the genetics society

www.nature.com/hdy

ARTICLE

Check for updates

Copy number variation and evolution of MHC class I and II genes in lovebirds (*Agapornis*, Psittaculidae, Psittaciformes)

Derek Kong Lam [1] and Simon Yung Wa Sin [1]✉

Gene duplication and loss play an important role in the evolution of the major histocompatibility complex (MHC). Variations in copy number and sequence diversity of MHC genes can have significant fitness consequences. Here, we characterized both MHC class I and class II genes in a group of parrots—lovebirds (*Agapornis* spp.) using cloning and sequencing, quantitative PCR, and depth-of-coverage (DoC) analysis with whole-genome re-sequencing data. We identified copy number variation in MHC class II genes, with *A. roseicollis* having a single MHCIIB gene copy, whereas *A. canus* possesses at least three gene copies. Conversely, the copy number of class I genes is invariable, with only one copy identified in each *Agapornis* species. Phylogenetic reconstructions revealed both concerted evolution and trans-species polymorphism of MHC genes. In both MHC class I and II genes, sequences from the recently diverged eye-ringed species (e.g., *A. fischeri*, *A. personatus*, and *A. nigrigenis*) and their sister species *A. roseicollis* showed an intercalating pattern with no species-specific clustering, consistent with trans-species polymorphism. In contrast, sequences from the early-diverged species (e.g., *A. canus* and *A. pullarius*) clustered by species, which is typical for avian MHC genes undergoing concerted evolution. The pattern of MHC copy number variation and modes of evolution observed are associated with the timescale of species divergence. We suggest that future studies should include both MHC class I and II genes and multiple species spanning a range of divergence time to enhance our understanding of the evolution of avian MHC diversity.

Heredity; https://doi.org/10.1038/s41437-025-00815-4

INTRODUCTION

Members of the major histocompatibility complex (MHC) multigene family are among the most polymorphic genes in jawed vertebrates (Gaigher et al. 2023; Robinson et al. 2017). MHC class II molecules, which are heterodimers encoded by distinct MHCIIA and MHCIIB genes, bind peptide fragments from extracellular pathogens for presentation at the cell surface (Brown et al. 1988; Joffre et al. 2012). MHC class I molecules instead typically detect peptide fragments from intracellular pathogens, although cross-presentation of extracellular antigens is also known to occur (Joffre et al. 2012). The ability of MHC genes to bind diverse pathogenic antigens relies on the peptide-binding sites (PBS), which experience intensive site-specific pathogen-mediated selection (Lam et al. 2023; Minias et al. 2018; Radwan et al. 2020; Sin et al. 2014; Spurgin and Richardson 2010). Given its central role in adaptive immunity, coevolutionary arms races with pathogens have driven rapid evolution in the MHC gene family, with different evolutionary patterns observed across different clades (Radwan et al. 2020).

The evolutionary pattern and history of MHC genes differ markedly between birds and mammals (Minias et al. 2023). Mammalian MHC genes tend to cluster by locus rather than species, reflecting an ancient gene duplication history where duplication preceded speciation (Nei and Rooney 2005; Takahashi et al. 2000). This evolutionary history has led to widespread trans-species polymorphism, where ancestral genetic variants are shared across species and are maintained as distinct MHC lineages (Azevedo et al. 2015; Hughes and Nei 1990). However, not all mammals share the same set of orthologous MHC genes, as some MHC lineages have been independently lost or duplicated in different taxa (Hughes and Nei 1989; Nei and Rooney 2005). This pattern of MHC evolution observed in mammals is consistent with the model of birth-and-death evolution (Nei and Rooney 2005).

In addition to the birth-and-death model typical of mammals, concerted evolution is proposed as an alternative mechanism to explain the evolution of MHC genes in birds (Edwards et al. 1999; Nei and Rooney 2005). Concerted evolution refers to the process whereby genes within a multigene family become homogenized through crossing-over and gene conversion, with early evidence for this phenomenon coming from studies in mammals (Geliebter and Nathenson 1987). Early analyses of MHC genes in several Passeriformes and Galliformes found species-specific clustering of MHC sequences with little evidence of orthologous relationships between species, suggesting that recent duplications or an increased rate of homogenization between loci have occurred (Edwards et al. 1999; Wittzell et al. 1999). However, a later study in the order Strigiformes revealed a different story, with MHCIIB lineages not showing species-specific clustering but instead resembling the locus-specific signature typically observed in mammals (Burri et al. 2008). Large-scale phylogenetic reconstructions further confirmed these findings and revealed two ancient avian class II lineages (i.e., DAB1 and DAB2) that can be traced back to the basal avian radiation (Burri et al. 2010; Goebel et al. 2017). Together, these findings suggest that evolution of the avian

[1]School of Biological Sciences, The University of Hong Kong, Hong Kong, China. Associate editor: Sara Knott. ✉email: sinyw@hku.hk

Received: 25 November 2023 Revised: 29 November 2025 Accepted: 1 December 2025
Published online: 19 December 2025

MHC is both highly dynamic and also clade-specific, with passerines potentially exhibiting a more dynamic pattern of evolution compared to other avian groups like the owls. Further investigation of MHC evolution in diverse avian clades with varying species divergence times is needed to better understand the evolutionary dynamics of this important gene family (Burri et al. 2008; Goebel et al. 2017; Hess and Edwards 2002).

A wider sampling across birds can also help clarify how rearrangements to MHC genomic architecture might influence the sequence diversity of constituent gene family members and how this diversity, in turn, might correlate with differences in life history, pathogen resistance, and sexual selection across species. While pathogen-mediated selection undoubtedly plays an important role in MHC evolution, gene duplication and deletion, together with point mutation, ultimately generate the raw genetic variation upon which selection acts (e.g., Minias et al. 2020; Sin et al. 2022; Wong et al. 2022). New MHC gene copies can arise through gene duplication, potentially increasing an individual's fitness by providing a broader pathogen protection (Chappell et al. 2015; Radwan et al. 2020). However, the optimal MHC copy number is constrained by other intrinsic immunological factors such as the depletion of T-cell receptor repertoires and an increasing risk of autoimmune disease with increasing numbers of MHC gene duplicates (Lenz et al. 2015; Migalska et al. 2019). Over time, intraspecific variation in MHC copy number can also produce interspecific differences. For example, some passerines show over 30 class I gene copies compared to generally lower copy numbers in non-passerines (Minias et al. 2019b; O'Connor et al. 2016). Yet the functionality of all detected copies remains unclear, as some may be pseudogenes or non-classical MHC genes (O'Connor and Westerdahl 2021; Stervander et al. 2020). This information is important for understanding the adaptive significance and evolutionary forces driving copy number variation (O'Connor and Westerdahl 2021). Although some life-history traits (e.g., lifespan, migration behavior) are correlated with copy number variation observed in birds, the underlying mechanisms driving this association are still poorly understood (Minias et al. 2019b; O'Connor et al. 2018).

Recent advances in sequencing technologies have furthered our understanding of avian MHC genomic architecture, although MHC diversity and characteristics such as copy number and locus orthology remain poorly characterized in many groups (He et al. 2022; O'Connor et al. 2019). For example, only a handful of MHC studies have been previously undertaken in parrots, a large and diverse group of birds including many species that are currently considered threatened (Edwards et al. 1999; Hughes et al. 2008; Knafler et al. 2014). The lack of understanding of copy number variation and the evolution of MHC genes is further hampered by the high genomic complexity of this region and by technological limitations hindering our ability to resolve this complexity. MHC genotyping using traditional cloning-and-sequencing approaches or newer high-throughput sequencing methodologies is commonly applied without knowing MHC locus identity (Babik 2010; Sin et al. 2012a; Sin et al. 2012b), and alleles from multiple loci are often co-amplified by degenerate primers that target all gene copies (Burri et al. 2014). The species-specific gene copy number is thus often estimated based on the maximum number of alleles detected in any individual within the studied population (Minias et al. 2019b). This method is likely to underestimate the true copy number as it lacks information regarding zygosity. Quantitative PCR (qPCR) is an under-employed technique in this regard, which can accurately identify the presence of MHC hemizygosity in addition to estimating gene copy number (Gaigher et al. 2016; Lighten et al. 2014; Weaver et al. 2010; Wong et al. 2022). Other promising approaches are long-read sequencing (Westbrook et al. 2015) and depth-of-coverage (DoC) analysis from short-read data to estimate gene copy number from the read depth at target MHC loci relative to that of reference single-copy nuclear genes (but see Ammar et al. 2015; Viļuma et al. 2017; Wong et al. 2022). Since newly duplicated gene copies are expected to retain high sequence similarity to the parental MHC locus, reads from all descendant gene copies are expected to map to the parental sequence with high fidelity, and individuals with extra gene copies would therefore exhibit increased read depth at MHC loci relative to single-copy reference genes. This DoC approach offers several advantages relative to qPCR, such as not relying on specific primers to amplify polymorphic MHC genes and therefore reducing the risk of allelic dropout (Babik 2010; Sommer et al. 2013).

Here, we investigated the copy number variation and evolutionary history of both MHC class I and II genes in lovebirds (*Agapornis* spp.), which is a group of small African parrots that includes species popular in the global pet trade (Chan et al. 2021; Mori et al. 2020). Lovebirds provide a good model for studying MHC evolution in parrots because of their well-resolved phylogeny with species spanning a range of divergence timescales (Huynh et al. 2023). The nine extant *Agapornis* species diverged over approximately the past 10 million years. This includes four earlier-diverging species (gray-headed lovebirds *A. canus*, red-faced lovebirds *A. pullarius*, black-winged lovebirds *A. taranta*, and black-collared lovebirds *A. swindernianus*) and five more recently diverged species. The recent group consists of the rosy-faced lovebird (*A. roseicollis*) and the four eye-ringed species (Nyasa lovebirds *A. lilianae*, Fischer's lovebirds *A. fischeri*, masked lovebirds *A. personatus*, and black-cheeked lovebirds *A. nigrigenis*), with the latter group undergoing a rapid radiation within the last ~1 million years (Dilger 1960; Manegold and Podsiadlowski 2014). In this study, by employing multiple approaches (i.e., cloning and sequencing, qPCR, and whole-genome sequencing [WGS]), we aimed to (1) characterize both MHC class I and class II genes across lovebirds; (2) determine whether gene copy number variation exists among these species; and (3) investigate the selection and evolutionary patterns of both gene families. Our results provide valuable insights into MHC diversity and evolution in Psittaciformes, a group of birds that was poorly studied with respect to MHC genes.

