

ORIGINAL RESEARCH

“Balancing” balancing selection? Assortative mating at the major histocompatibility complex despite molecular signatures of balancing selection

Joel W. G. Slade  | Matthew J. Watson | Elizabeth A. MacDougall-Shackleton 

Department of Biology, University of Western Ontario, London, Ontario, Canada

Correspondence

Joel W.G. Slade, BEACON Center for the Study of Evolution in Action, Michigan State University, East Lansing, MI.
Email: sladejo1@msu.edu

Funding information

Natural Sciences and Engineering Research Council of Canada, Grant/Award Number: 293123-2012RGPIN; Society of Canadian Ornithologists; American Ornithologists' Union

Abstract

In vertebrate animals, genes of the major histocompatibility complex (MHC) determine the set of pathogens to which an individual's adaptive immune system can respond. MHC genes are extraordinarily polymorphic, often showing elevated nonsynonymous relative to synonymous sequence variation and sharing presumably ancient polymorphisms between lineages. These patterns likely reflect pathogen-mediated balancing selection, for example, rare-allele or heterozygote advantage. Such selection is often reinforced by disassortative mating at MHC. We characterized exon 2 of MHC class II, corresponding to the hypervariable peptide-binding region, in song sparrows (*Melospiza melodia*). We compared nonsynonymous to synonymous sequence variation in order to identify positively selected sites; assessed evidence for trans-species polymorphisms indicating ancient balancing selection; and compared MHC similarity of socially mated pairs to expectations under random mating. Six codons showed elevated ratios of nonsynonymous to synonymous variation, consistent with balancing selection, and we characterized several alleles similar to those occurring in at least four other avian families. Despite this evidence for historical balancing selection, mated pairs were significantly more similar at MHC than were randomly generated pairings. Nonrandom mating at MHC thus appears to partially counteract, not reinforce, pathogen-mediated balancing selection in this system. We suggest that in systems where individual fitness does not increase monotonically with MHC diversity, assortative mating may help to avoid excessive offspring heterozygosity that could otherwise arise from long-standing balancing selection.

KEYWORDS

assortative mating, balancing selection, major histocompatibility complex, positive selection, song sparrows, trans-species polymorphism

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2019 The Authors. *Ecology and Evolution* published by John Wiley & Sons Ltd.

1 | INTRODUCTION

In jawed vertebrate animals, the major histocompatibility complex (MHC) is a key component of adaptive immune defense. MHC genes encode cell-surface glycoproteins that recognize and bind antigenic peptides, present them to T cells, and, in the case of nonself antigens, initiate an adaptive immune response (Trowsdale, 1995). MHC genotype determines the suite of pathogens to which an individual can respond, making these loci critical to disease resistance and subject to intense pathogen-mediated selection. Although both directional (e.g., good genes, locally good genes) and balancing selection (e.g., rare-allele advantage, heterozygote advantage) are well established to occur at MHC, the latter mode of selection is particularly well-studied (Bernatchez & Landry, 2003). Supporting the widespread importance of balancing selection at MHC, these loci are the most polymorphic in the vertebrate genome: In some systems, upwards of a thousand alleles are maintained (Robinson, 2003).

Balancing selection at MHC may take the form of negative frequency-dependent selection, arising from arms races with pathogens (Slade & McCallum, 1992), and/or heterozygote advantage (Niskanen et al., 2014). Contemporary balancing selection can be demonstrated directly, for example, if disease resistance is higher in individuals bearing rare alleles (Phillips et al., 2018) or highly heterozygous genotypes (Doherty & Zinkernagel, 1975; McClelland, Penn, & Potts, 2003; Penn, Damjanovich, & Potts, 2002). In addition to such snapshots of current selection regimes, historical balancing selection has also been inferred over longer timescales. Signatures of past balancing selection include the occurrence of trans-species polymorphisms, thought to be maintained by long-standing balancing selection that originated before the lineages became reproductively isolated (Hedrick, 1998; Hess & Edwards, 2002; Klein, 1980). Past balancing selection can also be reflected in positive selection, inferred from an excess of nonsynonymous relative to synonymous sequence variation within a population. Although positive selection is rare across most of the genome (Yang & Swanson, 2002), this pattern is observed relatively frequently at MHC (Hughes & Hughes, 1995; Piertney & Oliver, 2006), suggesting that parasite-mediated balancing selection has often favored new and rare variants at these loci.

The survival advantage associated with particular combinations of alleles suggests that provided individuals can assess the genetic similarity of potential mates, sexual selection in the form of disassortative mating may further reinforce the balancing effects of pathogen-mediated natural selection. Pathogen-mediated balancing selection and disassortative mating are both expected to increase individual diversity (heterozygosity): the former by maintaining multiple alleles within the population, and the latter by increasing the likelihood of dissimilar parental genotypes coming together to produce offspring. To the extent that disease resistance varies positively with MHC diversity (e.g., Brambilla, Keller, Bassano, & Grossen, 2018; Hughes & Nei, 1989), disassortative mating at MHC should enhance offspring quality. Indeed, preferences for MHC-dissimilar mates have been reported for many animal species (Kamiya, O'Dwyer, Westerdahl, Senior, & Nakagawa, 2014), including mammals (e.g., mice *Mus musculus*

