

Characterization of MHC class II variation in feral pigeons – copy number variation, polymorphism, and selection

Maciej Kamiński

University of Lodz

Piotr Minias

pminias@op.pl

University of Lodz

Research Article

Keywords:

Posted Date: October 27th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-11288379/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Abstract

The Major Histocompatibility Complex (MHC) genes encode proteins that play a key role in pathogen recognition by the adaptive immune system of jawed vertebrates. While the architecture and variation of MHC regions are often well characterized in species of high cultural and economic importance, the domestic pigeon *Columba livia domestica* remains poorly characterized from an immunogenetic perspective. The aim of this study was to characterize MHC class II (MHC-II) variation in feral populations of the domestic pigeon. To this end, we sampled 200 feral pigeons in five major cities in Poland and genotyped them at two ancient avian MHC-II gene lineages, DAB1 and DAB2. We detected moderate allelic diversity within each gene lineage, with 19 DAB1 and 24 DAB2 alleles detected across all populations. Genotype reconstruction indicated the presence of both single-copy and multi-copy haplotypes at DAB1 and DAB2, being consistent with intraspecific MHC gene copy number variation. We found that DAB1 and DAB2 exon 2 sequences clustered separately, suggesting pronounced divergence between the lineages. We also found a stronger signal of recombination (gene conversion) and a weaker signal of diversifying selection at DAB1 than at DAB2, suggesting distinct evolutionary trajectories. Our findings are consistent with the hypothesis that intraspecific MHC gene copy variation may be more widespread in birds than previously thought. Our study also highlights the potential for further immunogenetic research in the domestic pigeon, a species of high cultural and economic importance.

Introduction

The genes of the Major Histocompatibility Complex (MHC) form a key component of the adaptive immune system in jawed vertebrates, as they are essential for effective pathogen recognition. In general, MHC genes encode cell surface molecules, which bind short peptides originating predominantly from intracellular (MHC class I, MHC-I) or extracellular (mostly MHC class II, MHC-II) sources and present them to T cells, allowing the detection of non-self (foreign) antigenic peptides presented by host cells (Murphy and Weaver 2022). The peptides are loaded into the peptide-binding groove of MHC molecules, forming MHC-peptide complexes, which directly interact with T-cell receptors (Murphy and Weaver 2022). Recognition of MHC-presented foreign peptides by T cells not only triggers a cascade of adaptive immune responses aimed at pathogen clearance, but may also stimulate other (e.g. innate) immune pathways (Rodgers and Cook 2005). Consequently, MHC-based immunosurveillance is considered one of the central mechanisms by which the hosts detect and respond to pathogens, making it highly relevant in the evolutionary and ecological research on host-pathogen interactions (Sommer 2005, Piertney and Oliver 2006).

One of the most remarkable features of MHC genes is their extraordinary polymorphism (often referred to as hyper-polymorphism). In fact, the MHC is recognized among the most polymorphic genomic regions in vertebrates, with tens of thousands of alleles described in humans (Barker et al. 2026) and hundreds to thousands of alleles reported in some natural vertebrate populations (Bollmer et al. 2021, Biedrzycka et al. 2017). This extreme polymorphism is maintained through different forms of balancing selection, including heterozygote advantage (overdominance), negative frequency-dependent selection, and fluctuating

selection (Radwan et al. 2020). Over evolutionary timescales, balancing selection has favoured the accumulation and maintenance of nonsynonymous substitutions within MHC genes, primarily in the regions encoding the peptide-binding groove (exons 2 and 3 of MHC-I genes; exon 2 of MHC-IIA and MHC-IIB genes) (Hughes and Nei 1989). From an evolutionary perspective, MHC diversity has also been enhanced through gene duplication, which promoted the expansion of the MHC gene family in some species and produced substantial interspecific variation in gene copy number. In birds, the most extensive MHC expansion occurred in passerines (tens of MHC gene copies recorded in some species), whereas non-passerines usually retain relatively low numbers of MHC genes (Minias et al. 2019, He et al. 2021). Despite extensive duplications, avian MHC genes are strongly homogenized by gene conversion (consistent with concerted evolution; Hess and Edwards 2002, Minias 2024) and lose the signature of one-to-one orthology. The notable exceptions include two ancient and broadly conserved MHC-II gene lineages (DAB1/DAA1 and DAB2/DAA2) that retain recognizable orthologous relationships across birds (Burri et al. 2010, Minias et al. 2023) and ratite DAA3 lineage that retains the signal of orthology with MHC-IIA genes in non-avian reptiles (Minias and Babik 2024).

MHC gene copy number variation may also occur among haplotypes within species; however, empirical evidence for such intraspecific variation is scarce, largely due to technical challenges in sequencing, assembling, and annotating complex gene families (Peel et al. 2022). In birds, evidence for intraspecific MHC gene copy number variation originates from only a handful of species, mostly characterized by relatively simple MHC architecture. For example, it has been shown that the number of MHC-IIA and MHC-IIB gene pairs per haplotype varied between one and three in the crested ibis *Nipponia nippon* (Taniguchi et al. 2014) and between one and two in the oriental stork *Ciconia boyciana* (Tsuji et al. 2017). Similarly, pedigree-based segregation analyses in the barn owl *Tyto alba* suggested departures from the expected architecture of two tightly linked MHC-I loci, implying copy number variation among haplotypes (Gaigher et al. 2016). It nevertheless remains unresolved whether intraspecific MHC gene copy variation represents a rare taxon-specific phenomenon or whether it is a universal feature of avian MHC genes.

The research on the MHC has expanded significantly since the advent of next-generation sequencing approaches (Babik 2010). Currently, broad-scale population-level screening of MHC variation is often complemented by analyses of high-quality long-read genome assemblies, which provide important insights into the genomic architecture of the MHC region (Malmstrøm et al. 2016, Westerdahl et al. 2022, Silver et al. 2024, Babik et al. 2026). Despite these methodological advances, MHC variation in model and domesticated organisms still remains far better characterized compared to a wide range of non-model species (Bernatchez and Landry 2003). Nevertheless, even among domesticated organisms, the resolution of MHC variation and architecture is uneven. While species such as the domestic chicken *Gallus gallus* have particularly well-characterized MHC architecture (e.g. Kaufmann et al. 1999, Halabi and Kaufman 2022), others, such as the domestic pigeon *Columba livia domestica* remain comparatively understudied in this respect. The domestic pigeon is one of the most widely recognized bird species worldwide, not only due to its early domestication and cultural importance, but also because of the global distribution of its urban feral populations (Hernández-Alonso et al. 2023). It is also considered an important model organism for studies in evolutionary and developmental biology, morphology, and

neurobiology (Helms and Brugmann 2007, Wasserman et al. 2024). Despite genomic research and the availability of well-annotated genomic assemblies for the domestic pigeon (Shapiro et al. 2013, Murray et al. 2017), to the best of our knowledge, immunogenetic variation has not been comprehensively investigated in this species. This gap is particularly notable given the ecological importance of feral pigeons in urban ecosystems (Capoccia et al. 2018), their role in pathogen transmission, and associated public health concerns for humans (Haag-Wackernagel and Moch 2004, Mia et al. 2022). Taken together, the domestic pigeon appears as a particularly informative yet still underexploited system for investigating MHC variation.