MATERIALS AND METHODS

Sample collection and nucleic acid extraction

In total, bird tissue ($n = 30$) or blood samples ($n = 19$) were collected from museums or local breeders (Supplementary Information S1). Eighteen *A. roseicollis*, nine *A. fischeri*, eight *A. personatus*, one *A. nigrigenis*, one *A. lilianae*, three *A. pullarius*, three *A. taranta*, one *A. swindernianus*, and three *A. canus* were included in the study. Two closely related species (i.e., *Loriculus galgulus*, Blue-crowned hanging parrot, and *Bolbopsittacus lunulatus*, Guaiabero) were included as outgroups. Fresh blood samples were collected for *A. roseicollis* and *A. personatus* from local breeders and stored in Queen's lysis buffer (Seutin et al. 1991) until genomic DNA (gDNA) extraction or in RNAlater until RNA extraction. gDNA was extracted from blood or tissue using the EZNA Tissue DNA Kit (Omega, USA) following the manufacturer's protocol. Total RNA was extracted from blood samples using TRIzol Reagent (Invitrogen, USA) and quantified by the NanoDrop One Microvolume UV-Vis Spectrophotometer (Thermo Scientific, USA). One microgram of RNA was used for cDNA synthesis with the QuantiTect Reverse Transcription kit (Qiagen, Germany).

Whole-genome sequencing (WGS)

Whole-genome re-sequencing was performed on all *Agapornis* species using museum samples ($n = 36$) to facilitate primer design and DoC analysis for copy number estimation. Whole-genome library preparation and sequencing were performed following Lam and Sin (2020) and Huynh et al. (2023). In brief, gDNA was sheared to target size on a Covaris S220 platform, and the PrepX ILM 32i DNA Library Kit (TaKaRa, US) was used to prepare DNA fragment libraries of 220 bp insert size. The libraries were then quantified with qPCR (KAPA library quantification kit) and sequenced using an Illumina HiSeq 2500 instrument (High Output, PE 125 bp reads), aiming for ~5× coverage per sample (Grayson et al. 2017).

MHC primer design for genotyping

Most previous MHC studies of class I genes only targeted exon 3 for genotyping (Minias et al. 2021a). However, the peptide-binding groove of MHC class I molecules is formed by both the α1 and α2 domains that are encoded by exon 2 and exon 3, respectively, and we therefore analyzed both of these exons. We found that both MHC class I and class II genes were fragmented in the available draft genome of *A. roseicollis* (GenBank: NDXB 00000000.1; Supplementary Information S2), so we therefore limited our analysis to exons 2 and 3 of MHCI (located on scaffold NDXB01002742.1; 6259 bp) and exon 2 of MHCIIB (located on scaffold NDXB010078282.1; 1224 bp).

MHC primers were designed from multiple sequence alignment of MHC sequences extracted from the WGS data of *Agapornis* spp. from this study, as well as from publicly available genome assemblies for additional parrot species (see Supplementary File for Accession Numbers). We re-constructed MHC sequences from WGS data in different *Agapornis* species using cross-species mapping. The draft genome of *A. roseicollis* (GenBank: NDXB 00000000.1) served as the reference for mapping. In brief, the raw sequence reads of the whole-genome re-sequencing data were filtered and trimmed using Trimmomatic 0.35 (Bolger et al. 2014) with default settings. The processed reads were mapped onto the draft genome of *A. roseicollis* using the BWA-MEM algorithm in Burrows-Wheeler Aligner (BWA 0.7.17), followed by PCR duplicate removal using SAMtools 1.3.1 (Danecek et al. 2021; Li and Durbin 2009). IGV 2.4.14 was used to visualize the mapping results and to extract the consensus sequences (Thorvaldsdóttir et al. 2013).

MHCI and MHCIIB sequences from other parrot species, as well as several non-psittaciforms (Supplementary Information S3), were identified from their draft genomes with BLAST similarity searches and were aligned with the re-constructed MHC sequences of different *Agapornis* species in MEGA7 (Kumar et al. 2016). Based on this alignment, several degenerated primers were designed targeting the highly conserved regions inside or flanking MHCI exon 2, MHCI exon 3, and MHCIIB exon 2 (Supplementary Information S2 and S4) using NetPrimer (Premier Biosoft International). Primers were tested for allelic dropout using PCR and cloning and sequencing by comparing different primers amplifying the same region.

MHC genotyping and expression analysis

MHC genotyping was performed by cloning and sequencing of PCR products amplified from gDNA of *A. roseicollis* ($n = 18$ individuals), *A. fischeri* ($n = 3$), *A. personatus* ($n = 4$), *A. nigrigenis* ($n = 1$), *A. pullarius* ($n = 3$), and *A. canus* ($n = 3$). PCR was performed in 30 µl reaction mix containing 1X GoTaq Reaction Buffer, 1% DMSO, 2 mM $MgCl_2$, 0.2 mM dNTP, 0.33 µM forward primer, 0.33 µM reverse primer, 10-50 ng gDNA, and 0.75 unit GoTaq Polymerase (M3001; Promega, USA). The PCR began with incubation at 95 °C for 2 min, followed by 32 cycles of: 30 s at 95 °C, 30 s at 52–60 °C (Supplementary Information S4), and 90 s at 72 °C, and ended with a final extension step at 72 °C for 10 min. PCR products were electrophoresed on 1% agarose gel, followed by gel purification using PureLink Quick Gel Extraction Kit (Invitrogen). The purified PCR products were then ligated into a T-Vector pMD19 (TaKaRa) using a DNA ligation kit (TaKaRa). Transformation and blue-white screening were performed in *Escherichia coli* competent cells (DH5-alpha). PCR was performed on the positive clones using M13 primers (Messing 1983), and the resulting PCR products with the expected band size were sent to BGI (Hong Kong) for Sanger sequencing. The identity of the MHC sequences was confirmed by BLAST. We initially sequenced 10 clones per species, and additional clones were sequenced for species that appeared to have more MHC copies after the initial assessment. We designed multiple primer pairs that covered the region of interest, and the alleles identified were consistent across primer pairs. Based on the initial assessment result for each species, between 7 and 60 clones were sequenced for each individual. Allele identity was confirmed by identical sequences derived from at least two independent PCRs. The gene copy number estimation for each species based on cloning and sequencing was consistent with the results obtained from qPCR and DoC. Expression analysis was performed using cloning and sequencing on cDNA for *A. roseicollis* following the methods outlined above.

Phylogenetic, recombination, and selection analyses

Sequence data were aligned in MEGA7. Phylogenetic trees were constructed for MHCI exon 2, exon 3, and MHCIIB exon 2 separately using MrBayes 3.1.2 (Ronquist and Huelsenbeck 2003). JModelTest was used to ascertain the best-fitting model of nucleotide substitution, which was identified as the general time-reversible model with gamma-distributed rates in each case (Posada 2008). A Markov chain Monte Carlo (MCMC) sampling was run for 30,000,000 iterations, with the first 25% of tree samples discarded as burn-in. Convergence was assessed from the average standard deviations of split frequencies (0.0008, 0.005, and 0.002, for MHCI exon 2, MHCI exon 3, and MHCIIB exon 2, respectively). MHC sequences of other parrots (including budgerigar *Melopsittacus undulatus*, Puerto-Rican parrot *Amazona vittata*, Yellow-billed amazon *Amazona collaria*, etc.) and other birds (including *Gallus gallus*, *Coturnix japonica*, *Petroica australis*, *Anser anser*) were included for the phylogenetic analysis. Sequences from *L. galgulus* and *B. lunulatus*, which were obtained from cloning and sequencing using the *Agapornis* MHC primers, were also included in phylogenetic analyses.

We used Genetic Algorithm for Recombination Detection (GARD) implemented in Datamonkey to detect recombination breakpoints for each species prior to running selection inference (Kosakovsky Pond et al. 2006). Sites subject to positive selection and/or purifying selection were detected using four methodologies, with the partitioned recombination tree input for the analysis. MEME (Murrell et al. 2012), which detects sites under episodic positive selection, and FEL (Kosakovsky Pond and Frost 2005), FUBAR (Murrell et al. 2013), and SLAC (Kosakovsky Pond et al. 2006) methods, which all detect sites under pervasive positive or negative selection, were each implemented in Datamonkey (Weaver et al. 2018). Putative PBS and non-PBS codons were defined based on the crystallographic structure of human MHC class I and class II molecules (Brown et al. 1993; Saper et al. 1991) to contrast selection within versus outside of the peptide-binding region. The ω ratio [non-synonymous (d_N) over synonymous (d_S) substitution, d_N/d_S] was calculated in MEGA7. An ω ratio larger than 1 indicates positive selection, while lower than 1 indicates purifying selection. Selection analyses were performed for five *Agapornis* species (*A. roseicollis*, *A. fischer*i, *A. personatus*, *A pullarius*, and *A. canus*) separately for MHCIIB. For MHCI, we only performed this analysis on *A. roseicollis* due to limited samples for other species. We also performed an interspecific selection analysis using all *Agapornis* sequences identified.