domesticus; Yamazaki et al., 1979), fish (e.g., brown trout *Salmo trutta*; Forsberg, Dannewitz, Petersson, & Grahn, 2007), reptiles (e.g., tuatara *Sphenodon punctatus*; Miller, Moore, Nelson, & Daugherty, 2009), seabirds (e.g., blue petrels *Halobaena caerulea*; Leclaire, Strandh, Mardon, Westerdahl, & Bonadonna, 2017), and passerine birds (e.g., savannah sparrows *Passerculus sandwichensis*; Freeman-Gallant, Johnson, Saponara, & Stanger, 2002). Nonrandom mating at MHC is thought to be facilitated by olfactory cues, because MHC gene products affect not only pathogen resistance but also individual odor (Brennan & Zufall, 2006; Milinski et al., 2005).

Although disassortative mating at MHC is taxonomically widespread among vertebrate animals, it is not universal. Some species mate assortatively, rather than disassortatively, at MHC (mammals: European badgers *Meles meles* (Sin et al., 2015); wolves *Canis lupus* (Galaverni et al., 2016); amphibians: tiger salamanders *Ambystoma tigrina* (Bos, Williams, Gopurenko, Bulut, & Dewoody, 2009); passerine birds: house sparrows *Passer domesticus* (Bonneaud, Chastel, Federici, Westerdahl, & Sorci, 2006)). Assortative mating at MHC may provide material benefits to the choosy individual, such as reduced risk of infection with novel pathogens carried by MHC-dissimilar immigrants (Lewis, 1998). Such assortative mating may also confer genetic benefits to offspring, for example, by reducing outbreeding depression in hybrid zones, reducing the disruption of coadapted sets of alleles (Roberts, 2009), and reducing susceptibility to autoimmune disorders associated with excessive individual diversity at MHC (Wegner, Kalbe, Kurtz, Reusch, & Milinski, 2003). Genetic benefits to assortative mating may be particularly salient in systems in which MHC is highly polygenic due to extensive duplication (Bollmer, Dunn, Whittingham, & Wimpee, 2010; Hess & Edwards, 2002; Minias, Pikus, Whittingham, & Dunn, 2018). In such systems, assortative mating may allow individuals to avoid the disadvantages of producing offspring with an excessive number of alleles at MHC. If so, rather than uniformly reinforcing MHC diversity, sexual selection may under some circumstances counteract—or “balance”—the effects of historical balancing selection.

Finally, some species appear to mate randomly with respect to MHC profiles. This pattern is particularly well-documented in passerine birds (e.g., great reed warblers *Acrocephalus arundinaceus* (Westerdahl, 2004); common yellowthroats *Geothlypis trichas* (Bollmer, Dunn, Freeman-Gallant, & Whittingham, 2012); great tits *Parus major* (Sepil et al., 2015), although not restricted to this group (e.g., brown bears *Ursus arctos*; Kuduk et al., 2014). Random mating at MHC, particularly among free-living animals, could suggest that any genetic benefits associated with MHC-mediated pairing may be outweighed by material benefits or other critical traits. Alternatively, this pattern might reflect sensory constraints and an inability to assess potential mates' MHC profiles. Thus, naturally occurring patterns of mate choice are particularly illuminating for species with documented phenotypic cues of MHC genotype.

We examined MHC-mediated mate choice in free-living song sparrows (*Melospiza melodia*; Figure 1). In the focal population, chemical composition of preen oil (the major source of body odor in birds) conveys information regarding MHC profiles (Slade et al.,

2016). To our knowledge, song sparrows are the only passerine bird for which this relationship has been demonstrated. We then related patterns of social mate choice to evidence for historical pathogen-mediated balancing selection, which we inferred from signatures of positive molecular evolution and trans-species polymorphisms. MHC molecules are generally categorized into two major classes: Classes I and II interact primarily with intracellular (e.g., viruses) and extracellular antigens (e.g., bacteria), respectively (Klein, 1986). We focused on the hypervariable second exon of class II for two reasons. First, chemical composition of preen oil covaries with MHC class II genotype (Slade et al., 2016), raising the possibility that olfactory cues of class II similarity could facilitate nonrandom mating. Second, this region shows the highest individual- and population-level variability in song sparrows (Slade, Sarquis-Adamson, Gloor, Lachance, & MacDougall-Shackleton, 2017) and other passerines (Minias et al., 2018). Song sparrows show as many as 26 alleles per individual at MHC class II, implying at least 13 loci (Slade, Sarquis-Adamson, et al., 2017), and common yellowthroats as many as 39 alleles per individual, implying at least 20 loci (Bollmer et al., 2010).

To assess MHC-mediated mating, we identified breeding pairs in the wild, calculated genetic distance between mates, and compared observed distances to those expected under random mating. To assess historical evidence of balancing selection, presumably mediated by pathogen resistance, we examined patterns of molecular evolution (i.e., ratios of nonsynonymous to synonymous variation), and surveyed for trans-species polymorphisms by comparing song sparrow sequences at MHC class II to those from other passerine species. Ultimately, we sought to determine whether nonrandom mating at MHC affects individual diversity in a manner congruent with, or opposing, the inferred effects of past selection.

