Edgar RC, Haas BJ, Clemente JC, Quince C, Knight R (2011) UCHIME improves sensitivity and speed of chimera detection. Bioinformatics 27:2194–2200

Ekblom R, Saether SA, Jacobsson P, Fiske P, Sahlman T, Grahn M et al. (2007) Spatial pattern of MHC class II variation in the great snipe (*Gallinago media*). Mol Ecol 16:1439–1451

Excoffier L, Lischer HE (2010) Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. Mol Ecol Res 10:564–567

Gazzinelli RT, Denkers EY (2006) Protozoan encounters with Toll-like receptor signalling pathways: implications for host parasitism. Nat Rev Immunol 6:895–906

Gerdol M, Lucente D, Buonocore F, Poerio E, Scapigliati G, Mattiucci S et al. (2019) Molecular and structural characterization of MHC class II β genes reveals high diversity in the cold-adapted icefish *Chionodraco hamatus*. Sci Rep. 9:5523

Gemayel R, Vinces MD, Legendre M, Verstrepen KJ (2010) Variable tandem repeats accelerate evolution of coding and regulatory sequences. Ann Rev Genet 44:445–477

Gibson D, Bray R, Harris E (2005) Host–parasite database of the Natural History Museum, London. https://www.nhm.ac.uk/research-curation/scientific-resources/taxonomy-systematics/host-parasites/database/index.jsp

Gillingham MA, Béchet A, Courtiol A, Rendón-Martos M, Amat JA, Samraoui B et al. (2017) Very high MHC Class IIB diversity without spatial differentiation in the mediterranean population of greater Flamingos. BMC Evol Biol 17:56

Given AD, Mills JA, Baker AJ (2002) Isolation of polymorphic microsatellite loci from the red-billed gull (*Larus novaehollandiae scopulinus*) and amplification in related species. Mol Ecol Notes 2:416–418

Gonzalez-Quevedo C, Spurgin LG, Illera JC, Richardson DS (2015) Drift, not selection, shapes toll-like receptor variation among oceanic island populations. Mol Ecol 24:5852–5863

Goudet J (1995) FSTAT (version 1.2): a computer program to calculate F-statistics. J Hered 86:485–486

Grueber CE, Wallis GP, Jamieson IG (2013) Genetic drift outweighs natural selection at toll-like receptor (TLR) immunity loci in a re-introduced population of a threatened species. Mol Ecol 22:4470–4482

Guernier V, Hochberg ME, Guégan JF (2004) Ecology drives the worldwide distribution of human diseases. PLoS Biol 2:e141

Gutiérrez JS, Piersma T, Thieltges DW (2019) Micro-and macroparasite species richness in birds: the role of host life history and ecology. J Anim Ecol 88:1226–1239

Herdegen M, Babik W, Radwan J (2014) Selective pressures on MHC class II genes in the guppy (*Poecilia reticulata*) as inferred by hierarchical analysis of population structure. J Evol Biol 27:2347–2359

Herdegen-Radwan M, Phillips KP, Babik W, Mohammed RS, Radwan J (2021) Balancing selection versus allele and supertype turnover in MHC class II genes in guppies. Heredity 126:548–560

Indykiewicz P, Podlaszczuk P, Janiszewska A, Minias P (2018) Extensive gene flow along the urban–rural gradient in a migratory colonial bird. J Avian Biol 49:e01723

Jakubas D, Indykiewicz P, Kowalski J, Iciek T, Minias P (2020) Intercolony variation in foraging flight characteristics of black-headed gulls *Chroicocephalus ridibundus* during the incubation period. Ecol Evol 10:5489–5505

Janeway JCA, Travers P, Walport M, Shlomchik MJ (2001) Immunobiology: the immune system in health and disease. Garland Science, New York

Joffre OP, Segura E, Savina A, Amigorena S (2012) Cross-presentation by dendritic cells. Nat Rev Immunol 12:557–569

Jombart T (2008) Adegenet: a R package for the multivariate analysis of genetic markers. Bioinformatics 24:1403–1405

Jost L (2008) GST and its relatives do not measure differentiation. Mol Ecol 17:4015–4026

Kanarek G, Zaleśny G (2014) Extrinsic-and intrinsic-dependent variation in component communities and patterns of aggregations in helminth parasites of great cormorant (*Phalacrocorax carbo*) from NE Poland. Parasitol Res 113:837–850

Kaufman J (2020) From chickens to humans: the importance of peptide repertoires for MHC class I alleles. Front Immunol 11:601089

Kawecki TJ, Ebert D (2004) Conceptual issues in local adaptation. Ecol Lett 7:1225–1241

Kohyama TI, Omote K, Nishida C, Takenaka T, Saito K, Fujimoto S, Masuda R (2015) Spatial and temporal variation at major histocompatibility complex class IIB genes in the endangered Blakiston's fish owl. Zoological Lett 1:13