The aim of this study was to characterize MHC-II variation in a central European population of the feral pigeon. To this end, we sampled 200 birds from five major cities in Poland and genotyped them at MHC-II B genes. Using available genetic resources, we designed primers that allowed lineage-specific genotyping of the DAB1 and DAB2 genes. We characterized and compared polymorphism, selection, and gene conversion between these two ancient MHC-II gene lineages and, finally, we used allele composition data to investigate presence of intraspecific copy number variation.

Material and Methods

Field procedures and sampling

Feral pigeons were captured and sampled in 2020 and 2021 in five major cities in Poland: Gdańsk, Krakow, Łódź, Warsaw, and Wrocław (see Kamiński et al. 2025 for detailed description of study plots). Birds were baited with dry seeds and captured using a noose trap. In total, 200 pigeons were captured (40 individuals per population). From each captured bird we collected ca. 40 µl of blood by the puncture of the ulnar vein. Blood was collected into Eppendorf tubes with 96% ethanol and stored at 4°C until analysis.

MHC genotyping

To separately genotype DAB1 and DAB2 gene lineages we designed lineage-specific primers based on feral pigeon MHC-II sequences (GenBank accession nos: KJ162459 and KJ162460) obtained by Burri et al. (2014). All primers were located in the intronic regions flanking exon 2. Forward primers *CoLi_DAB1_F* (5'-GAGCCTGGTGTGGGGAAC-3') and *CoLi_DAB2_F* (5'-ACATCCCCATTCTGCACGAG-3') were designed to selectively amplify DAB1 and DAB2 lineages, whereas the same reverse primer *CoLi_DAB_R* (5'-GGACTGGGCTTGTGTCTGTC-3') was designed and used for both gene lineages. For amplifications we used fusion primers containing Illumina Nextera Transposase adapter sequences (Illumina Corp., San Diego, CA, USA) and 8-bp barcodes for amplicon identification. DNA was isolated with Bio-Trace DNA Purification Kit (EURx, Gdańsk, Poland) and PCR amplifications were conducted in a total volume of 30 µl containing 15 µl of 2X HotStarTaq Plus MasterMix Kit (HotStarTaq Plus DNA Polymerase, PCR Buffer with 3 mM MgCl₂, and 400 µM of each dNTP) (Qiagen, Venlo, Netherlands), 9 µl of deionized water, 1.5 µl of each primer (final concentration of 1.25 µM), and 3 µl of genomic DNA solution. The following PCR conditions were used: initial denaturation (95°C for 5 min) followed by 30 cycles of denaturation (95°C for

60s), annealing (50°C for 30s) and elongation (72°C for 60s), and a final extension (72°C for 10 min). All PCR products were purified using AMPure XP Beads for DNA Cleanup (Beckman Coulter, Brea, CA, UAS) and their concentrations were assessed with a Quant-iT PicoGreen dsDNA assay kit (Invitrogen, Thermo Fisher Scientific, Waltham, MA, USA). Each library (DAB1 and DAB2) was prepared separately from equimolar amounts of PCR products using Illumina DNA/RNA UD Index Sets A-D and Q5 Hot Start High-Fidelity 2X Master Mix (New England Biolabs, Ipswich, MA, USA). DAB1 library was sequenced in 2024 using the 2x300bp Illumina MiSeq platform, while DAB2 library was sequenced in 2026 using the PE300 AVITI platform (Element Bioscience, San Diego, CA, USA). Sequences were trimmed in fastp v.0.23.4 (Chen 2023) and quality control was conducted in FastQC v.0.12.1 (Andrews 2010). Raw paired-end reads were merged using the BBMerge algorithm in Genious v.10.0.5 (Biomatters Ltd., Auckland, New Zealand).

MHC allele calling

Allele calling was conducted using a stand-alone version of AmpliSAS software (Sebastian et al. 2016, 2024). The maximum per amplicon number of reads used for allele calling was set to 5000. First, all merged sequences were de-multiplexed using unique barcode combinations. Second, the reads resulting from genotyping errors were identified based on a set of criteria recommended for Illumina sequencing (substitution errors—1%, indel errors—0.001%, minimum dominant frequency—25%; Biedrzycka et al. 2017) and clustered with reads identified as true alleles based on the dominant sequencing depth. Third, sequence clusters identified as chimeras or characterized by low (< 3%) per amplicon depth were filtered out. We also discarded all amplicons with total sequencing depth of < 200 reads. Reproducibility of allele calling was 100%, as assessed using 16 DAB1 and 11 DAB2 technical replicates (i.e. independent amplicons from the same DNA extracts). DAB1 and DAB2 genotypes were successfully obtained for 195 and 199 individuals, respectively. All sequences were aligned and trimmed to exonic regions (270 bp for DAB1 and 273 bp for DAB2 due to the presence of 3-bp insertion) in Geneious v10.0.5. No stop codons or frameshift mutations were detected, suggesting that all sequences represented putatively functional MHC-II allelic variants.

MHC genotype reconstruction

Allele composition within each gene lineage was highly structured among individuals, which facilitated manual reconstruction of genotypes and inference of copy number variation. As DAB1 and DAB2 were genotyped separately, multi-locus genotypes were inferred only within each lineage, and the linkage between the two lineages was not investigated. Manual genotype reconstruction and assignment of alleles to different gene copies were independently evaluated using an automatic genotype reconstruction algorithm implemented in MHC-Typer v1.1 (Huang et al. 2019). Briefly, the algorithm determines the number of putative gene copies (loci) and infers locus-specific allele configurations from multi-locus allele phenotypes (defined as the set of alleles detected in an individual). Putative allele configurations are resolved by applying penalties for missing errors (no alleles assigned to a locus-specific genotype) and mismatch errors (more than two alleles assigned to a locus-specific genotype). The resulting allele

configurations are then compared and ranked based on the maximum likelihood and the Bayesian Information Criterion (BIC). The analyses allowed for null alleles (initial frequency set to 0.05), empty phenotypes resulting from negative amplification, and copy number variation. Analyses were run using the annealing and optimization parameter values recommended by Huang et al. (2019), including a missing error penalty of -2 and a mismatch error penalty of -80. Departure from Hardy-Weinberg equilibrium was tested for the most polymorphic (in terms of allele richness) gene copies within each lineage (DAB1A and DAB2A) using GenAlex v6.5 (Peakall and Smouse 2021).

Sequence polymorphism, recombination, and selection

Measures of sequence polymorphism, including the number of segregating sites, total number of mutations, and nucleotide diversity, were calculated using DnaSP v.6.10.3 (Rozas et al. 2017). Sequence polymorphism was not estimated for DAB1B and DAB2B gene copies, which showed very low allelic richness of one and two alleles, respectively (see results for details).

Recombination signals, including putative gene conversion events, were assessed using seven different algorithms (Maxchi, BootScan, Genconv, SiScan, RDP, Chimaera, and 3Seq), as implemented in RDP v.4.97 (Martin et al. 2015). Analyses were conducted using default settings and Bonferroni correction for multiple comparisons. A putative recombination event was considered supported when detected by at least two independent algorithms. Distinct recombination events were identified based on differences in the inferred positions of recombination breakpoints. Recombination was assessed separately for DAB1 and DAB2 lineages, using all sequences available within each lineage.