2 | MATERIALS AND METHODS

2.1 | Study animals and field methods

Fieldwork focused on a long-term study population of migratory song sparrows breeding near Newboro, Ontario, Canada (44.6338°N,



FIGURE 1 Song sparrow (*Melospiza melodia*). Photo credit: Tasha Kelly

76.3308°W). During spring 2014 (April 14–June 2) and 2015 (April 13–June 6), corresponding to nesting, egg-laying, incubating, and provisioning offspring in the study population, we captured adult song sparrows; collected blood samples for genetic analysis; and identified socially mated pairs, using trapping records and behavioral observations as detailed below.

We captured song sparrows in two-celled, seed-baited Potter traps, which we checked once per hour, three times a day, between 07:00 and 11:00. We captured 69 birds in 2014 (44 males, 25 females) and 87 in 2015 (49 males, 38 females). These figures include 28 birds that were captured in both study years (25 males, three females); thus, our sample comprises 128 individuals (68 males, 60 females).

From each bird, we collected ~25 μ l of whole blood via brachial venipuncture the first time it was captured each year. We blotted blood onto high wet-strength filter paper saturated with 0.5 M Na-EDTA (pH 8.0), allowed the blot to air-dry, then stored it at room temperature awaiting DNA extraction. We identified sex based on the presence (male) or absence (female) of a cloacal protuberance, supplemented by unflattened wing length (measured with dial calipers to the nearest 0.1 mm). If not already banded, we outfitted the bird with a numbered aluminum leg band (Environment and Climate Change Canada Scientific Banding Permit 10691B), and a unique combination of three colored plastic leg bands to permit individual identification in the field. We recorded any other song sparrows in the trap (i.e., trapped in the other cell at the same time) and released birds at their site of capture. In almost all cases, birds were later resighted and/or recaptured, implying that they were resident breeders. Animal procedures were approved by the Animal Use Subcommittee at the University of Western Ontario (protocols 2008-054 and 2015-047 to EAM-S) and conducted under the required federal permits.

Song sparrows are socially monogamous, and extrapair mating is rare in the study population: Fewer than 20% of nests contain extrapair offspring (Potvin & MacDougall-Shackleton, 2009). For these reasons, and because high rates of nest failure make it challenging to collect blood from nestlings for genetic parentage analysis, we focused on social rather than genetic mate choice. We identified socially mated pairs through opportunistic behavioral observations on color-banded individuals (e.g., copulations or copulation solicitations; female and male observed foraging together; female and male observed provisioning the same nest) combined with trapping records. We identified pairs based on trapping records if one or both of the following conditions were met: (a) Two individuals of opposite sex were trapped together at the same time (8.7% of all trapping records), or (b) two individuals of opposite sex were captured in the same trap within 48 hr of one another, and no other song sparrows were trapped at that location during the entire field season (8.8% of all trapping records). Based on these criteria, we identified 18 socially mated pairs in 2014 and 22 in 2015. These figures include one pair that remained the same over both field seasons; to avoid pseudoreplication, we included this pairing in the 2014 mate choice analysis only. These figures also include three individuals (one male,

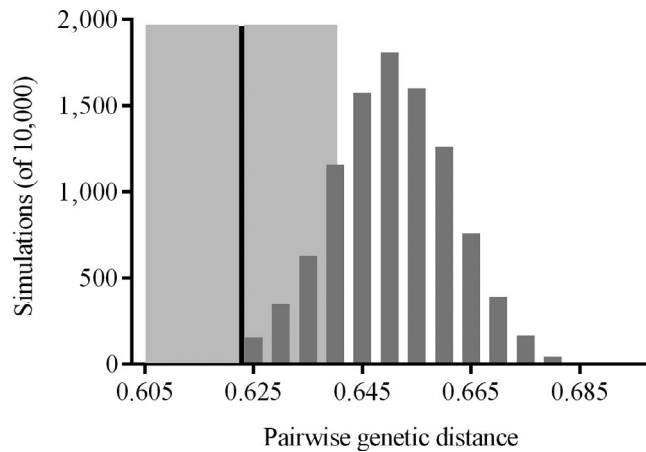


FIGURE 2 Frequency distribution of pairwise genetic distances at exon 2 of MHC class II, generated by Monte Carlo simulations of 10,000 randomized male–female pairings based on MHC genotypes characterized in the population. The vertical line at 0.623 denotes the average pairwise genetic distance between observed mates, inferred from behavioral and trapping records. Shading corresponds to ± 1 SE around this average. Mated pairs were more similar at MHC than expected under random mating (two-tailed $p = 0.018$)

two females) that were each identified as part of a mated pair in 2014, but had a different mate in 2015: These were included in both years' mate choice analyses.

2.2 | Characterizing MHC

For each of the 128 birds sampled, we used polymerase chain reaction (PCR) to amplify the hypervariable second exon of class II MHC. We used primers *SospMHCint1f* (Slade et al., 2016) and *Int2r.1* (Edwards, Gasper, & March, 1998), which should bind within introns 1 and 2 respectively, to amplify exon 2. In addition to the priming sequence, each primer included a unique “barcode” sequence of eight base pairs; four “wobble” bases; and an adaptor sequence for the Illumina MiSeq platform. PCRs were conducted in a total volume of 30 μ l, including 12.5 μ l of GoTaq® Hot Start Master Mix (Promega), 0.2 μ M of each primer, and 25–60 ng of template DNA. Cycling conditions were 3 min at 94°C; 28 cycles of 30 s at 94°C, 30 s at 62°C, and 45 s at 72°C; and a final extension step of 10 min at 72°C. We confirmed amplification by running a portion of each PCR product on a 2% agarose gel. We then pooled together all samples for a given field season (2014, 2015) to form a library. Each season's library was run on a separate Illumina MiSeq flow cell at the London Regional Genomics Center.