Copy number variation analysis

Quantitative PCR. Since only MHCIIB was found to have CNV, we focused on MHCIIB exon 03 for copy number variation using qPCR in the analysis of copy number variation. The immunoglobulin-like exon 03 of MHCIIB genes is relatively conserved compared to the highly variable peptide-binding exon 2; therefore, targeting this exon for qPCR can increase the likelihood of designing universal primers that will amplify all MHCII gene sequences in an individual. We aligned *Agapornis* MHCIIB exon 03 sequences obtained from cloning and sequencing of both gDNA and cDNA with other parrot species using MEGA7, and designed a qPCR primer pair flanking a conserved region that was specific to exon 03 of *Agapornis* species. Both β-actin (*ACTB*) and *GAPDH*, which are routinely employed as qPCR standards and which are expected to occur as single-copy nuclear genes, were used as reference genes in the qPCR assay. To construct a standard, we cloned the amplicon into the T-Vector pMD19 and linearized using ScaI-HF (NEB). The amplification efficiencies of these primers were between 85.01-91.5%. Only samples with unfragmented gDNA were used for qPCR. qPCR was performed in triplicate for each sample in a 15-µl reaction mix containing 2X iTaq Universal SYBR Green Supermix (Bio-Rad, US), 0.4 µM forward primer, 0.4 µM reverse primer (Supplementary Information S4), and 20 ng gDNA using CFX96 Torch Real-Time PCR Detection System (Bio-Rad, US). The thermal cycle program consisted of 01 cycle of 2 min at 95 °C, followed by 40 cycles of 5 s at 95°C and 10 s at 60 °C, and ended with a melt curve analysis ramping from 65 to 95 °C with 0.5 °C/5 s increment. We used the ΔΔCq method to calculate gene copy number (CN) following the equation $CN = 2 \times 2^{-\Delta\Delta C_q}$ (Weaver et al. 2010), where the calibrator was set to be one of the *A. roseicollis* individuals.

Depth-of-coverage (DoC) analysis. Only one locus was identified for each of MHCI and MHCIIB genes in the publicly available genome assembly for *A. roseicollis*, with each gene placed on a separate, short scaffold (GenBank: NDXB 00000000.1; 63407 and 71984 nt long, respectively; Supplementary Information S2). The fragmented MHC region of this assembly could affect the mapping results and, hence, the estimation of DoC and inference of gene copy number. We therefore used an alternative, high-quality *A. roseicollis* genome assembly (in prep) as the reference genome for mapping. We performed cross-species read mapping of all WGS data to this *A. roseicollis* reference using NextGenMap v0.5.5 (Sedlazeck et al. 2013). Reads mapping onto target exons of MHCI, MHCIIA, and MHCIIB and reference genes ACTB, GAPDH, and ribosomal protein 32 (RLP32) were

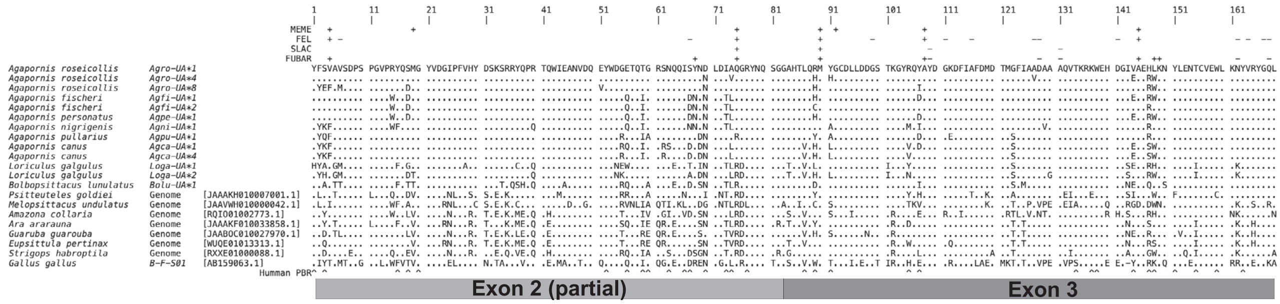

Fig. 1 Amino acid sequence identity for the MHC class I exon 2 and 3 (partial) in Psittaciformes. Only selected sequences from each species are shown here. See Supplementary Information S5 for alignment of all identified MHCI sequences. Sequences other than Agapornithinae (i.e., *Agapornis*, *Loriculus*, and *Bolbopsittacus*) were extracted from draft genomes available online. The accession numbers for the contig/scaffold/chromosome containing MHC class I exon 2 and 3 are provided in the supplementary file. The peptide-binding sites (PBS; ^) were defined according to human HLA molecules (Brown et al. 1993; Saper et al. 1991). Interspecific codon-based selection analyses were only performed by inputting all *Agapornis* sequences using FEL, SLAC, FUBAR, and MEME. "+" indicates positively selected sites while "−" indicates negatively selected sites ($p < 0.10$). "." = identical amino acid with the *Agro-UA*1*.

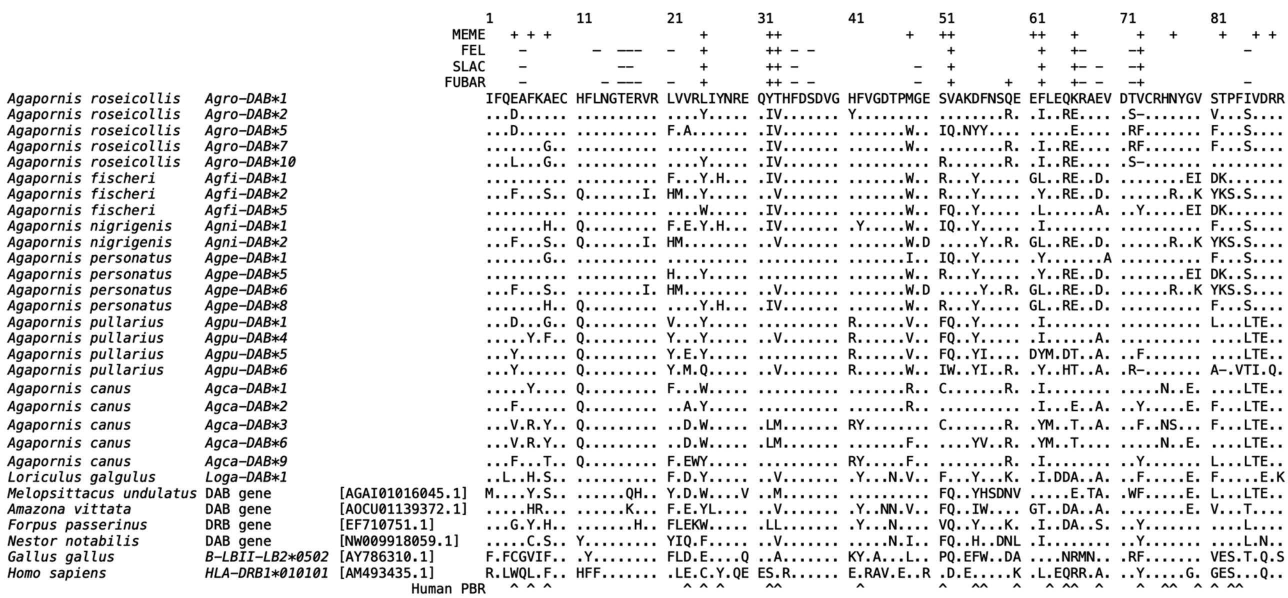

Fig. 2 Amino acid sequence identity for the MHC class II exon 2 in Psittaciformes. Only selected sequences from each species are shown here. See Supplementary Information S6-7 for alignments of all identified MHCIIB sequences. All sequences from Agapornithinae (i.e., *Agapornis* and *Loriculus*) were from cloning-and-sequencing data, while the sequences from *A. taranta* were re-constructed using whole-genome re-sequencing data. Sequences other than Agapornithinae (i.e., *Agapornis* and *Loriculus*) were extracted from draft genomes available online. The accession numbers for the contig/scaffold/chromosome containing MHC class II exon 2 are provided in the Supplementary File. The peptide-binding sites (PBS; ^) were defined according to human HLA molecules (Brown et al. 1993; Saper et al. 1991). Interspecific codon-based selection analyses were only performed by inputting all *Agapornis* sequences using FEL, SLAC, FUBAR, and MEME. "+" indicates positively selected sites while "−" indicates negatively selected sites ($p < 0.10$). "." = identical amino acid with the *Agro-DAB*1*.

counted using SAMtools 1.3.1. The copy number for each individual was calculated using the ΔΔCq approach, where the calibrator was set to be one of the *A. roseicollis* individuals (Weaver et al. 2010; Wong et al. 2022). We also investigated MHCIIA copy number variation using DoC.

Research ethics statement

This study received approval from the Animal Research Ethics Committee (4749-18) of the University of Hong Kong and the Department of Health [(18-674) in DH/SHS/8/2/3 Pt. 17] of the HKSAR Government.

RESULTS

Isolation of MHC genes from gDNA and cDNA using PCR, cloning, and sequencing

We sequenced a total of 1002 clones in 32 individuals, with 352 clones sequenced for MHCI (mean = 11 clones per individual) and 650 clones for MHCIIB (mean = 20.3 clones per individual). Among *Agapornis* spp., the total number of unique nucleotide sequences identified for MHCI and MHCIIB genes were 28 and 49, respectively (Figs. 1 and 2, Supplementary Information S5–S7). Some MHCI nucleotide sequences (e.g., *Agro-UA*5* and *Agro-UA*6*, *Agfi-UA*1* and *Agfi-UA*4*, *Agpu-UA*1* and *Agpu-UA*3*, etc.; Supplementary Information S5) shared the same amino acid translation, resulting in 23 unique MHCI amino acid sequences. No internal stop codons or frameshifts were observed, suggesting that all sequences are putatively functional. Transcription analysis was only performed for *A. roseicollis*. In individuals with both gDNA and cDNA available, the genotypes were consistent between gDNA and cDNA, confirming the expression of the MHCI and MHCIIB alleles in *A. roseicollis*.

The number of MHCI sequences isolated from cloning and sequencing was 1–2 per individual in the investigated species (i.e., *A. canus*, *A. pullarius*, *A. nigrigenis*, *A. personatus*, *A. fischeri*, and *A. roseicollis*), suggesting the presence of a single MHCI locus in the

Table 1. Summary of the genotyping results of MHC class I exon 2 + 3 and class II exon 2 in *Agapornis*.