Selection was assessed using the ratio of nonsynonymous (amino acid altering) to synonymous (silent) nucleotide substitution rates (dN/dS). The $dN/dS > 1$ indicates positive (diversifying) selection reflecting an excess of non-synonymous substitutions, $dN/dS \approx 1$ is consistent with neutral evolution, whereas $dN/dS < 1$ indicates negative (purifying) selection reflecting the removal of non-synonymous substitutions. Pervasive diversifying and purifying selection (acting consistently across the phylogeny) was analysed using Fast Unconstrained Bayesian AppRoximation (FUBAR; Murrell et al. 2013), whereas episodic diversifying selection (acting on a subset of phylogenetic branches) was analysed using Mixed Effects Model of Evolution (MEME) (Murrell et al. 2012). Both methods quantified selection at the codon level. Analyses were conducted at the Datamonkey web server (Weaver et al. 2018) with default settings and phylogenies inferred from input sequence alignments. Recombinant regions in the input alignments were separated using RDP4 software, as the presence of recombinant sequences may affect reconstruction of allele phylogenies used to infer substitution rates and, thus, introduce biases in selection analyses (Anisimova et al. 2003). Default thresholds of posterior probability > 0.90 (FUBAR) or $p < 0.1$ (MEME) were used to identify sites under selection. Selection analyses were conducted separately for each gene lineage (DAB1 and DAB2), using either all available sequences or sequences assigned to the most polymorphic gene copy (DAB1A and DAB2A). The dN/dS ratios were calculated for *i*) all codon sites, *ii*) sites corresponding to human MHC-II peptide-binding region (PBR), as inferred from the crystallographic structure of the MHC-II molecule (Brown et al. 1993), and *iii*) sites identified as positively selected.

MHC allele clustering

To resolve clustering of DAB1 and DAB2 sequences assigned to different gene copies we used Bayesian tree-building algorithms as implemented in MrBayes v.3.2.6 (Ronquist et al. 2012). The model of nucleotide substitution was selected based on the maximum likelihood analysis conducted in Mega v.6.0 (Tamura et al. 2013). The little tinamou *Crypturellus soui* (Palaeognathae: Tinamiformes) DAB1 and DAB2 sequences (as extracted from the reference VWPX01015387 genome by Minias et al. 2023) were used as outgroups. The clustering was conducted using two independent chains, each consisting of one million iterations, burn-in length of 250 000 iterations, and sampling frequency of 500. The average effective sample size for model parameters was 1197 ± 240 [SD]. Convergence of independent chains was inferred from the average standard deviation of split frequencies (ASDSF) approaching zero (Lakner et al. 2008) and the potential scale reduction factors (PSRF) approaching one (Gelman and Rubin 1992). Extended majority-rule consensus topology was determined in Geneious v.10.0.5.

Results

Allele numbers and copy number variation

In total, we identified 19 DAB1 and 24 DAB2 alleles across all individuals. The number of DAB1 alleles per individual ranged from one to four, whereas the number of DAB2 alleles ranged from one to five. Most birds carried one or two DAB1 and DAB2 alleles (90.3% and 86.9% of individuals, respectively), consistent with the presence of a single gene copy of each lineage, hereafter referred to as DAB1A and DAB2A. Observed genotype frequencies were consistent with Hardy-Weinberg expectations for both DAB1A ($\chi^2 = 60.91$, $df = 105$, $p = 1.00$) and DAB2A ($\chi^2 = 121.54$, $df = 105$, $p = 1.00$).

Three to five alleles were recorded in only 19 individuals at DAB1 and 24 individuals at DAB2, indicating the presence of additional gene copies. These multi-locus genotypes consistently comprised: *i*) one or two DAB1A/DAB2A alleles shared with other individuals (marked in orange in Figs. 1 and 2), and *ii*) two additional alleles that were not detected in individuals carrying a single gene copy of each lineage, except for one DAB2 genotype with four such alleles (see below). The composition of these additional alleles was highly structured, as all multi-locus DAB1 genotypes contained the same allele (marked in green in Fig. 1) together with one of the three additional alleles (marked in blue in Fig. 1). Similarly, most multi-locus DAB2 genotypes contained one of two 270-bp alleles (marked in green in Fig. 2) and one of three 273-bp alleles (marked in blue in Fig. 2).

The allele combinations observed with each lineage were consistent with heterozygous genotypes comprising two haplotypes that differed in copy number: one haplotype containing a single gene copy (DAB1A or DAB2A) and the other carrying three gene copies (DAB1A plus DAB1B and DAB1C, or DAB2A plus DAB2B and DAB2C). The DAB1B gene copy was monomorphic, containing a single allele (DAB1B*01; Fig. 1), DAB2B contained two alleles (DAB2B*01–02; Fig. 2), whereas both DAB1C and DAB2C contained three alleles (DAB1C*01–03 and DAB2C*01–03, respectively; Figs. 1 and 2), all of which were restricted to

multi-locus haplotypes. For DAB2, one individual carried two alleles at both DAB2B and DAB2C, consistent with a genotype comprising two copies of the three-locus DAB2 haplotype (CoLi-DAB2-GT26 in Fig. 2). Two additional individuals lacked DAB2C alleles, most likely reflecting two-locus DAB2 haplotypes in which DAB2C gene copy was absent. The assignment of alleles to three DAB1 and three DAB2 gene copies, as described above, was supported by the maximum-likelihood genotype reconstruction, which identified these configurations as having the highest likelihood and lowest BIC values, with no missing or mismatch errors (Supplementary Tables S1 and S2).

Sequence polymorphism, recombination and selection

We found moderate to high polymorphism across all DAB1 sequences (50 segregating sites, 0.067 nucleotide diversity) and DAB2 sequences (69 segregating sites, 0.093 nucleotide diversity) (Table 1). In both gene lineages, DAB1A/DAB2A sequences showed considerably higher nucleotide diversity than DAB1C/DAB2C sequences (Table 1). We detected stronger recombination signal in DAB1 than DAB2 sequences. At DAB1, four independent recombination events were detected, involving 12 out of 19 sequences (63%), with one to five recombinant sequences associated with each event. At DAB2, we detected two independent recombination events with 6 out of 24 sequences (25%) identified as recombinants (two and four recombinant sequences per event).

Table 1

Polymorphism and selection at MHC class II exon 2 in the feral pigeon. All sequences were assigned to three gene copies within each gene lineages (DAB1 and DAB2). The signature of selection (the ratio of nonsynonymous to synonymous substitution rates, dN/dS) was calculated for *i*) all codon sites, *ii*) sites corresponding to human MHC-II peptide-binding region (PBR), as inferred from the crystallographic structure of the MHC-II molecule, and *iii*) sites identified as positively selected.