We used a pipeline (Gloor et al., 2010) together with individually unique combinations of barcode sequences to assign MHC reads to individuals. We identified and removed chimeric sequences using UCHIME (Edgar, Haas, Clemente, Quince, & Knight, 2011). We also filtered out any sequence occurring in fewer than 1% of an individual's total reads, as these rare sequences might represent errors in PCR or sequencing. This 1% threshold was established as described in Slade, Sarquis-Adamson, et al. (2017); briefly, we used bacterial

cloning (Promega pGEM-T Easy Vector System) to generate colonies that each contained a single allele. We included these colonies on the flow cell run and used the frequency of secondary reads to establish the error rate associated with PCR and sequencing. After filtering, we aligned the remaining sequences in MEGA 7.0 (Kumar, Stecher, & Tamura, 2016). We trimmed each to a length of 216–219 base pairs (72–73 codons; this variation in length reflects the presence of a 3-bp indel), corresponding to codons 8–80 of exon 2 of MHC class II (total length 90 codons; Minias et al., 2018) in order to restrict our analysis to the region of highest read quality and to maintain consistency with previous studies (Slade, Sarquis-Adamson, et al., 2017; Slade et al., 2016; Slade, Watson, & MacDougall-Shackleton, 2017). Hereafter, references to specific codon numbers (e.g., codon 2, codon 6) refer to positions within the trimmed region that we analyzed, not within the full exon 2. We confirmed similarity to other passerine sequences by querying the Basic Local Alignment Search tool (BLAST; Altschul, Gish, Miller, Myers, & Lipman, 1990) implemented in GenBank.

2.3 | Mate choice analysis

To test whether song sparrows pair nonrandomly at MHC class II, we compared genetic distances between socially mated pairs to the randomized set of possible opposite-sex combinations based on adults captured at the site during the corresponding field season. Because some allelic pairs are more similar in sequence than others, rather than simply quantify allele sharing, we used UniFrac (Lozupone & Knight, 2005) to take into account phylogenetic distances between alleles. Based on a phylogeny of all alleles detected, the algorithm calculates the distance between two individuals' translated MHC repertoires, such that two individuals with the identical set of alleles would have a UniFrac distance of zero and two individuals with alleles derived from completely different clades in the reference tree would have a UniFrac distance of one (Lozupone & Knight, 2005). We first constructed a maximum-likelihood phylogeny of all alleles recovered, over both field seasons, using Whelan and Goldman (2001) substitution with five discrete gamma categories. Based on this phylogeny, we calculated pairwise amino acid UniFrac distances for observed pairings (i.e., between socially mated pairs). Because genotype data were binary (i.e., presence or absence of an allele within an individual's genotype) rather than continuous, we calculated unweighted (qualitative) rather than weighted (quantitative) UniFrac. UniFrac distances were calculated using the package GUniFrac (Chen et al., 2012), implemented in R 3.4.0 (R Core Team, 2017).

We used a Monte Carlo simulation (Manly, 1997) implemented in a custom Microsoft Excel macro (Neff, Garner, Heath, & Heath, 2008) to generate 10,000 randomized male–female adult pairings. Genotypes from which randomized pairings were drawn were sorted by year, such that, for example, an adult female captured at the field site in 2014 was assumed to have had the potential to pair with any of the adult males present in that year. We then calculated pairwise genetic distance for each randomized pairing using unweighted

UniFrac as described above. Finally, we calculated the average genetic distance over the 40 observed social pairings and compared this to the expected distribution of genetic distances based on random mating.

2.4 | Sequence evolution

We tested for signatures of positive selection using all available genotypes from the study population in 2014 and 2015 (128 breeding adults). After filtering out chimeras and very low-frequency reads as described above, we tested for positive selection across our sequenced portion of exon 2, using PAMLx 1.3.1 (Xu & Yang, 2013). Statistical significance was assessed by a Z-test, bootstrapped at 500 replicates, and conducted in MEGA 7.0 (Kumar et al., 2016). To ensure we confidently identified positively selected sites, we tested for positive selection at each codon site separately using multiple complementary approaches implemented in datamonkey.org (Weaver et al., 2018) in addition to PAMLx 1.3.1. Specifically, we tested for positive selection using a mixed-effects model of evolution (MEME; Murrell et al., 2012), fixed-effects likelihood (FEL; Kosakovsky Pond & Frost, 2005), single-likelihood ancestor counting (SLAC; Kosakovsky Pond & Frost, 2005), and fast, unconstrained Bayesian approximation (FUBAR; Murrell et al., 2013). Finally, because positive selection is expected to be strongest at peptide-binding codons, we compared the set of codons our analyses identified as having experienced past positive selection to the set of codons identified as peptide-binding in humans (Brown et al. 1993), as well as to those codons recently identified as positively selected in passerines in general (Minias et al., 2018).