	Class I												Class II							
		Exon 2 + intron 2 + exon 3 (partial)			Exon 2 (partial)				Exon 3					Exon 2						
	No. of individual	No. of nt allele	No. of aa allele	No. of nt allele per individual (mean)	No. of nt amplified	No. of codons amplified	Variable nt site	Variable aa site	No. of nt amplified	No. of codon amplified	Variable nt site	Variable aa site	No. of individual	No. of nt allele	No. of aa allele	No. of nt allele per individual (mean)	No. of nt amplified	No. of codon amplified	Variable nt site	Variable aa site
A. roseicollis	18	11	10	1–2 (1.83)	249	82	12	8	276	91	15	7	18	14	14	1–2 (1.78)	267–270	88–89	38	22
A. fischeri	3	4	3	2 (2)	249	82	10	5	276	91	9	6	3	8	8	2–3 (2.67)	270	89	60	30
A. personatus	3	2	2	1–2 (1.33)	249	82	2	2	276	91	8	5	4	8	8	2–3 (2.5)	270	89	62	31
A. nigrigenis	1	1	1	1 (1)	249	82	/	/	276	91	/	/	1	2	2	2 (2)	270	89	45	26
A. pullarius	3	4	3	1–2 (1.67)	249	82	1	1	276	91	5	2	3	8	8	3–4 (3.67)	264–270	87–89	55	28
A. canus	3	6	4	2 (2)	249	82	2	2	276	91	3	2	3	9	9	5 (5)	270	89	50	26

nt nucleotide, *aa* amino acid.

genus *Agapornis* (Table 1). However, the estimated number of MHCIIB genes varied among species. *A. roseicollis* and *A. nigrigenis* had 1–2 sequences isolated per individual. Up to three or four sequences were found per individual in *A. fischeri* (2–3 sequences), *A. personatus* (2–3 sequences), and *A. pullarius* (3–4 sequences), suggesting a minimum of two loci in each of these species, and five sequences per individual were found for *A. canus*, providing evidence for at least three gene copies (Table 1). Despite large sequencing efforts for *A. canus* MHCIIB genes (no. of clones per individual = 48), there were still sequences (i.e., 1–2 per individual) that were not verified by two independent PCRs in this species only.

Copy number variation

We used both qPCR and in silico DoC analyses to further confirm the interspecific copy number variation identified from cloning and sequencing of MHCIIB PCR products. The qPCR analysis of MHCIIB exon 3 indicated a 2.8-fold or 3.6-fold higher gene copy number in *A. fischeri* compared to *A. roseicollis*, depending on whether ACTB or GAPDH was used as the reference gene, respectively (Fig. 3A, B). A similar pattern was observed *in A. personatus*, with 1.79-fold or 1.27-fold higher MHCIIB copy number than *A. roseicollis* (Fig. 3A, B). The variation within species could be due to technical variation or potential intraspecific copy number variation of MHCIIB genes. The copy number calculated based on qPCR for *A. canus* (4.45 or 5.73-fold higher than *A. roseicollis*; Fig. 3A, B) was in line with the cloning-and-sequencing result that shows a much higher copy number in *A. canus* than in *A. roseicollis*.

The DoC analysis using low-coverage WGS data was concordant with the cloning-and-sequencing and qPCR results (Fig. 3C, D), indicating copy number variation of MHCIIB genes among *Agapornis* species. Estimated MHCIIB copy numbers in all other *Agapornis* species were at least 2-fold higher than *A. roseicollis*, with *A. canus* having the highest copy number, which was >5-fold higher than *A. roseicollis*. Furthermore, since MHCIIA genes were found to be ~540 bp upstream of the MHCIIB in the genome assembly of *A. roseicollis* (Supplementary Information S8), and MHCIIA and MHCIIB genes usually appear as a unit in different avian genomes, we also examined whether MHCIIA showed a similar pattern of copy number variation as MHCIIB among *Agapornis* species. However, DoC analysis showed a similar copy number for MHCIIA genes for all species (~1; Fig. 3C, D). DoC-based estimates of MHCI gene copy number (Fig. 3C, D) also indicate a single MHCI locus is present in each *Agapornis* species, which agrees with the results from the cloning-and-sequencing approach.

Sequence polymorphism

Sequence polymorphism was higher in MHCIIB than MHCI genes among *Agapornis* species (Table 1, Figs. 1 and 2, Supplementary Information S5–S7). MHCIIB contained 22-30 non-synonymous mutations within the amplified exon 2 region (88 amino acids in total), whereas MHCI contained only 1–8 non-synonymous mutations within the longer amplified exon 2 and 3 region (169 amino acids in total). In MHCIIB, amino acid deletions were also found at position 73 for *A. roseicollis* and at positions 73 and 82 for *A. pullarius* (Supplementary Information S7). The higher number of variable sites in *A. fischeri*, *A. personatus*, *A. pullarius*, and *A. canus* than in *A. roseicollis* could potentially be attributed to a greater number of MHCIIB genes (Table 1).

Recombination and selection

Two recombination breakpoints were detected in exon 3 of MHCI sequences (positions 222 and 274) for *A. roseicollis* using GARD (Fig. 1; Supplementary Information S9). In MHCIIB, two to four recombination breakpoints were detected in *A. roseicollis*, *A. fischeri*, *A. personatus*, and *A. canus* (Fig. 3; Supplementary

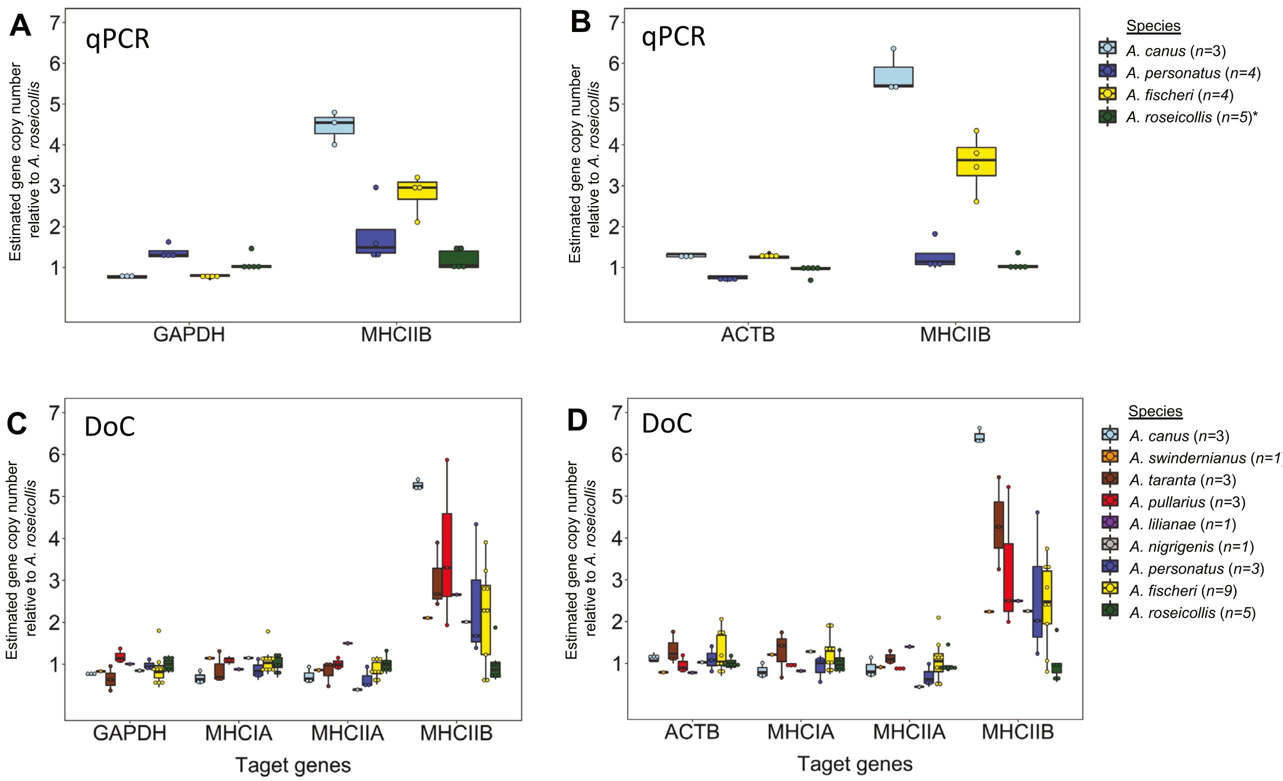

Fig. 3 The copy number variation (CNV) of MHC genes among *Agapornis* species. The CNV was estimated using **A**, **B** qPCR and **C**, **D** depth-of-coverage (DoC) analysis. For qPCR analysis, **A** GAPDH or **B** ACTB genes were used as the invariant reference gene for normalization. **C**, **D** The DoC analysis was performed using low-coverage (~4×) whole-genome re-sequencing data. The invariant reference genes used for normalization were **C** GAPDH + RLP32 or **D** ACTB + RLP32. Each dot represents one individual.

Information S9). *A. fischeri* and *A. personatus* shared the same three breakpoints (at positions 78, 141, and 200). No recombination event was detected for *A. pullarius*.

Signatures of positive selection were detected for both MHCI and MHCIIB genes. Positive selection was detected in the PBS of both MHCI and MHCII genes of *A. roseicollis* ($\omega = 4.16$ in MHCI, $\omega = 1.71$ in MHCIIB; Supplementary Information S10) but not in non-PBS codons ($\omega = 0.56$ in MHCI, $\omega = 0.70$ in MHCIIB; Supplementary Information S10). However, the number of positively selected sites was generally lower for MHCI than for MHCIIB genes. For example, only 3 positively selected sites were detected in MHCI in *A. roseicollis* using MEME, compared to 6 sites detected in MHCIIB (Supplementary Information S5, S7, and S11). For the MHCIIB genes in different *Agapornis* species, the ω ratios were similar ($\omega = 2.021$–2.761 in PBS, $\omega = 0.674$–0.804 in non-PBS; Supplementary Information S10). *A. roseicollis* showed a relatively lower ω ratio at PBS than other species ($\omega = 1.71$), perhaps due to the presence of a single MHCIIB locus. Similar patterns were observed for interspecific selection analysis, e.g., more positively selected sites in MHCIIB sequences compared to MHCI (Figs. 1 and 2). Compared to the intraspecific analysis, some sites appear to be under positive selection before species divergence.

Phylogenetic analysis

Sequences from *A. canus* and *A. pullarius* each formed a monophyletic cluster to the exclusion of other *Agapornis* spp. for MHCI exon 2 (Fig. 4). Sequences from the eye-ringed species (i.e., *A. fischeri*, *A. personatus*, and *A. nigrigenis*) as well as *A. roseicollis* were instead paraphyletic or polyphyletic, with one of the sequences (i.e., Agfi-UA*3) intercalated within the non-eye-ringed group comprising *A. canus*, *A. taranta*, *A pullarius*, and *A. nigrigenis*. However, the degree of intercalating observed within the eye-ringed species and *A. roseicollis* was much higher in exon 3 sequences than in exon 2 sequences (Figs. 4 and 5). The MHCI exon 3 tree also showed a clear separation between the recently diverged species (e.g., *A. roseicollis*, *A. fischeri*, *A. personatus*, and *A. nigrigenis*) and the early-diverged species (e.g., *A. canus* and *A. pullarius*).