Kozakiewicz CP, Burridge CP, Funk WC, VandeWoude S, Craft ME, Crooks KR et al. (2018) Pathogens in space: advancing understanding of pathogen dynamics and disease ecology through landscape genetics. Evol Appl 11:1763–1778

Kryazhimskiy S, Plotkin JB (2008) The population genetics of dN/dS. PLoS Genet 4:e1000304

Leclaire S, van Dongen WF, Voccia S, Merkling T, Ducamp C, Hatch SA et al. (2014) Preen secretions encode information on MHC similarity in certain sex-dyads in a monogamous seabird. Sci Rep 4:6920

Levy H, Fiddaman SR, Vianna JA, Noll D, Clucas GV, Sidhu JK et al. (2020) Evidence of pathogen-induced immunogenetic selection across the large geographic range of a wild seabird. Mol Biol Evol 37:1708–1726

Lighten J, Papadopulos AS, Mohammed RS, Ward BJ, Paterson IG, Baillie L (2017) Evolutionary genetics of immunological supertypes reveals two faces of the Red Queen. Nat Comm 8:1294

Loiseau C, Zoorob R, Robert A, Chastel O, Julliard R, Sorci G (2011) *Plasmodium relictum* infection and MHC diversity in the house sparrow (*Passer domesticus*). Proc R Soc B 278:1264–1272

Martin DP, Murrell B, Golden M, Khoosal A, Muhire B (2015) RDP4: detection and analysis of recombination patterns in virus genomes. Virus Evol 1:vev003

Miller HC, Allendorf F, Daugherty CH (2010) Genetic diversity and differentiation at MHC genes in island populations of tuatara (*Sphenodon* spp.). Mol Ecol 19:3894–3908

Minias P, Drzewińska-Chańko J, Włodarczyk R (2021a) Evolution of innate and adaptive immune genes in a non-model waterbird, the common tern. Infect Genet Evol 95:105069

Minias P, Janiszewska A, Pikus E, Zadworny T, Anderwald D (2021b) MHC reflects fine-scale habitat structure in white-tailed eagles, *Haliaeetus albicilla*. J Hered 112:335–345

Minias P, Pikus E, Whittingham LA, Dunn PO (2018) A global analysis of selection at the avian MHC. Evolution 72:1278–1293

Minias P, Vinkler M (2022) Selection balancing at innate immune genes: adaptive polymorphism maintenance in Toll-like receptors. Mol Biol Evol 39:msac102

Murrell B, Moola S, Mabona A, Weighill T, Sheward D, Kosakovsky Pond SL, Scheffler K (2013) FUBAR: a fast, unconstrained Bayesian approximation for inferring selection. Mol Biol Evol 30:1196–1205

Murrell B, Wertheim JO, Moola S, Weighill T, Scheffler K, Kosakovsky Pond SL (2012) Detecting individual sites subject to episodic diversifying selection. PLoS Genet 8:e1002764

Nelson-Flower MJ, Germain RR, MacDougall-Shackleton EA, Taylor SS, Arcese P (2018) Purifying selection in the toll-like receptors of song sparrows *Melospiza melodia*. J Hered 109:501–509

Péron G, Lebreton JD, Crochet PA (2010) Breeding dispersal in black-headed gull: the value of familiarity in a contrasted environment. J Anim Ecol 79:317–326

Promerová M, Králová T, Bryjová A, Albrecht T, Bryja J (2013) MHC class IIB exon 2 polymorphism in the Grey partridge (*Perdix perdix*) is shaped by selection, recombination and gene conversion. PLoS ONE 8:e69135

Quéméré E, Galan M, Cosson JF, Klein F, Aulagnier S, Gilot-Fromont E et al. (2015) Immunogenetic heterogeneity in a widespread ungulate: the European roe deer (*Capreolus capreolus*). Mol Ecol 24:3873–3887

Quesniaux V, Fremond C, Jacobs M, Parida S, Nicolle D, Yeremeev V et al. (2004) Toll-like receptor pathways in the immune responses to mycobacteria. Microbes Infect 6:946–959

Radwan J, Babik W, Kaufman J, Lenz TL, Winternitz J (2020) Advances in the evolutionary understanding of MHC polymorphism. Trends Genet 36:298–311

Ranathunge C, Chimahusky ME, Welch ME (2022) A comparative study of population genetic structure reveals patterns consistent with selection at functional microsatellites in common sunflower. Mol Genet Genom 297:1329–1342

Rozas J, Ferrer-Mata A, Sánchez-DelBarrio JC, Guirao-Rico S, Librado P, Ramos-Onsins SE, Sánchez-Gracia A (2017) DnaSP 6: DNA sequence polymorphism analysis of large data sets. Mol Biol Evol 34:3299–3302

Sallaberry-Pincheira N, González-Acuña D, Padilla P, Dantas GP, Luna-Jorquera G, Frere E et al. (2016) Contrasting patterns of selection between MHC I and II across populations of Humboldt and Magellanic penguins. Ecol Evol 6:7498–7510