In PAMLx 1.3.1, we used the *codeml* command (Yang, 2007) to calculate ω , defined as the ratio of nonsynonymous substitutions per nonsynonymous site (dN) to synonymous substitutions per synonymous site (dS). Codons (sites) at which ω is less than, equal to, or greater than one is interpreted as having undergone purifying (negative) selection (ω_0), neutral evolution (ω_1), and positive selection (ω_2), respectively. *Codeml* does not make a priori assumptions as to which codons are likely to experience each type of selection (Yang & Swanson, 2002).

To evaluate which model(s) of sequence evolution best explained the observed variation in MHC class II exon 2 sequence in *codeml*, we used an information theoretic approach (Burnham & Anderson, 2002), ranking candidate models generated by PAML based on Akaike's information criterion (AIC). Candidate models were M1a (nearly neutral; $\omega_0 < 1$, $\omega_1 = 1$), M2a (positive selection; $\omega_2 > 1$), M8 (positive selection with β distribution; $0 < \omega < 1$, $\omega > 1$), and M7 (null model counterpart to M8). Models M1a and M2a use the ω distribution to model parameters, and models M7 and M8 use the β distribution, constrained to range between 0 and 1 (Yang, 2000). We were particularly interested in the explanatory power of M2a relative to M1a, and of M8 relative to M7, because models M2a and M8 permit some codons to experience positive selection, whereas the null models M1a and M7 assume that codons experience neutral evolution, but also test for purifying selection. We identified positively

selected codons using the Bayes empirical Bayes approach (Yang, Wong, & Nielsen, 2005) implemented in PAMLx 1.3.1 (Xu & Yang, 2013).

2.5 | Trans-species polymorphisms

Trans-species polymorphisms involve two or more alleles, each occurring in two or more species. We conducted a nonexhaustive survey for such polymorphisms, including variants that while not identical across species, are more similar to their heterospecific counterpart than to one or more conspecific alleles. We identified the ten most common alleles in the study population of song sparrows and queried them against BLAST in GenBank (Altschul et al., 1990). For each allele, we identified any heterospecific sequences with at least 94% DNA sequence similarity. In all, we retrieved 25 such sequences from other songbirds.

We constructed a maximum-likelihood phylogeny of the ten most common song sparrow alleles and the 25 heterospecific sequences, using a Jukes–Cantor model. Interdigitation of song sparrow alleles with those of other species would suggest that allelic diversification occurred prior speciation events and that balancing selection is of ancient origin. Conversely, if the ten song sparrow alleles comprised a monophyletic clade, this would suggest that allelic diversification has occurred more recently than speciation events and that balancing selection is of relatively recent origin.

3 | RESULTS

We identified 278 unique DNA sequences (i.e., differing at one or more base pairs) from the 128 individuals genotyped, with an average (\pm SEM) of 14.1 ± 0.2 alleles per individual. Number of alleles per individual did not differ between the two years of the study (unpaired *t* test, $t_{154} = 1.42$, $p = 0.16$). Sequences have been deposited to GenBank (accession numbers KX263957-KX264148; KX375230-KX375341; MF197785-MF197843; MH670952-MH671105).

Pairwise distance at MHC was lower, on average, for observed pairings (i.e., socially mated pairs; mean \pm SEM = 0.623 ± 0.018) than for simulated pairings (0.650 ± 0.0001 ; two-tailed $p = 0.018$; Figure 2). Thus, socially mated pairs were more similar at MHC than expected under random mating.

The observed pattern of molecular evolution, averaged across the sequenced portion of MHC class II exon 2, was consistent with positive selection (Z-test, $Z = 2.17$, $p = 0.016$). Supporting this, of the four candidate models of codon evolution generated in PAML, the positive selection model (M2a) was the best supported, followed by the positive selection model with beta distribution (M8; Table 1); that is, both models allowing positive selection (i.e., M2a and M8) received comparable levels of support (AIC = 2) and were far better supported than models disallowing positive selection (i.e., M1a and M7; AIC ≥ 464 , Table 1). The positive selection model (M2a) indicated that 94.3% of sites (68–69 sites) have experienced purifying selection, 4.8% (3–5 sites) have evolved through neutral drift, and

TABLE 1 AIC-ranked codon maximum-likelihood models of sequence evolution, based on 518 DNA sequences of MHC class II, exon 2 recovered from song sparrows

Model	lnL	AIC	Δ AIC	Parameter estimates	Positively selected codons
Positive selection ω (M2a)	-11,072	22,126	--	$p_0 = 0.945$, $p_1 = 0.046$, $p_2 = 0.009$ $\omega_0 = 0.086$, $\omega_1 = 1$, $\omega_2 = 3.10$	1, <u>2</u> , 24, 26, 42*, 46*, 56, <u>60*</u> , <u>62*</u> , <u>73*</u>
Positive selection β (M8)	-11,105	22,220	94	$p_0 = 0.959$, $p_1 = 0.041$, $p = 0.041$, $q = 4.61$, $\omega = 2.25$	1, 2, <u>19*</u> , 24, 26, <u>27*</u> , 42*, 46*, 56, <u>60*</u> , 62, <u>63*</u> , <u>67</u> , <u>73*</u>
Nearly neutral (M1a)	-11,290	22,588	462	$p_0 = 0.977$, $p_1 = 0.023$, $\omega_0 = 0.049$, $\omega_1 = 1$	Not allowed
β (M7)	-11,359	22,722	596	$p = 0.021$, $q = 0.164$	Not allowed

Note. The estimated proportion of sites subject to purifying selection, neutral evolution (drift), and positive selection are denoted by p_0 , p_1 , and p_2 respectively. Underlining denotes codons that correspond to antigen-binding sites in humans (Brown et al. 1993); asterisks denote codons that are positively selected in passerines in general (Minias et al., 2018).