Gene lineage	Gene copy	No. alleles	No. segregating sites	Total no. mutations	Nucleotide diversity	dN/dS		
						All sites	Human PBR sites	Positively selected sites
DAB1	DAB1A	15	50	51	0.072	0.59	0.83	1.39
	DAB1B	1	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
	DAB1C	3	12	12	0.030	n.a.	n.a.	n.a.
	All	19	50	52	0.067	0.64	0.93	1.62
DAB2	DAB2A	19	69	79	0.090	0.86	0.96	2.90
	DAB2B	2	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
	DAB2C	3	9	9	0.022	n.a.	n.a.	n.a.
	All	24	69	79	0.093	0.79	0.88	2.21

In contrast, the signature of positive diversifying selection was stronger at DAB2 than DAB1. We identified five sites under pervasive diversifying selection across DAB1 sequences (one of which was apparent only

at DAB1A) and two additional sites showed signature of episodic diversifying selection (Fig. 2). At DAB2, we detected seven sites under pervasive diversifying selection and additional eight sites showed signature of episodic diversifying selection (Fig. 2). Most of these sites, both at DAB1 (71%) and DAB2 (60%), corresponded to the human PBR sites (Fig. 2). The dN/dS ratio across positively selected sites was 1.62 for DAB1 and 2.21 for DAB2 sequences, respectively (Table 1). Only one site showed evidence of purifying selection at DAB2 (Fig. 2).

Allele phylogenetic clustering

DAB1 and DAB2 sequences formed separate, well-supported clusters (100% posterior probability; Fig. 4). However, we found no evidence of substantial divergence among sequences assigned to different gene copies within either lineage. In fact, most of DAB1B/DAB2B and DAB1C/DAB2C sequences were deeply embedded within the DAB1A/DAB2A clusters (Fig. 4), which may indicate either recent duplication or homogenization through inter-locus gene conversion.

Discussion

Our study provided an in-depth characterization of MHC- II variation across feral populations of an important domesticated bird species, the feral pigeon. Importantly, we separately genotyped two ancient avian MHC-II gene lineages, DAB1 and DAB2. Limited polymorphism and highly structured allele composition facilitated effective genotype reconstruction and inference of intraspecific copy number variation. Specifically, within each gene lineage, we inferred the presence of haplotypes comprising either a single gene copy or multiple (in most cases three) gene copies. This finding adds the feral pigeon to the growing, but still limited, number of avian species showing evidence of intraspecific MHC copy number variation. Additionally, comparisons between the two lineages indicated slightly higher polymorphism, a stronger signal of diversifying selection, and a weaker signal of recombination (gene conversion) at DAB2 than at DAB1. These differences suggest that the two ancient MHC-II gene lineages may follow distinct evolutionary trajectories.

It has been well established that MHC genes show substantial variation in copy number among species (Bentkowski and Radwan 2019). In birds, the number of MHC-II genes ranges from one or two copies in some non-passerine bird species to tens of copies in many passerines, a variation thought to have primarily arisen in response to fluctuating selection, whereby different species evolve different optimal gene copy number through lineage-specific patterns of gene duplication and deletion (Minias et al. 2019). In contrast, intraspecific variation in MHC gene copy number among haplotypes has rarely been reported. It remains an open question whether this scarcity represents a true biological phenomenon (i.e. most species show no copy number variation among haplotypes) or instead reflects technical challenges in resolving MHC haplotypes. To date, intraspecific MHC copy number variation was inferred through direct haplotype reconstruction using a combination of cloning and targeted sequencing of the MHC region (Taniguchi et al. 2014; Tsuji et al. 2017). Although assembling and inferring haplotypes within MHC genomic regions remains challenging, particularly when the number of gene copies is high (Westerdahl et al. 2022), recent high-quality haplotype-resolved genomes may gradually help overcome these challenges.

For instance, the pseudo-haplotype-resolved genome of the bluethroat *Luscinia svecica* revealed substantial structural differences between haplotypes, including copy number variation in both MHC class I and II genes (Strand et al. 2026). Alternatively, intraspecific copy number variation can be inferred indirectly through pedigree-based segregation analyses, as in the barn owl (Gaigher et al. 2016), or through genotype reconstruction and allele assignment to individual gene copies, as in the common buzzard *Buteo buteo* (Winternitz et al. 2023). The latter approaches, however, are feasible primarily in species with relatively simple MHC architecture and limited levels of polymorphism. Here, we took advantage of these characteristics to reconstruct genotypes and infer copy number variation in the feral pigeon.

Interestingly, we detected a similar pattern of copy number variation within each of the two ancient avian MHC-II gene lineages, DAB1 and DAB2. Most individuals across our study populations were characterized by genotypes putatively consisting of two single-copy haplotypes, resulting in one or two alleles detected within each lineage (representing homo- and heterozygotes, respectively). Nevertheless, a relatively small proportion of individuals carried more than two alleles of either DAB1 or DAB2, indicating the presence of multiple gene copies. Manual and simulation-based genotype reconstruction provided consistent support for the presence of three-locus haplotypes in these individuals, which in most cases formed genotypes comprising one three-locus and one single-locus haplotype (only one genotype consisted of two three-locus haplotypes). Additionally, at DAB2, we detected two individuals with putative haplotypes consisting of two gene copies. Multi-locus haplotypes were present in all five study populations representing different major cities in Poland, suggesting their widespread occurrence. The two additional gene copies at these haplotypes were, however, characterized by low allelic diversity (1–3 alleles detected per gene copy). Hence, the functional significance of these additional gene copies in terms of pathogen recognition capacity and their potential fitness consequences remain to be investigated.

Phylogenetic clustering of exon 2 sequences provided compelling evidence for pronounced divergence between the DAB1 and DAB2 gene lineages. This was unexpected, as previous analyses of owls revealed that exon 2 sequences did not cluster according to gene lineage, even within highly supported intraspecific clusters, which suggests shuffling of these exons among genes by recombination (gene conversion) (Burri et al. 2008). In fact, the signal of DAB1 and DAB2 orthology in owls and other bird species has usually been retained only in the short upstream region of exon 3 (Burri et al. 2008, 2010). In the feral pigeon, the clustering of exon 2 sequences by gene lineage rather than gene copy suggests that multi-locus haplotypes arose independently within the two lineages. On the other hand, we detected no evidence of divergence among sequences originating from different gene copies within either lineage. This may be explained by relatively recent origin of these duplications (leaving limited time for sequence divergence to accumulate between gene copies) or by homogenization of duplicated gene copies through gene conversion. The latter scenario is consistent with concerted evolution characteristic for avian MHC genes (Hess and Edwards 2002).

Finally, our study indicated that DAB1 and DAB2 genes followed distinct evolutionary trajectories. While the overall level of polymorphism was generally similar between the two lineages (only slightly higher at DAB2), we detected substantial differences in recombination and selection. Specifically, DAB1 showed

much stronger signal of recombination, with four detected recombination events involving ~ 60% of sequences, compared with two recombination events involving only 25% sequences at DAB2. In contrast, DAB2 showed a stronger signature of diversifying selection, with 15 sites identified as under diversifying selection and a higher dN/dS ratio, compared with seven sites under diversifying selection and a lower dN/dS ratio at DAB1. The stronger signal of diversifying selection at DAB2 is consistent with a broader analysis across diverse avian orders, which also revealed a higher number of positively selected sites at DAB2 than at DAB1 (Minias et al. 2023). These differences suggest that the two ancient MHC-II gene lineages may be subject to different pathogen-driven selection regimes, potentially allowing them to contribute differently to pathogen recognition by the immune system (Minias et al. 2023).