The phylogenetic pattern of MHCIIB exon 2 was similar to that of MHCI exon 3, with sequences from *A. roseicollis* and the eye-ringed species being intercalated with each other while sequences from *A. canus* and *A. pullarius* each showed species-specific clustering (Fig. 6). Since *A. roseicollis* only had one MHCIIB locus, as suggested by cloning and sequencing, qPCR, and DoC analysis, locus-specific clustering of MHCIIB exon 2 was unlikely.

A previous phylogenetic analysis by Goebel et al. (2017) identified only sequences belonging to the DAB1 ancestral lineage of MHCIIB genes in Psittaciformes, but it was based upon a limited sample size. Further support for this finding can be found in recent phylogenetic studies using a larger number of parrot genomes (Burri et al. 2008; Minias et al. 2023). Here, we show that all identified MHCIIB sequences in *Agapornis* belong to the DAB1 lineage (Supplementary Information S12). Despite the clear phylogenetic support for DAB1-lineage genes in *Agapornis*, only 9 of 16 lineage-specific nucleotides previously identified by Burri et al. (2008) and Goebel et al. (2017) were conserved in this group (Supplementary Information S13). The presence of only DAB1 lineage genes in *Agapornis* was further confirmed by the inclusion of MHCIIB exon 3 sequences extracted from the genomes of other Psittaciformes. All sequences across Psittaciformes were more closely related to DAB1 than DAB2 sequences (Supplementary Information S13).

DISCUSSION

We characterized both MHCI and MHCIIB genes in the parrot genus *Agapornis* using cloning and sequencing, qPCR, and WGS

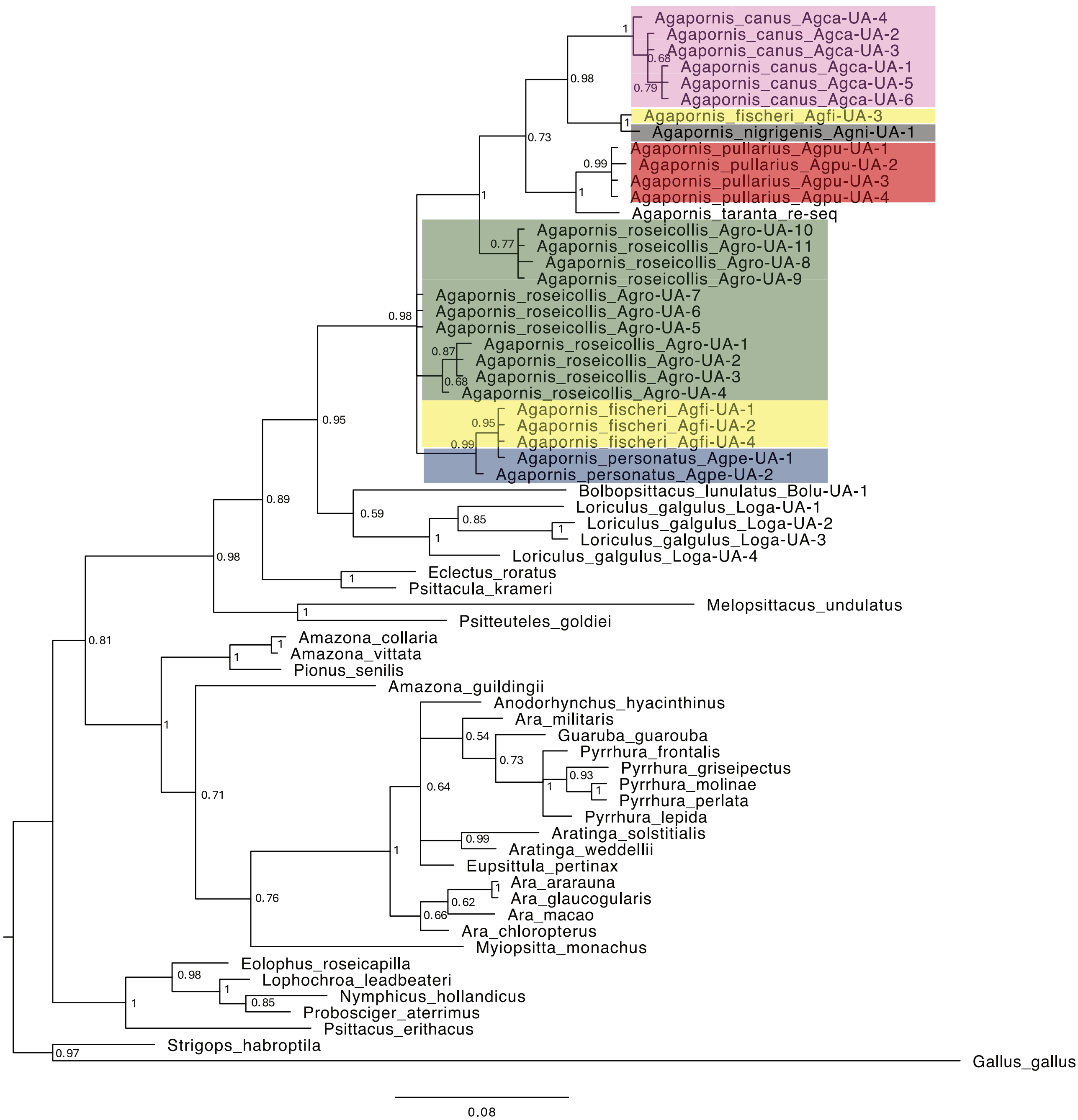

Fig. 4 Phylogenetic tree of MHC class I exon 2 sequences in Psittaciformes. Sequences other than Agapornithinae (i.e., *Agapornis*, *Loriculus*, and *Bolbopsittacus*) were extracted from the draft genomes available online. The accession numbers for the contig/scaffold/chromosome containing MHC class I exon 2 are provided in the Supplementary File. The *Agapornis taranta* sequence was extracted from the mapping result of the whole-genome re-sequencing data. Sequences of different *Agapornis* species identified from cloning and sequencing are marked with boxes of different colors. The numbers by the tree nodes indicate the posterior probabilities.

data. We demonstrated copy number variation of MHCIIB loci and revealed the evolutionary dynamics of MHC genes within this genus. The evolutionary patterns we observed between the studied species suggest different mechanisms are at play. The early-diverged *Agapornis* species, including *A. canus* and *A. pullarius*, exhibited species-specific clustering of both MHCI and MHCIIB sequences, which is indicative of concerted evolution acting to homogenize allelic sequences originating from duplicated loci within each species. However, the more recently diverged species (e.g., *A. roseicollis*, *A. fischeri, A. personatus*, and *A. nigrigenis*) showed trans-species polymorphism of MHCIIB genes, likely due to incomplete lineage sorting. Introgression might also play a role, as hybridization resulting in fertile offspring has been reported for eye-ringed species from lovebird breeders. Additionally, the birth-and-death model explains the observed MHCIIB gene copy number variation across all studied *Agapornis* species, independent of their divergence times.

Positive selection on MHC genes: class I vs. class II

Non-passerine birds are generally shown to experience stronger selection at MHC class II genes than class I genes (Minias et al. 2018). For example, more positively selected sites are reported for

Fig. 5 Phylogenetic tree of MHC class I exon 3 sequences in Psittaciformes. Sequences other than Agapornithinae (i.e., *Agapornis*, *Loriculus*, and *Bolbopsittacus*) were extracted from draft genomes available online. The accession numbers for the contig/scaffold/chromosome containing MHC class I exon 3 are provided in the Supplementary File. Sequences of different *Agapornis* species identified from cloning and sequencing are marked with boxes of different colors. The numbers by the tree nodes indicate the posterior probabilities.

the PBS-encoding exon 2 of MHCIIB genes than for MHCI exon 3 in flamingo and prairie grouse (Gillingham et al. 2016; Minias et al. 2016). Consistent with this trend, *A. roseicollis* showed more positively selected sites in MHCIIB than MHCI, as well as a higher number of variable amino acid sites and alleles in MHCIIB (Table 1).

Differences in the selection intensity acting upon class I and II genes are thought to reflect their differential roles in defending the host from intracellular and extracellular pathogens (Minias et al. 2018, 2021b). Stronger selection on class II genes of non-passerines may be partly driven by their larger body size, which can support a richer extracellular parasite fauna, although lovebirds are not exceptionally large (Forshaw 2010; Minias et al. 2018; Morand and Poulin 1998). Alternatively, the colonial habit of *A. roseicollis* could contribute to elevated class II selection (Minias et al. 2017; Ndithia et al. 2007). A recent meta-analysis in birds suggests a positive relationship between selection at MHC class II genes and the degree of sociality by showing that colonial species have a higher rate of non-synonymous substitutions than solitary species (Minias et al. 2017). However, it is unknown if class I genes also show elevated non-synonymous substitutions in colonial species. Our methodological approach for selection analysis has some limitations, however. The sample sizes for some species limited the power of our intraspecific selection analyses. Additionally, the interspecific analysis could be influenced by the heterogeneity in loci and sequences among species, as we analyzed data from multiple loci concurrently.