TABLE 2 Positively selected sites indicated by mixed-effects model of evolution (MEME), fixed-effects likelihood (FEL), single-likelihood ancestor counting (SLAC), and a fast, unconstrained Bayesian approximation for inferring selection (FUBAR)

Test	Positively selected sites
MEME	<u>2</u> , <u>7*</u> , <u>17</u> , <u>19*</u> , 24, 25, <u>27*</u> , 41, 42*, 46*, <u>49*</u> , <u>50</u> , 56, <u>60*</u> , <u>62*</u> , 66, 67, <u>71</u> , <u>73*</u>
FEL	<u>2</u> , <u>7*</u> , <u>19*</u> , 24, 42*, 46*, 56, <u>60*</u> , <u>62*</u> , 67, <u>73*</u>
SLAC	<u>2</u> , <u>7*</u> , <u>17</u> , <u>19*</u> , 24, 25, <u>27*</u> , 41, 42*, 46*, <u>49*</u> , <u>59*</u> , 62*, 67, <u>73*</u>
FUBAR	<u>2</u> , <u>7*</u> , <u>19*</u> , 24, <u>27*</u> , 42*, 46*, 62*, 67, <u>73*</u>

Note. Underlining denotes codons that correspond to antigen-binding sites in humans (Brown et al. 1993); asterisks denote codons that are positively selected in passerines in general (Minias et al., 2018).

1% (~1 site) have experienced positive selection (Table 1). Similarly, the positive selection model with β distribution (M8) indicated that 99.1% (72–73 sites) of sites have experienced purifying selection, and 0.9% (~1 site) have experienced positive selection. Under M2a, the Bayes empirical Bayes approach identified ten codons with signatures of positive selection. Five of these have also been identified as positively selected in passerines by Minias et al. (2018), and three correspond to antigen-binding sites in humans (Brown et al. 1993; Table 1). Under M8, the Bayes empirical Bayes approach identified fourteen codons with signatures of positive selection, seven of which matched sites in passerines (Minias et al., 2018), and six of which correspond to the antigen-binding sites based on alignment with human sequence (Brown et al. 1993; Table 1). Generally consistent with the above findings, each of the alternative tests for positive selection implemented in Datamonkey 2.0 (Weaver et al., 2018) identified several codons showing the signature of positive selection (19, 11, 15, and 10 codons identified by MEME, FEL, SLAC, and FUBAR, respectively; Table 2). Across all tests for positive selection (i.e., PAML models M2a and M8; MEME, FEL, SLAC, and FUBAR tests in Datamonkey 2.0), six codons (2, 24, 42, 46, 62, and 73) were consistently identified by all methods as having undergone positive selection. Of these, codons 2 and 73 correspond to peptide-binding sites in human MHC class II (Brown et al. 1993), and codons 42, 46, 62, and 73 have also been identified as positively selected across passerines in general (Minias et al., 2018).

The ten most common MHC class II alleles recovered from song sparrows showed 94%–98% sequence similarity to 25 sequences from eight other songbird species, belonging to five different

families (Passerellidae: savannah sparrow *Passerculus sandwichensis*; Thraupidae: large cactus finch *Geospiza conirostris*, medium ground finch *G. fortis*, woodpecker finch *Cactospiza pallida*; Parulidae: common yellowthroat *Geothlypis trichis*; Icteridae: red-winged blackbird *Agelaius phoeniceus*; Emberizinae: meadow bunting *Emberiza cioides*, Jankowski's bunting *E. jankowskii*; Figure 3). None of the 35 alleles investigated had 100% sequence identity to others published to GenBank; thus, we found no allelic pairs in song sparrows that were shared at 100% sequence similarity by another species. However, song sparrow alleles did not cluster as a single monophyletic clade (Figure 3). Instead, we observed several well-supported clades in which one or more song sparrow alleles were more similar to heterospecific alleles within the clade than to conspecific alleles outside the clade (Figure 3). For example, song sparrow alleles SOSP-DAB*18, *19, and *21 were more similar to alleles from Jankowski's bunting, meadow bunting, red-winged blackbird, and the three species of Galapagos finch than to any other song sparrow alleles. Similarly, song sparrow allele SOSP-DAB*4 was more similar to the common yellowthroat allele Getr-DAB*809 than to any other song sparrow allele. Reciprocally, Getr-DAB*809 was more similar to SOSP-DAB*4 than to a different common yellowthroat allele in the phylogeny (Figure 3).