In conclusion, our characterization of MHC-II variation in the feral pigeon provided evidence for MHC copy number variation within both the DAB1 and DAB2 gene lineages. This finding is in line with the hypothesis that intraspecific MHC gene copy variation may represent a relatively widespread phenomenon in birds. We suggest that analyses of high-quality phased genome assemblies may provide further insights into the true incidence of MHC gene copy number variation in birds and other vertebrates. Our study also provided evidence for moderate MHC-II polymorphism across the feral pigeon populations sampled in Poland. This finding is consistent with relatively low mtDNA diversity reported in feral pigeons (Prakas et al. 2025) and could possibly reflect the recurrent history of bottlenecks inherent in the domestication process (Pacheco et al. 2020). Comparisons of MHC variation between feral and natural wild pigeon populations would provide a direct test of this hypothesis. Finally, our study highlights the potential for further immunogenetic research in the domestic pigeon, a species of high cultural and economic importance.

Declarations

Ethics statement

The study was performed in accordance with the national legislation and was approved by the Local Bioethical Commission for Experiments on Animals in Lodz (1/ŁB154/2020 and 36/ŁB154 ptp/2020) and the General Directorate for Environmental Protection in Poland (DZP-WG.6401.15.2020.KS and DZPWG.6401.15.2020.TŁ.3).

Funding

The study was financially supported by the research grant Preludium 16 of the National Science Centre in Poland (2018/31/N/NZ8/00832).

Author Contribution

MK: conceptualization, methodology, investigation, data curation, writing – original draft. PM: conceptualization, methodology, formal analysis, writing-original draft, visualization.

Acknowledgement

We thank Joanna Drzewińska-Chańko for support in laboratory work.

Data Availability

All sequences will be deposited in GenBank.

References

1. Andrews, S. (2010). FASTQC: A quality control tool for high throughput sequence data. Available at: <http://www.bioinformatics.babraham.ac.uk/projects/fastqc>
2. 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.
3. Babik, W. (2010). Methods for MHC genotyping in non-model vertebrates. *Molecular Ecology Resources*, 10, 237–251.
4. Babik, W., Dudek, K., Palomar, G., Marszałek, M., Dubin, G., Yun, M. H., & Migalska, M. (2026). MHC in newts illuminates the evolutionary dynamics of complex regions in giant genomes. *Genome Research*, 36, 303–317.
5. Barker, D. J., Natarajan, R. H., Cooper, M. A., Hopper, S. J., Yates, A. D., Parham, P., Marsh, S. G. E., & Robinson, J. (2026). The IPD-IMGT/HLA database: recent developments in sequence submission. *Nucleic Acids Research*, 54, D1152-D1158.
6. Bentkowski, P., & Radwan, J. (2019). Evolution of major histocompatibility complex gene copy number. *PLoS Computational Biology*, 15, e1007015.
7. Bernatchez, L., & Landry, C. (2003). MHC studies in nonmodel vertebrates: what have we learned about natural selection in 15 years? *Journal of Evolutionary Biology*, 16, 363–377.
8. Biedrzycka, A., O'Connor, E., Sebastian, A., Migalska, M., Radwan, J., Zając, T., Bielański, W., Solarz, W., Ćmiel, A., & Wester Dahl, H. (2017). Extreme MHC class I diversity in the sedge warbler (*Acrocephalus schoenobaenus*); selection patterns and allelic divergence suggest that different genes have different functions. *BMC Evolutionary Biology*, 17, 159.
9. Bollmer, J. L., Dunn, P. O., Freeman-Gallant, C. R., & Whittingham, L. A. (2012). Social and extra-pair mating in relation to major histocompatibility complex variation in common yellowthroats. *Proceedings of the Royal Society B*, 279, 4778–4785.
10. Brown, J. H., Jardetzky, T. S., Gorga, J. C., Stern, L. J., Urban, R. G., Strominger, J. L., & Wiley, D. C. (1993). Three-dimensional structure of the human class II histocompatibility antigen HLA-DR1. *Nature*, 364, 33–39.
11. Burri, R., Promerová, M., Goebel, J., & Fumagalli, L. (2014). PCR-based isolation of multigene families: lessons from the avian MHC class IIB. *Molecular Ecology Resources*, 14, 778–788.

12. Burri, R., Salamin, N., Studer, R. A., Roulin, A., & Fumagalli, L. (2010). Adaptive divergence of ancient gene duplicates in the avian MHC class II β . *Molecular Biology and Evolution*, 27, 2360–2374.
13. Capoccia, S., Boyle, C., & Darnell, T. (2018). Loved or loathed, feral pigeons as subjects in ecological and social research. *Journal of Urban Ecology*, 4, juy024.
14. Chen, S. (2023). Ultrafast one-pass FASTQ data preprocessing, quality control, and deduplication using fastp. *iMeta*, 2, e107.
15. Gaigher, A., Burri, R., Gharib, W. H., Taberlet, P., Roulin, A., & Fumagalli, L. (2016). Family-assisted inference of the genetic architecture of major histocompatibility complex variation. *Molecular Ecology Resources*, 16, 1353–1364.
16. Gelman, A., & Rubin, D. B. (1992). Inference from iterative simulation using multiple sequences. *Statistical Science*, 7, 457–472.
17. Haag-Wackernagel, D., & Moch, H. (2004). Health hazards posed by feral pigeons. *Journal of Infection*, 48, 307–313.
18. Halabi, S., & Kaufman, J. (2022). New vistas unfold: Chicken MHC molecules reveal unexpected ways to present peptides to the immune system. *Frontiers in Immunology*, 13, 886672.
19. He, K., Minias, P., & Dunn, P. O. (2021). Long-read genome assemblies reveal extraordinary variation in the number and structure of MHC loci in birds. *Genome Biology and Evolution*, 13, evaa270.
20. Helms, J. A., & Brugmann, S. A. (2007). The origins of species-specific facial morphology: the proof is in the pigeon. *Integrative and Comparative Biology*, 47, 338–342.
21. Hernández-Alonso, G., Ramos-Madrigal, J., van Grouw, H., Ciucani, M. M., Cavill, E. L., Sinding, M. H. S., Gopalakrishnan, S., Pacheco, G., & Gilbert, M. T. P. (2023). Redefining the evolutionary history of the rock dove, *Columba livia*, using whole genome sequences. *Molecular Biology and Evolution*, 40, msad243.
22. Hess, C. M., & Edwards, S. V. (2002). The evolution of the Major Histocompatibility Complex in birds. *Bioscience*, 52, 423–431.
23. Huang, K., Zhang, P., Dunn, D. W., Wang, T., Mi, R., & Li, B. (2019). Assigning alleles to different loci in amplifications of duplicated loci. *Molecular Ecology Resources*, 19, 1240–1253.
24. Hughes, A. L., & Nei, M. (1989). Nucleotide substitution at major histocompatibility complex class II loci: evidence for overdominant selection. *Proceedings of the National Academy of Sciences*, 86, 958–962.
25. Kamiński, M., Chyb, A., Matson, K. D., & Minias, P. (2025). Constitutive innate immune defenses in relation to urbanization and population density in an urban bird, the feral pigeon *Columba livia domestica*. *Integrative Zoology*, 20, 1136–1148.
26. Kaufman, J., Jacob, J., Shaw, J., Walker, B., Milne, S., Beck, S., & Salomonsen, J. (1999). Gene organisation determines evolution of function in the chicken MHC. *Immunological reviews*, 167, 101–117.
27. Lakner, C., Van Der Mark, P., Huelsenbeck, J. P., Larget, B., & Ronquist, F. (2008). Efficiency of Markov chain Monte Carlo tree proposals in Bayesian phylogenetics. *Systematic Biology*, 57, 86–103.