Saper MA, Bjorkman P, Wiley DC (1991) Refined structure of the human histocompatibility antigen HLA-A2 at 2.6 Å resolution. J Mol Biol 219:277–319

Sarquis-Adamson Y, MacDougall-Shackleton EA (2016) Song sparrows *Melospiza melodia* have a home-field advantage in defending against sympatric malarial parasites. R Soc Open Sci 3:160216

Sebastian A, Herdegen M, Migalska M, Radwan J (2016) Amplisas: a web server for multilocus genotyping using next-generation amplicon sequencing data. Mol Ecol Res 16:498–510

Shapiro MD, Kronenberg Z, Li C, Domyan ET, Pan H, Campbell M et al. (2013) Genomic diversity and evolution of the head crest in the rock pigeon. Science 339:1063–1067

Shultz AJ, Sackton TB (2019) Immune genes are hotspots of shared positive selection across birds and mammals. eLife 8:e41815

Slade JW, Sarquis-Adamson Y, Gloor GB, Lachance MA, MacDougall-Shackleton EA (2017) Population differences at MHC do not explain enhanced resistance of song sparrows to local parasites. J Hered 108:127–134

Spurgin LG, Illera JC, Jorgensen TH, Dawson DA, Richardson DS (2014) Genetic and phenotypic divergence in an island bird: isolation by distance, by colonization or by adaptation? Mol Ecol 23:1028–1039

Spurgin LG, Richardson DS (2010) How pathogens drive genetic diversity: MHC, mechanisms and misunderstandings. Proc R Soc B 277:979–988

Spurgin LG, Van Oosterhout C, Illera JC, Bridgett S, Gharbi K, Emerson BC, Richardson DS (2011) Gene conversion rapidly generates major histocompatibility complex diversity in recently founded bird populations. Mol Ecol 20:5213–5225

Stephens M, Donnelly P (2003) A comparison of Bayesian methods for haplotype reconstruction from population genotype data. Am J Hum Genet 73:1162–1169

Temperley ND, Berlin S, Paton IR, Griffin DK, Burt DW (2008) Evolution of the chicken Toll-like receptor gene family: a story of gene gain and gene loss. BMC Genomics 9:62

Tirard C, Helfenstein F, Danchin E (2002) Polymorphic microsatellites in the black-legged kittiwake *Rissa tridactyla*. Mol Ecol Notes 2:431–433

Van Oosterhout C, Hutchinson WF, Wills DP, Shipley P (2004) MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. Mol Ecol Notes 4:535–538

Velová H, Gutowska-Ding MW, Burt DW, Vinkler M (2018) Toll-like receptor evolution in birds: gene duplication, pseudogenization, and diversifying selection. Mol Biol Evol 35:2170–2184

Waterhouse RM, Kriventseva EV, Meister S, Xi Z, Alvarez KS, Bartholomay LC et al. (2007) Evolutionary dynamics of immune-related genes and pathways in disease-vector mosquitoes. Science 316:1738–1743

Weaver S, Shank SD, Spielman SJ, Li M, Muse SV, Kosakovsky Pond SL (2018) Datamonkey 2.0: a modern web application for characterizing selective and other evolutionary processes. Mol Biol Evol 35:773–777

Włodarczyk R, Bouwhuis S, Bichet C, Podlaszczuk P, Chyb A, Indykiewicz P et al. (2022) Contrasting haemoparasite prevalence in larid species with divergent ecological niches and migration patterns. Parasitology 149:1479–1486

Xu W, Zhou X, Fang W, Chen X (2020) Genetic diversity of toll-like receptor genes in the vulnerable Chinese egret (*Egretta eulophotes*). PLoS ONE 15:e0233714

Yang J, Zhou M, Zhong Y, Xu L, Zeng C, Zhao X, Zhang M (2021) Gene duplication and adaptive evolution of Toll-like receptor genes in birds. Dev Comp Immunol 119:103990

ACKNOWLEDGEMENTS

We thank all participants in fieldwork, especially Jacek Betleja, Beata Dulisz, and Jarosław Kowalski. We thank three anonymous reviewers for constructive comments on the earlier drafts of the manuscript.

AUTHOR CONTRIBUTIONS

PM designed the study. PI, ML, JN, and TJ performed fieldwork and collected samples. PP and ACh performed laboratory analyses. PM and PP performed statistical and bioinformatic analyses. PM wrote the first draft of the manuscript, all authors revised the manuscript for intellectual content and approved the final draft.

COMPETING INTERESTS

The authors declare no competing interests.

ETHICS APPROVAL

Catching of birds was authorised by the Polish Academy of Sciences, with the approval of the General Directorate for Environmental Protection, Poland. The study was performed under the permit from the Local Bioethical Committee for Experiments on Animals in Łódź and appropriate Regional Environmental Protection Directorates in Poland.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41437-023-00645-2.

Correspondence and requests for materials should be addressed to Piotr Minias.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.