4 | DISCUSSION

The diversifying effects of disassortative mating, that is, increasing genetic diversity of individuals and maintaining multiple alleles within

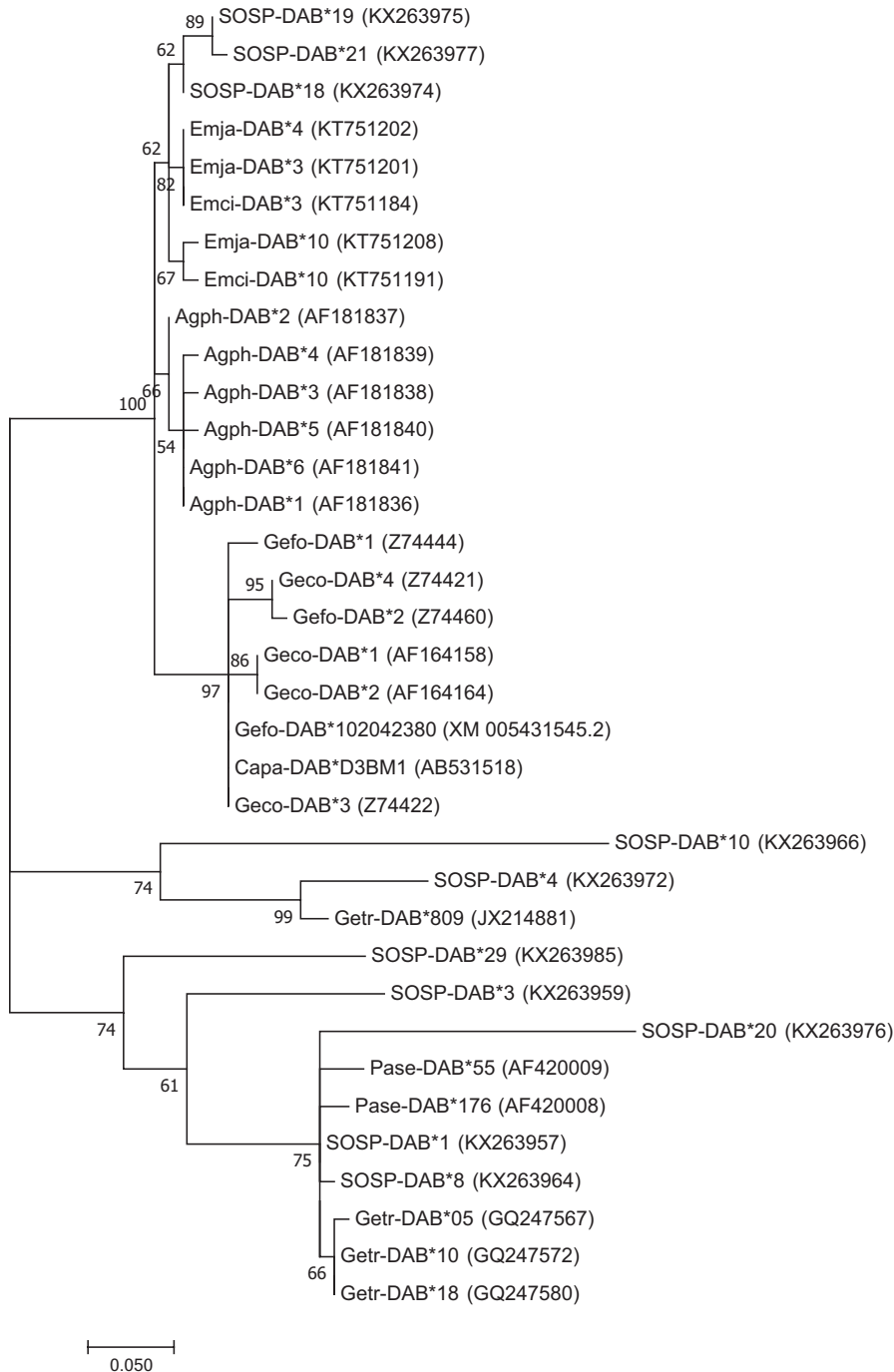


FIGURE 3 Unrooted phylogenetic tree of the ten most common alleles at MHC class II, exon 2 in song sparrows in this study (SOSP-DAB*), plus 25 other sequences retrieved from GenBank with 94%–98% identity published for other songbird species. These additional sequences are denoted as Emja (Jankowski's bunting, *Emberiza jankowskii*), Emci (meadow bunting, *Emberiza cioides*), Agph (red-winged blackbird, *Agelaius phoeniceus*), Gefo (medium ground finch, *Geospiza fortis*), Geco (Española cactus finch, *Geospiza conirostris*), Capa (woodpecker finch, *Cactospiza pallida*), Getr (common yellowthroat, *Geothlypis trichas*), and Pase (Savannah sparrow, *Passerculus sandwichensis*). All sequence names are followed by their GenBank accession number. Phylogeny was inferred by maximum likelihood based on the Jukes–Cantor model (log likelihood = -905.29). Bootstrap values, or the percentage of trees in which the associated alleles clustered together, are shown left of the nodes. Tree is drawn to scale, with branch lengths measured in the number of substitutions per site

populations, are similar to those of heterozygote advantage and negative frequency-dependent selection. The general prevalence of balancing selection at MHC (Hedrick, 1998; Piertney & Oliver, 2006), combined with the discovery that these loci affect not just disease resistance but also odor, thus providing a plausible mechanism for nonrandom mating (Yamazaki et al., 1979), led to the reasonable initial expectation that when mating is nonrandom with respect to MHC, it should be disassortative such that mate choice operates in concert with the diversifying effects of balancing selection mediated by pathogens.