Copy number variation of MHCIIB genes in *Agapornis*

We determined the copy number of MHC genes in *Agapornis* using three distinct methods: cloning and sequencing, qPCR, and DoC. For the MHCIIB genes, cloning and sequencing targeted exon 2, while qPCR targeted exon 3. The DoC analysis covers all exons of the MHCIIB genes. These complementary approaches yielded

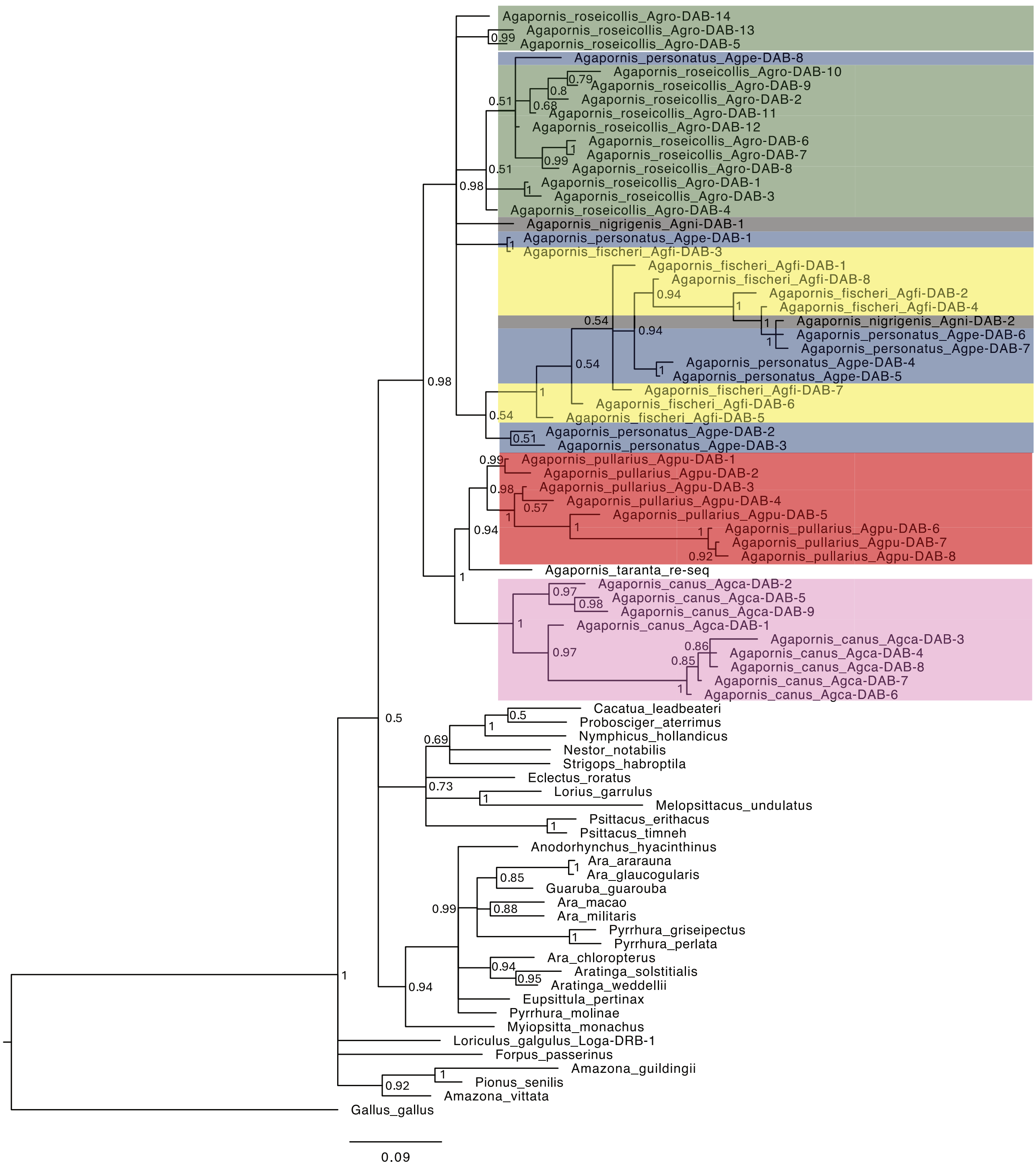

Fig. 6 Phylogenetic tree of MHC class II exon 2 sequences in Psittaciformes. Sequences other than Agapornithinae (i.e., *Agapornis*, *Loriculus*, and *Bolbopsittacus*) were extracted from draft genomes available online. The *Agapornis taranta* sequence was extracted from the mapping result of the whole-genome re-sequencing data. Sequences of different *Agapornis* species identified from cloning and sequencing are marked with boxes of different colors. The numbers by the tree nodes indicate the posterior probabilities.

consistent estimates of a single MHCIIB gene copy in *A. roseicollis*, whereas other *Agapornis* species harbor multiple copies of MHCIIB.

Ancestral birds are hypothesized to have low MHC gene copy numbers, with both MHC class I and II genes later becoming highly duplicated in different passerine lineages, but remaining low in most non-passerines (Minias et al. 2019b). Consistent with this general observation, we found evidence for only a single class I gene in the genus *Agapornis*, and a moderate number of class II genes ranging from 1 to at least 3 copies. This MHCIIB copy number is still relatively low compared to passerines (Bollmer et al. 2010). For example, common yellowthroats (*Geothlypis trichas*) have an average of 27.7 class II alleles per individual (Bollmer et al. 2012). In contrast, both the green-rumped parrot (*Forpus passerinus*) and kakapo (*Strigops habroptilus*) were shown to have only one or two copies of the MHC class II gene, respectively (Hughes et al. 2008; Knafler et al. 2014). Similarly, most parrot genomes currently available in GenBank contain low gene copy numbers (<3), although it is possible that multiple MHC genes are artificially collapsed into a single copy during genome assembly from short-read data. One notable exception to the generally low MHC copy number estimates for parrots is the red-crowned parrot (*Amazona viridigenalis*), where 8.5 alleles on average and a maximum of 12 alleles per individual were recorded for MHCI genes, signaling a minimum of six gene copies (Knafler et al. 2016). Taken together, it appears that MHC copy number is generally low within parrots, but there are exceptions, such as in *Agapornis* and *Amazona*. Further studies are required to understand how common MHC gene expansion is in parrots.

In *Agapornis*, the more recently diverged species, specifically the eye-ringed species (i.e., *A. fischeri*, *A. personatus*, *A. nigrigenis*, and *A. lilianae*), along with *A. roseicollis*, exhibit lower copy numbers of MHCIIB genes. *A. roseicollis* likely has only one copy, while *A. personatus*, *A. fischeri*, and *A. nigrigenis* each appear to have at least two copies. In contrast, *A. canus*, which is the sister to all other *Agapornis* species, possesses the highest number of MHCIIB gene copies, with at least three or more genes. Other species, such as *A. taranta* and *A. pullarius*, exhibit intermediate levels of gene copies between *A. canus* and the eye-ringed species. Based on the phylogeny of *Agapornis* (Huynh et al. 2023), two possible scenarios to explain the MHC class II gene copy number change are (1) the common ancestor of *Agapornis* had many gene copies like *A. canus* but some copies were lost during evolution; and (2) the common ancestor of *Agapornis* had an intermediate gene copy number, and extra copies were gained in *A. canus* while only one persisted in *A. roseicollis*. A chromosome-level genome assembly of *Loriculus galgulus* (Osipova et al. 2024), which is a sister species to *Agapornis*, contains a single MHCIIB gene, further suggesting that the common ancestor to *Agapornis* was also unlikely to possess an expanded MHC gene repertoire, and the second scenario is more likely.

The question of what evolutionary forces underlie copy number variation of MHC genes has been frequently raised; however, there are few empirical studies addressing this question (Minias et al. 2019b, 2020). The gain or loss of MHC gene copies in *Agapornis* might be linked to demographic history or founder effects within this group. For example, successive founder effects could influence the copy number of MHC genes in each founder population during their colonization of and dispersal throughout Africa (Huynh et al. 2023; Manegold and Podsiadlowski 2014; Schweizer et al. 2011). But this requires intraspecific copy number variation before speciation and subsequent fixation of gene copy number in different species. Alternatively, and not mutually exclusive with neutral processes, natural selection might shape the MHCIIB copy number variation in *Agapornis* after speciation (Minias et al. 2018; Kloch et al. 2010; Wegner et al. 2003).

The evolutionary dynamics of avian MHC genes: species divergence timescale matters

The relationship between species divergence time and evolutionary pattern (e.g., species-specific clustering) was observed in both MHC class I and II genes of *Agapornis*, reflecting a dynamic interplay between trans-species polymorphism and concerted evolution. Most of our understanding of the avian MHC evolution comes from studying class II genes, and we have limited knowledge on the evolution of class I genes in birds in a broader sense, although many studies focus exclusively on MHCI in passerines (Minias et al. 2018, 2019b, 2021a; O'Connor et al. 2016). In fact, evidence of both trans-species polymorphism (e.g., in a relatively short timescale in the family Accipitridae) and species-specific clustering (e.g., in a relatively long timescale in birds of prey) exists in MHC class I genes (Alcaide et al. 2008; Minias et al. 2019a). It is likely that the evolution of avian MHC class I genes is also highly dynamic, as in class II genes.

By comparing the interspecific and intraspecific selection analysis, we found that some sites are likely under positive selection before species divergence, a pattern consistent with the long-term maintenance of alleles via trans-species polymorphism. Our results further shed light on the timescales at which the diversification and homogenization of avian MHC genes have occurred by identifying alleles in the genus *Agapornis* with estimated divergence times ranging from <1 to >20 million years ago (mya; Huynh et al. 2023; Schweizer et al. 2011). Several studies have reported the existence of trans-species polymorphism at different taxonomic levels and timescales, e.g., between two genera of the tribe Tetraonini (Centrocercus and Tympanuchus, ~6–8 Mya; Minias et al. 2016), two genera of the family Procellariidae (Halobaena and Pachyptila, ~25 Mya; Strandh et al. 2011), three sister clades in the order Passeriformes (Fringillinae, Carduelinae, and Emberizidae, ~26 Mya; Minias et al. 2021b), as well as within the family Accipitridae (~20–30 Mya; Minias et al. 2019a) and the genus *Anas* (~0.1–10 Mya; Qin et al. 2021). Trans-species polymorphism in these taxa could arise through strong balancing selection that maintains ancestral MHC lineages across speciation events. Alternatively, but not mutually exclusive, a slower rate of concerted evolution could allow for the persistence of these ancient lineages in some avian clades (Goebel et al. 2017). The contrasting patterns we observed in *Agapornis* highlight that the evolutionary mode—whether dominated by trans-species polymorphism or concerted evolution—could be a key factor shaping MHC diversity, with the timescale of species divergence being a critical determinant.