28. Malmstrøm, M., Matschiner, M., Tørresen, O. K., Star, B., Snipen, L. G., Hansen, T. F., Baalsrud, H. T., Nederbragt, A. J., Hanel, R., Salzburger, W., Stenseth, N. C., Jakobsen, K. S., & Jentoft, S. (2016). Evolution of the immune system influences speciation rates in teleost fishes. *Nature Genetics*, 48, 1204–1210.
29. Martin, D. P., Murrell, B., Golden, M., Khoosal, A., & Muhire, B. (2015). RDP4: Detection and analysis of recombination patterns in virus genomes. *Virus Evolution*, 1, vev003.
30. Mia, M. M., Hasan, M., & Hasnath, M. R. (2022). Global prevalence of zoonotic pathogens from pigeon birds: A systematic review and meta-analysis. *Heliyon*, 8, e09732.
31. Minias, P. (2024). Evolutionary variation in gene conversion at the avian MHC is explained by fluctuating selection, gene copy numbers and life history. *Molecular Ecology*, 33, e17453.
32. Minias, P., & Babik, W. (2024). Palaeognaths reveal evolutionary ancestry of the avian Major Histocompatibility Complex class II. *Genome Biology and Evolution*, 16, evae211.
33. Minias, P., Edwards, S. V., & Babik, W. (2023). Ancient duplication, coevolution, and selection at the MHC class IIA and IIB genes of birds. *Frontiers in Immunology*, 14, 1250824.
34. Minias, P., Pikus, E., Whittingham, L. A., & Dunn, P. O. (2019). Evolution of copy number at the MHC varies across the avian tree of life. *Genome Biology and Evolution*, 11, 17–28.
35. Murphy, K., & Weaver, C. (2022). *Janeway's immunobiology*. Garland Science.
36. Murray, G. G., Soares, A. E., Novak, B. J., Schaefer, N. K., Cahill, J. A., Baker, A. J., et al. (2017). Natural selection shaped the rise and fall of passenger pigeon genomic diversity. *Science*, 358, 951–954.
37. Murrell, B., Moola, S., Mabona, A., Weighill, T., Sheward, D., Kosakovsky Pond, S. L., & Scheffler, K. (2013). FUBAR: a fast, unconstrained bayesian approximation for inferring selection. *Molecular Biology and Evolution*, 30, 1196–1205.
38. Murrell, B., Wertheim, J. O., Moola, S., Weighill, T., Scheffler, K., & Kosakovsky Pond, S. L. (2012). Detecting individual sites subject to episodic diversifying selection. *PLoS Genetics*, 8, e1002764.
39. Pacheco, G., van Grouw, H., Shapiro, M. D., Gilbert, M. T. P., & Vieira, F. G. (2020). Darwin's fancy revised: an updated understanding of the genomic constitution of pigeon breeds. *Genome Biology and Evolution*, 12, 136–150.
40. Peakall, R. O. D., & Smouse, P. E. (2012). GenAEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research—an update. *Bioinformatics*, 28, 2537–2539.
41. Peel, E., Silver, L., Brandies, P., Zhu, Y., Cheng, Y., Hogg, C. J., & Belov, K. (2022). Best genome sequencing strategies for annotation of complex immune gene families in wildlife. *GigaScience*, 11, giac100.
42. Piertney, S. B., & Oliver, M. K. (2006). The evolutionary ecology of the major histocompatibility complex. *Heredity*, 96, 7–21.
43. Prakas, P., Bea, A., Švažas, S., Vaitkuvienė, D., Stepanovas, D., & Butkauskas, D. (2025). Mitochondrial D-loop sequencing reveals lower genetic variability in Feral Pigeons (*Columba livia* f. *urbana*) than in Woodpigeons (*Columba palumbus*) and European Turtle Doves (*Streptopelia turtur*). *Baltic Forestry*, 31, id779-id779.

44. Radwan, J., Babik, W., Kaufman, J., Lenz, T. L., & Winternitz, J. (2020). Advances in the evolutionary understanding of MHC polymorphism. *Trends in Genetics*, 36, 298–311.
45. Rodgers, J. R., & Cook, R. G. (2005). MHC class Ib molecules bridge innate and acquired immunity. *Nature Reviews Immunology*, 5, 459–471.
46. Ronquist, F., Teslenko, M., Van Der Mark, P., Ayres, D. L., Darling, A., Höhna, S., Larget, B., Liu, L., Suchard, M. A., & Huelsenbeck, J. P. (2012). MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Systematic Biology*, 61, 539–542.
47. Rozas, J., Ferrer-Mata, A., Sánchez-DelBarrio, J. C., Guirao-Rico, S., Librado, P., Ramos-Onsins, S. E., & Sánchez-Gracia, A. (2017). DnaSP 6: DNA sequence polymorphism analysis of large data sets. *Molecular Biology and Evolution*, 34, 3299–3302.
48. Sebastian, A., Herdegen, M., Migalska, M., & Radwan, J. (2016). amplisas: a web server for multilocus genotyping using next-generation amplicon sequencing data. *Molecular Ecology Resources*, 16, 498–510.
49. Sebastian, A., Migalska, M., & Gaczorek, T. (2024). AmpliSAS and AmpliHLA: web server and local tools for MHC typing of non-model species and human using NGS data. In: Boegel, S. (eds) *HLA Typing. Methods in Molecular Biology*, vol 2809. Humana, New York, NY.
50. Shapiro, M. D., Kronenberg, Z., Li, C., Domyan, E. T., Pan, H., Campbell, M., et al. (2013). Genomic diversity and evolution of the head crest in the rock pigeon. *Science*, 339, 1063–1067.
51. Silver, L. W., Hogg, C. J., & Belov, K. (2024). Plethora of new marsupial genomes informs our knowledge of marsupial MHC class II. *Genome Biology and Evolution*, 16, evae156.
52. Sommer, S. (2005). The importance of immune gene variability (MHC) in evolutionary ecology and conservation. *Frontiers in Zoology*, 2, 16.
53. Strand, M. A., Enevoldsen, E. L., Tørresen, O. K., Skage, M., Ferrari, G., Tooming-Klunderud, A., Leder, E., Lifjeld, J. T., Johnsen, A., & Jakobsen, K. S. (2026). A chromosome-level genome assembly of the bluethroat (*Luscinia s. svecica*; Linnaeus, 1758) uncovers the complex MHC region. *Journal of Heredity*, esag058 (online).
54. Tamura, K., Stecher, G., Peterson, D., Filipski, A., & Kumar, S. (2013). MEGA6: molecular evolutionary genetics analysis version 6.0. *Molecular Biology and Evolution*, 30, 2725–2729.
55. Taniguchi, Y., Matsumoto, K., Matsuda, H., Yamada, T., Sugiyama, T., Homma, K., Kaneko, Y., Yamagishi, S., & Iwaisaki, H. (2014). Structure and polymorphism of the major histocompatibility complex class II region in the Japanese Crested Ibis, *Nipponia nippon*. *PLoS ONE*, 9, e108506.
56. Tsuji, H., Taniguchi, Y., Ishizuka, S., Matsuda, H., Yamada, T., Naito, K., & Iwaisaki, H. (2017). Structure and polymorphisms of the major histocompatibility complex in the Oriental stork, *Ciconia boyciana*. *Scientific Reports*, 7, 42864.
57. Wasserman, E. A., Turner, B. M., & Güntürkün, O. (2024). The pigeon as a model of complex visual processing and category learning. *Neuroscience Insights*, 2024, 19.
58. Weaver, S., Shank, S. D., Spielman, S. J., Li, M., Muse, S. V., & Kosakovsky Pond, S. L. (2018). Datamonkey 2.0: a modern web application for characterizing selective and other evolutionary

processes. *Molecular Biology and Evolution*, 35, 773–777.