Evolution of the MHCIIB in parrots

Analysis of MHC class II genes in *Agapornis*, together with publicly available sequences from other Psittaciforms, illustrates that parrots only retain the ancestral avian *DAB1* lineage of MHCIIB genes and have lost the *DAB2* lineage, consistent with a previous study (Goebel et al. 2017). Within *Agapornis*, there were multiple events of duplication and/or loss of *DAB1* gene copies. In mammals, the duplication of MHCIIB genes is often coupled with the duplication of the MHCIIA gene, as they are arranged in tandem in the genome (Bontrop 2006). Similar tandem arrays of MHCIIA and MHCIIB genes occur in some avian genomes (e.g., Golden eagle, *Aquila chrysaetos chrysaetos*, GCA_900496995.2; ruff, *Calidris pugnax*, GCF_001431845.1; Chen et al. 2015). However, *Agapornis* spp. appear to have an invariant MHCIIA copy number, whereas MHCIIB copy numbers vary moderately across species. Some avian species exhibit a genomic arrangement with a single MHCIIA gene located adjacent to multiple MHCIIB genes (e.g., Tufted duck, *Aythya fuligula*, GCF_009819795.1; Ren et al. 2011), which might also be the case for *Agapornis*. MHCIIB molecules from the two ancestral lineages (i.e., *DAB1* and *DAB2*) require different MHCIIA molecules coded by different genes (i.e., *DAA1* and *DAA2*, respectively, Minias et al. 2023). Since parrots seem to have only the *DAB1* lineage of MHCIIB genes, it is likely that a single MHCIIA molecule can effectively form heterodimers with the different MHCIIB molecules encoded by various genes.

CONCLUSIONS

Our study provides valuable insights into the diversity and evolution of MHC genes in parrots, a group with very limited knowledge on their MHC genes. We discovered copy number variation and distinct evolutionary patterns in the MHC genes of *Agapornis* and revealed how the species divergence timescale could potentially relate to MHC evolution in birds. Furthermore, the level of MHC diversity and copy number variation observed in *Agapornis* makes this group a promising system for future ecological and behavioral studies that investigate pathogen resistance and mate choice in the context of MHC sequence variation. We advocate for future studies to include both MHC class I and II genes and multiple species with a broad range of divergence times to enhance our understanding of the evolution of this important immune gene family in vertebrates.

DATA AVAILABILITY

The MHC sequences from this study were assigned the GenBank accession numbers OR843061−OR843142.

REFERENCES

Alcaide M, Edwards SV, Cadahía L, Negro JJ (2008) MHC class I genes of birds of prey: isolation, polymorphism and diversifying selection. Conserv Genet 10:1349.

Ammar R, Paton TA, Torti D, Shlien A, Bader GD (2015) Long read nanopore sequencing for detection of HLA and CYP2D6 variants and haplotypes. F1000Research 4:17.

Azevedo L, Serrano C, Amorim A, Cooper DN (2015) Trans-species polymorphism in humans and the great apes is generally maintained by balancing selection that modulates the host immune response. Hum Genom 9:21.

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

Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30:2114–2120.

Bollmer JL, Dunn PO, Freeman-Gallant CR, Whittingham LA (2012) Social and extra-pair mating in relation to major histocompatibility complex variation in common yellowthroats. Proc R Soc B Biol Sci 279:4778–4785.

Bollmer JL, Dunn PO, Whittingham LA, Wimpee C (2010) Extensive MHC class II B gene duplication in a passerine, the common Yellowthroat (Geothlypis trichas). J Hered 101:448–460.

Bontrop RE (2006) Comparative genetics of MHC polymorphisms in different primate species: duplications and deletions. Hum Immunol 67:388–397.

Brown JH, Jardetzky T, Saper MA, Samraoui B, Bjorkman PJ, Wiley DC (1988) A hypothetical model of the foreign antigen binding site of Class II histocompatibility molecules. Nature 332:845–850.

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

Burri R, Hirzel HN, Salamin N, Roulin A, Fumagalli L (2008) Evolutionary Patterns of MHC Class II B in Owls and Their Implications for the Understanding of Avian MHC Evolution. Mol Biol Evol 25:1180–1191.

Burri R, Promerová M, Goebel J, Fumagalli L (2014) PCR-based isolation of multigene families: lessons from the avian MHC class IIB. Mol Ecol Resour 14:778–788.

Burri R, Salamin N, Studer RA, Roulin A, Fumagalli L (2010) Adaptive Divergence of Ancient Gene Duplicates in the Avian MHC Class II. Mol Biol Evol 27:2360–2374.

Chan DTC, Poon ESK, Wong ATC, Sin SYW (2021) Global trade in parrots – Influential factors of trade and implications for conservation. Glob Ecol Conserv 30:e01784.

Chappell PE, Meziane EK, Harrison M, Magiera Ł, Hermann C, Mears L et al (2015) Expression levels of MHC class I molecules are inversely correlated with promiscuity of peptide binding. eLife 4:e05345.

Chen L-C, Lan H, Sun L, Deng Y-L, Tang K-Y, Wan Q-H (2015) Genomic organization of the crested ibis MHC provides new insight into ancestral avian MHC structure. Sci Rep 5:7963.

Danecek P, Bonfield JK, Liddle J, Marshall J, Ohan V, Pollard MO et al (2021) Twelve years of SAMtools and BCFtools. Gigascience 10:p.giab008.

Dilger WC (1960) The comparative ethology of the African parrot genus Agapornis. Z Tierpsychol 17:649–685.

Edwards SV, Hess CM, Gasper J, Garrigan D (1999) Toward an evolutionary genomics of the avian Mhc. Immunol Rev 167:119–132.

Forshaw JM (2010) Parrots of the world. Princeton University Press. 41 William Street, Princeton, New Jersey 70:08540.

Gaigher A, Burri R, Gharib WH, Taberlet P, Roulin A, Fumagalli L (2016) Family-assisted inference of the genetic architecture of major histocompatibility complex variation. Mol Ecol Resour 16:1353–1364.

Gaigher A, Rota A, Neves F, Muñoz-Mérida A, Blasco-Aróstegui J, Almeida T et al (2023) Extensive MHC class IIβ diversity across multiple loci in the small-spotted catshark (Scyliorhinus canicula). Sci Rep 13:3837.

Geliebter J, Nathenson SG (1987) Recombination and the concerted evolution of the murine MHC. Trends Genet 3:107–112.

Gillingham MA, Courtiol A, Teixeira M, Galan M, Bechet A, Cezilly F (2016) Evidence of gene orthology and trans-species polymorphism, but not of parallel evolution, despite high levels of concerted evolution in the major histocompatibility complex of flamingo species. J Evol Biol 29:438–454.

Goebel J, Promerová M, Bonadonna F, McCoy KD, Serbielle C, Strandh M et al (2017) 100 million years of multigene family evolution: origin and evolution of the avian MHC class IIB. BMC Genom 18:460.

Grayson P, Sin SYW, Sackton TB, Edwards SV (2017) Comparative genomics as a foundation for evo-devo studies in birds. Methods Mol Biol 1650:11–46.

He K, Liang CH, Zhu Y, Dunn P, Zhao A, Minias P (2022) Reconstructing macroevolutionary patterns in avian MHC architecture with genomic data. Front Genet 13:823686.

Hess CM, Edwards SV (2002) The Evolution of the Major Histocompatibility Complex in Birds: scaling up and taking a genomic approach to the major histo compatibilty complex (MHC) of birds reveals surprising departures from generalities found in mammals in both large-scale structure and the mechanisms shaping the evolution of the MHC. BioScience 52:423–431.

Hughes AL, Nei M (1989) Evolution of the major histocompatibility complex: independent origin of nonclassical class I genes in different groups of mammals. Mol Biol Evol 6:559–579.

Hughes AL, Nei M (1990) Evolutionary relationships of class II major-histocompatibility-complex genes in mammals. Mol Biol Evol 7:491–514.

Hughes CR, Miles S, Walbroehl JM (2008) Support for the minimal essential MHC hypothesis: a parrot with a single, highly polymorphic MHC class II B gene. Immunogenetics 60:219–231.

Huynh S, Cloutier A, Sin SYW (2023) Museomics and phylogenomics of lovebirds (Psittaciformes, Psittaculidae, Agapornis) using low-coverage whole-genome sequencing. Mol Phylogenetics Evol 185:107822.

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

Kloch A, Babik W, Bajer A, Siński E, Radwan J (2010) Effects of an MHC-DRB genotype and allele number on the load of gut parasites in the bank vole Myodes glareolus. Mol Ecol 19:255–265.

Knafler G, Ortiz-Catedral L, Jackson B, Varsani A, Grueber C, Robertson B et al (2016) Comparison of beak and feather disease virus prevalence and immunity-associated genetic diversity over time in an island population of red-crowned parakeets. Arch Virol 161:811–820.

Knafler GJ, Fidler A, Jamieson IG, Robertson BC (2014) Evidence for multiple MHC class II β loci in New Zealand's critically endangered kakapo, Strigops habroptilus. Immunogenetics 66:115–121.

Kosakovsky Pond SL, Frost SDW (2005) Not so different after all: a comparison of methods for detecting amino acid sites under selection. Mol Biol Evol 22:1208–1222.

Kosakovsky Pond SL, Posada D, Gravenor MB, Woelk CH, Frost SDW (2006) Automated phylogenetic detection of recombination using a genetic algorithm. Mol Biol Evol 23:1891–1901.

Kumar S, Stecher G, Tamura K (2016) MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. Mol Biol Evol 33:1870–1874.

Lam DK, Frantz AC, Burke T, Geffen E, Sin SYW (2023) Both selection and drift drive the spatial pattern of adaptive genetic variation in a wild mammal. Evolution 77:221–238.

Lam DK, Sin SYW (2020) Development and characterization of microsatellite markers in Rosy-faced and other lovebirds (Agapornis spp.) using next-generation sequencing. Mol Biol Rep 47:6417–6427.

Lenz TL, Deutsch AJ, Han B, Hu X, Okada Y, Eyre S et al (2015) Widespread non-additive and interaction effects within HLA loci modulate the risk of autoimmune diseases. Nat Genet 47:1085–1090.

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

Lighten J, van Oosterhout C, Paterson IG, McMullan M, Bentzen P (2014) Ultra-deep Illumina sequencing accurately identifies MHC class IIb alleles and provides evidence for copy number variation in the guppy (Poecilia reticulata). Mol Ecol Resour 14:753–767.