59. Westerdahl, H., Mellinger, S., Sigeman, H., Kutschera, V. E., Proux-Wéra, E., Lundberg, M., Weissensteiner, M., Churcher, A., Bunikis, I., Hansson, B., Wolf, J. B. W., & Strandh, M. (2022). The genomic architecture of the passerine MHC region: High repeat content and contrasting evolutionary histories of single copy and tandemly duplicated MHC genes. *Molecular Ecology Resources*, 22, 2379–2395.
60. Winternitz, J., Chakarov, N., Rinaud, T., Ottensmann, M., & Krüger, O. (2023). High functional allelic diversity and copy number in both MHC classes in the common buzzard. *BMC Ecology and Evolution*, 23, 24.

Figures

DAB1 genotype ID	DAB1A*01	DAB1A*02	DAB1A*03	DAB1A*04	DAB1A*05	DAB1B*01	DAB1C*01	DAB1C*02	DAB1C*03	No. all DAB1 alleles	No. DAB1A alleles	No. DAB1B alleles	No. DAB1C alleles
CoLi-DAB1-GT01										3	1	1	1
CoLi-DAB1-GT02										4	2	1	1
CoLi-DAB1-GT03										3	1	1	1
CoLi-DAB1-GT04										4	2	1	1
CoLi-DAB1-GT05										4	2	1	1
CoLi-DAB1-GT06										3	1	1	1
CoLi-DAB1-GT07										3	1	1	1
CoLi-DAB1-GT08										4	2	1	1
CoLi-DAB1-GT09										4	2	1	1
CoLi-DAB1-GT10										4	2	1	1
CoLi-DAB1-GT11										4	2	1	1
CoLi-DAB1-GT12										4	2	1	1
CoLi-DAB1-GT13										3	1	1	1
CoLi-DAB1-GT14										3	1	1	1
CoLi-DAB1-GT15										3	1	1	1
CoLi-DAB1-GT16										4	2	1	1
CoLi-DAB1-GT17										3	1	1	1
CoLi-DAB1-GT18										3	1	1	1
CoLi-DAB1-GT19										3	1	1	1

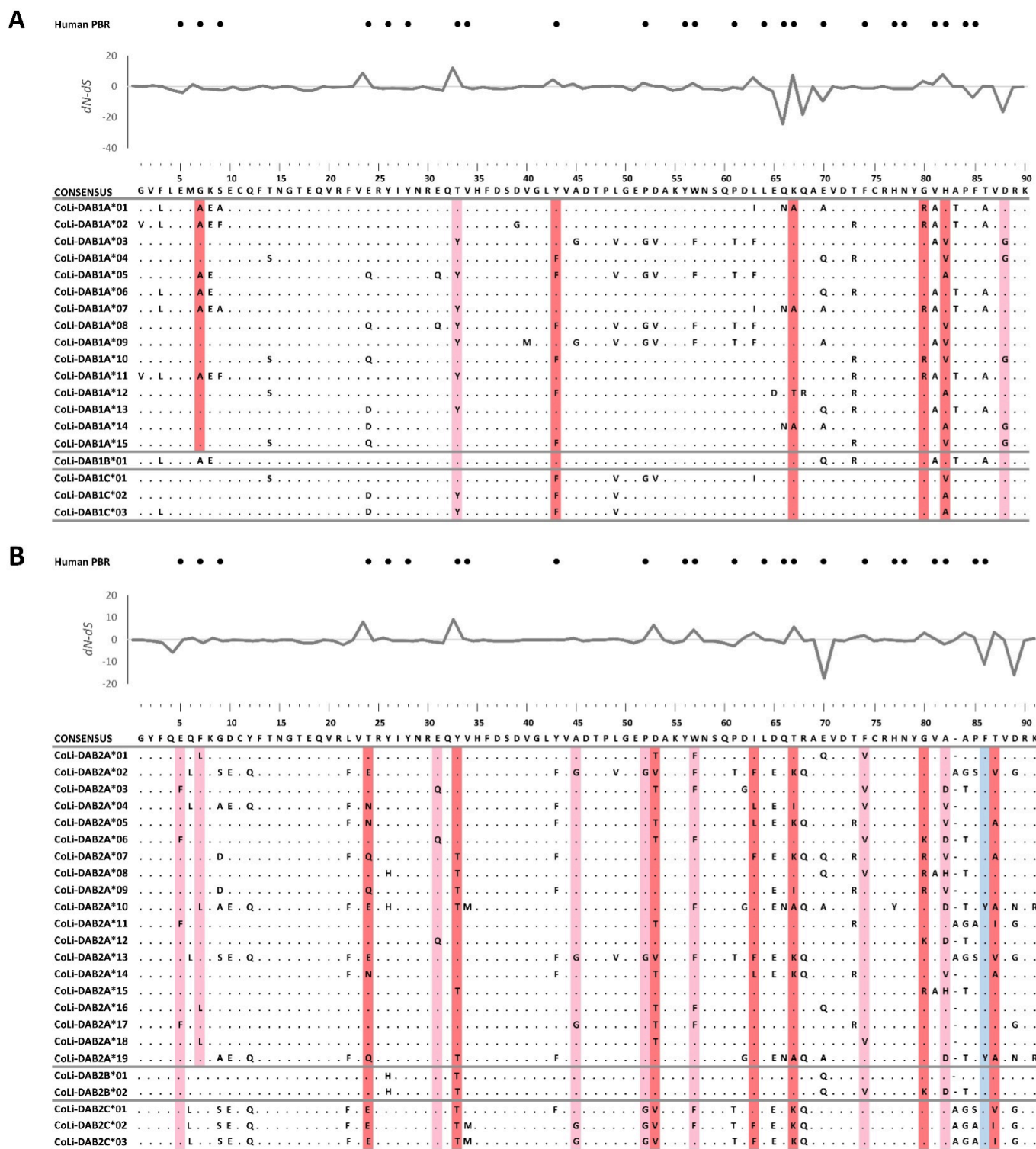
Figure 1

The MHC-II DAB1 genotype reconstruction for 19 feral pigeons with more than two alleles detected. Alleles were assigned to three loci marked with orange (DAB1A), green (DAB1B), and blue (DAB1C). The observed allele configurations indicate that genotypes are formed by two haplotypes characterized by copy number variation, i.e. by one-locus haplotype (DAB1A) and three-locus haplotype (DAB1A-DAB1B-DAB1C). DAB1B and DAB1C alleles were not detected in individuals with two copies of one-locus DAB1A haplotype.