Manegold A, Podsiadlowski L (2014) On the systematic position of the Black-collared Lovebird Agapornis swindernianus (Agapornithinae, Psittaciformes). J Ornithol 155:581–589.

Messing J (1983) New M13 vectors for cloning. Methods Enzymol 101:20–78.

Migalska M, Sebastian A, Radwan J (2019) Major histocompatibility complex class I diversity limits the repertoire of T cell receptors. Proc Natl Acad Sci USA 116:5021–5026.

Minias P, Bateson ZW, Whittingham LA, Johnson JA, Oyler-McCance S, Dunn PO (2016) Contrasting evolutionary histories of MHC class I and class II loci in grouse—effects of selection and gene conversion. Heredity 116:466–476.

Minias P, Edwards SV, Babik W (2023) Ancient duplication, coevolution, and selection at the MHC class IIA and IIB genes of birds. Front Immunol 14:1250824.

Minias P, Gutiérrez JS, Dunn PO (2020) Avian major histocompatibility complex copy number variation is associated with helminth richness. Biol Lett 16:20200194.

Minias P, He K, Dunn PO (2021a) The strength of selection is consistent across both domains of the MHC class I peptide-binding groove in birds. BMC Ecol Evol 21:80.

Minias P, Pikus E, Anderwald D (2019a) Allelic diversity and selection at the MHC class I and class II in a bottlenecked bird of prey, the White-tailed Eagle. BMC Evol Biol 19:2.

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

Minias P, Pikus E, Whittingham LA, Dunn PO (2019b) Evolution of copy number at the MHC varies across the avian tree of life. Genome Biol Evol 11:17–28.

Minias P, Whittingham LA, Dunn PO (2017) Coloniality and migration are related to selection on MHC genes in birds. Evolution 71:432–441.
Minias P, Włodarczyk R, Remisiewicz M, Cobzaru I, Janiszewski T (2021b) Distinct evolutionary trajectories of MHC class I and class II genes in Old World finches and buntings. Heredity 126:974–990.
Morand S, Poulin R (1998) Density, body mass and parasite species richness of terrestrial mammals. Evol Ecol 12:717–727.
Mori, Cardador E, Reino L, White L, Hernández-Brito RL, Le D, Louarn M et al (2020) Lovebirds in the air: trade patterns, establishment success and niche shifts of Agapornis parrots within their non-native range. Biol Invasions 22:421–435.
Murrell B, Moola S, Mabona A, Weighill T, Sheward D, Kosakovsky Pond SL et al (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.
Ndithia H, Perrin MR, Waltert M (2007) Breeding biology and nest site characteristics of the Rosy-faced Lovebird Agapornis roseicollis in Namibia. Ostrich 78:13–20.
Nei M, Rooney AP (2005) Concerted and birth-and-death evolution of multigene families. Annu Rev Genet 39:121–152.
O'Connor E, Westerdahl H, Burri R, Edwards S (2019) Avian MHC evolution in the era of genomics: phase 1.0. Cells 8:1152.
O'Connor EA, Cornwallis CK, Hasselquist D, Nilsson J, Westerdahl H (2018) The evolution of immunity in relation to colonization and migration. Nat Ecol Evol 2:841–849.
O'Connor EA, Strandh M, Hasselquist D, Nilsson JÅ, Westerdahl H (2016) The evolution of highly variable immunity genes across a passerine bird radiation. Mol Ecol 25:977–989.
O'Connor EA, Westerdahl H (2021) Trade-offs in expressed major histocompatibility complex diversity seen on a macroevolutionary scale among songbirds. Evolution 75:1061–1069.
Osipova E, Ko M-C, Petricek KM, Yung Wa Sin S, Brown T, Winkler S et al (2024) Convergent and lineage-specific genomic changes contribute to adaptations in sugar-consuming birds. bioRxiv https://doi.org/10.1101/2024.08.30.610474.
Posada D (2008) jModelTest: phylogenetic model averaging. Mol Biol Evol 25:1253–1256.
Qin S, Dunn PO, Yang Y, Liu H, He K (2021) Polymorphism and varying selection within the MHC class I of four Anas species. Immunogenetics 73:395–404.
Radwan J, Babik W, Kaufman J, Lenz TL, Winternitz J (2020) Advances in the evolutionary understanding of MHC polymorphism. Trends Genet 36:298–311.
Ren L, Yang Z, Wang T, Sun Y, Guo Y, Zhang Z et al (2011) Characterization of the MHC class II α-chain gene in ducks. Immunogenetics 63:667–678.
Robinson J, Guethlein LA, Cereb N, Yang SY, Norman PJ, Marsh SGE et al (2017) Distinguishing functional polymorphism from random variation in the sequences of >10,000 HLA-A, -B and -C alleles. PLoS Genet 13:e1006862.
Ronquist F, Huelsenbeck JP (2003) MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19:1572–1574.
Saper MA, Bjorkman PJ, Wiley DC (1991) Refined structure of the human histocompatibility antigen HLA-A2 at 2.6 A resolution. J Mol Biol 219:277–319.
Schweizer M, Seehausen O, Hertwig ST (2011) Macroevolutionary patterns in the diversification of parrots: effects of climate change, geological events and key innovations. J Biogeogr 38:2176–2194.
Sedlazeck FJ, Rescheneder P, von, Haeseler A (2013) NextGenMap: fast and accurate read mapping in highly polymorphic genomes. Bioinformatics 29:2790–2791.
Seutin G, White BN, Boag P (1991) Preservation of avian blood and tissue samples for DNA analyses. Can J Zool 69:82–90.
Sin SYW, Cloutier A, Nevitt G, Edwards SV (2022) Olfactory receptor subgenome and expression in a highly olfactory procellariiform seabird. Genetics 220:iyab210.
Sin YW, Annavi G, Dugdale HL, Newman C, Burke T, Macdonald DW (2014) Pathogen burden, co-infection and major histocompatibility complex variability in the European badger (*Meles meles*). Mol Ecol 23:5072–5088.
Sin YW, Dugdale HL, Newman C, Macdonald DW, Burke T (2012a) Evolution of MHC class I genes in the European badger (*Meles meles*). Ecol Evol 2:1644–1662.
Sin YW, Dugdale HL, Newman C, Macdonald DW, Burke T (2012b) MHC class II genes in the European badger (*Meles meles*): characterization, patterns of variation, and transcription analysis. Immunogenetics 64:313–327. 310.1007/s00251-00011-00578-00259.
Sommer S, Courtiol A, Mazzoni CJ (2013) MHC genotyping of non-model organisms using next-generation sequencing: a new methodology to deal with artefacts and allelic dropout. BMC Genom 14:542.
Spurgin LG, Richardson DS (2010) How pathogens drive genetic diversity: MHC, mechanisms and misunderstandings. Proc R Soc B Biol Sci 277:979–988.
Stervander M, Dierickx EG, Thorley J, Brooke MDL, Westerdahl H (2020) High MHC gene copy number maintains diversity despite homozygosity in a Critically Endangered single-island endemic bird, but no evidence of MHC-based mate choice. Mol Ecol 29:3578–3592.
Strandh M, Lannefors M, Bonadonna F, Westerdahl H (2011) Characterization of MHC class I and II genes in a subantarctic seabird, the blue petrel, Halobaena caerulea (Procellariiformes). Immunogenetics 63:653–666.
Takahashi K, Rooney AP, Nei M (2000) Origins and divergence times of mammalian class II MHC gene clusters. J Hered 91:198–204.
Thorvaldsdóttir H, Robinson JT, Mesirov JP (2013) Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. Brief Bioinform 14:178–192.
Viļuma A, Mikko S, Hahn D, Skow L, Andersson G, Bergström TF (2017) Genomic structure of the horse major histocompatibility complex class II region resolved using PacBio long-read sequencing technology. Sci Rep 7:45518.
Weaver S, Dube S, Mir A, Qin J, Sun G, Ramakrishnan R et al (2010) Taking qPCR to a higher level: analysis of CNV reveals the power of high throughput qPCR to enhance quantitative resolution. Methods 50:271–276.
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.
Westbrook CJ, Karl JA, Wiseman RW, Mate S, Koroleva G, Garcia K, Sanchez-Lockhart M, O'Connor DH, Palacios G (2015) No assembly required: Full-length MHC class I allele discovery by PacBio circular consensus sequencing. Hum Immunol 76:891–896.
Wegner KM, Kalbe M, Kurtz J, Reusch TB, Milinski M (2003) Parasite selection for immunogenetic optimality. Science 301:1343.
Wittzell H, Bernot A, Auffray C, Zoorob R (1999) Concerted evolution of two Mhc class II B loci in pheasants and domestic chickens. Mol Biol Evol 16:479–490.
Wong ATC, Lam DK, Poon ESK, Chan DTC, Sin SYW (2022) Intra-specific copy number variation of MHC class II genes in the Siamese fighting fish. Immunogenetics 74:327–346.

ACKNOWLEDGEMENTS

The authors would like to thank the University of Kansas Biodiversity Institute (KUBI) Ornithology Collection, University of Washington Burke Museum (UWBM) Ornithology Collection, Field Museum of Natural History (Zoology) Bird Collection, Louisiana State University Museum of Natural Science (LSUMZ), American Museum of Natural History (AMNH) Bird Collection, and San Diego Natural History Museum (SDNHM) Birds Collection, and bird owners for providing samples. The authors also thank Emily Shui Kei Poon and Bowen Wan for assistance on laboratory work, Charis May Ngor Chan for technical support, and Alison Cloutier for her comments and edits on the manuscript. This research was supported by the Seed Fund for Basic Research for New Staff (HKU). The computations were performed using the research computing facilities offered by the Information Technology Services at the University of Hong Kong.

AUTHOR CONTRIBUTIONS

Conceptualization: DKL and SYWS; methodology: DKL and SYWS; investigation: DKL; writing—original draft: DKL; writing—review and editing: DKL and SYWS; resources: SYWS; supervision: SYWS; funding acquisition: SYWS.

COMPETING INTERESTS

The authors declare no competing interests.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41437-025-00815-4.

Correspondence and requests for materials should be addressed to Simon Yung Wa Sin.

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.