DAB2 genotype ID	DAB2A*01	DAB2A*02	DAB2A*03	DAB2A*04	DAB2B*01	DAB2B*01	DAB2C*01	DAB2C*02	DAB2C*03	No. all DAB2 alleles	No. DAB2A alleles	No. DAB2B alleles	No. DAB2C alleles
CoLi-DAB2-GT01										3	1	1	1
CoLi-DAB2-GT02										3	1	1	1
CoLi-DAB2-GT03										3	1	1	1
CoLi-DAB2-GT04										3	1	1	1
CoLi-DAB2-GT05										3	1	1	1
CoLi-DAB2-GT06										3	1	1	1
CoLi-DAB2-GT07										3	1	1	1
CoLi-DAB2-GT08										3	1	1	1
CoLi-DAB2-GT09										3	1	1	1
CoLi-DAB2-GT10										4	2	1	1
CoLi-DAB2-GT11										3	1	1	1
CoLi-DAB2-GT12										3	1	1	1
CoLi-DAB2-GT13										3	1	1	1
CoLi-DAB2-GT14										3	1	1	1
CoLi-DAB2-GT15										3	1	1	1
CoLi-DAB2-GT16										3	1	1	1
CoLi-DAB2-GT17										3	1	1	1
CoLi-DAB2-GT18										4	2	1	1
CoLi-DAB2-GT19										3	1	1	1
CoLi-DAB2-GT20										3	1	1	1
CoLi-DAB2-GT21										3	1	1	1
CoLi-DAB2-GT22										3	1	1	1
CoLi-DAB2-GT23										3	1	1	1
CoLi-DAB2-GT24										3	1	1	1
CoLi-DAB2-GT25										3	1	1	1
CoLi-DAB2-GT26										5	1	2	2

Figure 2

The MHC-II DAB2 genotype reconstruction for 26 feral pigeons with more than two alleles detected. Alleles were assigned to three loci marked with orange (DAB2A), green (DAB2B), and blue (DAB2C). The observed allele configurations indicate that CoLi-DAB2-GT01-25 genotypes are formed by two haplotypes characterized by copy number variation, i.e. by one-locus haplotype (DAB2A) and three-locus haplotype (DAB2A-DAB2B-DAB2C), whereas CoLi-DAB2-GT26 genotype consists of two three-locus haplotypes. DAB1B and DAB1C alleles were not detected in individuals with two copies of one-locus DAB2A haplotype.



of the putative peptide-binding region (PBR) in humans (based on the crystallographic structure of MHC-II molecule) are indicated with large dots (•).

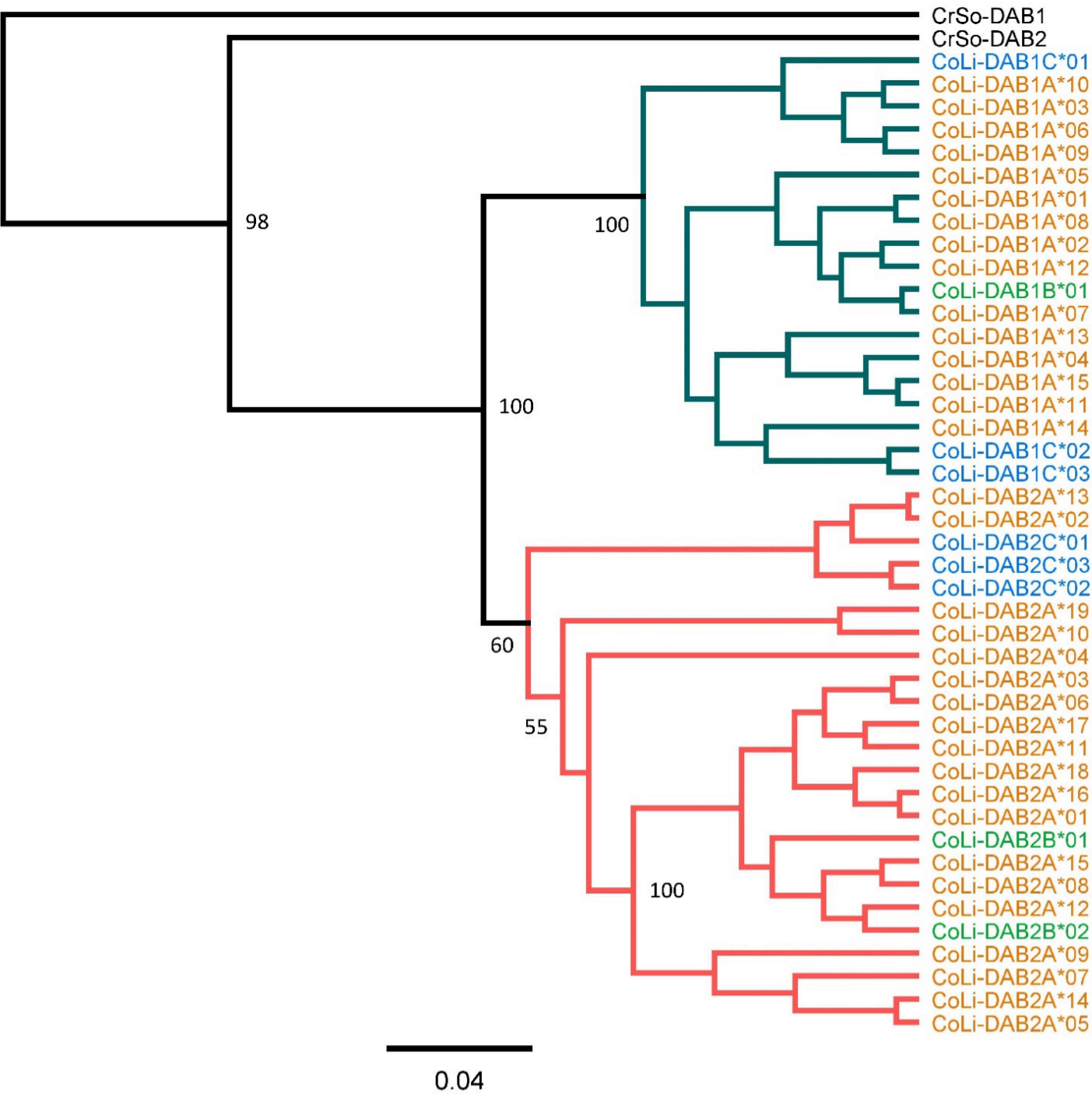


Figure 4

Consensus Bayesian topology of MHC class II exon 2 sequences in the feral pigeon. DAB1 sequences are marked in teal, while DAB2 sequences are marked in red. Alleles assigned to different gene copies were marked with orange (DAB1A/2A), green (DAB1B/2B), and blue (DAB1C/2C) node labels. Bayesian

posterior probabilities are provided for major clusters. Little tinamou (Palaeognathae: Tinamiformes) sequences CrSo-DAB1 and CrSo-DAB2 were used as outgroups.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [ESMMHCclassIIvariationinferalpigeons.docx](#)