As predicted, we found evidence of historical balancing selection at MHC class II in our study population of song sparrows,

presumably reflecting long-standing evolutionary interactions with pathogens. Several codons showed an excess of nonsynonymous relative to synonymous variation, indicating positive Darwinian selection consistent with balancing selection. Likewise, many of the alleles characterized in song sparrows were more similar to alleles found in other species (including some from different avian families) than to other song sparrow sequences: This pattern implies long-term balancing selection at MHC class II, which should promote the maintenance of multiple alleles at these loci and thus be associated with high levels of individual genetic diversity. In contrast, however, free-living song sparrows paired assortatively rather than

disassortatively at these loci, a pattern which should reduce individual genetic diversity of the resultant offspring. Indeed, nestling song sparrows in this population are less diverse at MHC class II than are adults (Watson, 2017), a pattern which might result from assortative mating. Collectively, our findings show that even when balancing selection is operating at MHC, nonrandom mating does not necessarily reinforce these diversifying effects.

Song sparrows in the study population showed several signals of past balancing selection. Although most codons appear to have been primarily subject to purifying selection, as is typical for functional coding loci (Yang & Swanson, 2002), six codons were consistently identified across multiple tests for sequence evolution as being likely to have experienced positive selection. Two of these, based on alignment to human MHC sequence, correspond to antigen-binding positions that are likely to experience particularly intense selection (Hughes & Hughes, 1995). Moreover, four of the six codons identified correspond to the consensus of positively selected sites across passerines (Minias et al., 2018). The remaining positively selected codon (site 24) did not correspond to an antigen-binding position based on sequence alignment with human MHC, nor was it reported to be positively selected site across passerines in general (Minias et al., 2018). Positive selection at MHC is widespread: In a recent review of the literature (25 publications on 25 vertebrate species), positive selection within the peptide-binding region of MHC was reported in every study (Slade, 2018). Although both balancing and directional selection can generate positive selection (Hedrick, 2007), positive selection at MHC is generally considered to result from balancing selection (i.e., new alleles entering a population through mutation or immigration tend to increase in frequency). Still, we do not conclusively rule out the possibility that the observed positive selection could reflect a transient excess of nonsynonymous variants resulting from directional rather than balancing selection.

Further support for the importance of balancing selection in shaping variation at the song sparrow MHC class II comes from comparing sequences in song sparrows to their homologues in other songbird species. While we did not observe trans-species polymorphisms as strictly defined, several alleles found in song sparrows were more similar to heterospecific alleles than to other conspecific alleles. For example, song sparrow sequences SOSP-DAB*18, *19, and *21 were more similar to sequences from Thraupidae, Icteridae, and Emberizinae than to other alleles from song sparrows (family Passerellidae), despite the long-standing divergence between Passerellidae and these other avian families (16.3–32.6 MY; Selvatti, Gonzaga, & de Moraes Russo, 2015). Although we cannot conclusively exclude the possibility of convergent sequence evolution, we think it probable that long-standing balancing selection has maintained these allelic lineages since before the divergence of Passerellidae.

The expectation that sexual selection should favor MHC-disassortative mating stems from the premise that offspring fitness (specifically, resistance to infectious disease) increases with increasing diversity at MHC. Clearly, offspring resulting from more MHC-dissimilar pairings should be more diverse at MHC, and presumably capable of responding to a wider diversity of antigens (Klein, 1986).

However, whether this translates into enhanced lifetime reproductive success, or even into superior disease resistance, is less clear. First, trade-offs between nonadditive (compatible gene) effects, such as heterozygote advantage, and additive (good gene) effects may generate stabilizing selection on MHC diversity. Thus, given a finite level of gene product, excessive diversity at MHC may dilute the protective effects of locally good alleles (de Boer & Perelson, 1993; Kubinak, Nelson, Ruff, & Potts, 2012; Wegner et al., 2003). Second, the risk of parasitic infection associated with insufficient diversity at MHC may trade off with the risk of autoimmune disorders associated with excessive diversity (Apanius, Penn, Slev, Ruff, & Potts, 1997; Bottazzo, Todd, Mirakian, Belfiore, & Pujol-Borrell, 1986; Wegner et al., 2003). Finally, pairings that are maximally dissimilar at MHC may generate outbreeding depression in offspring by disrupting coadapted gene complexes at MHC or other linked loci (Tregenza & Wedell, 2000). This risk may be magnified in hybrid zones, or when two or more locally adapted populations come into contact (Neff, 2004). We suggest that systems in which MHC is highly polygenic and polymorphic, such as song sparrows, are also likely to experience stabilizing rather than directional selection on MHC diversity. Optimal mate choice thus requires resolving trade-offs between the benefits and the costs of high MHC diversity, which promote disassortative and assortative mating, respectively.

Our study did not address directly whether pathogen-mediated balancing selection is currently operating in this population, for example, by comparing disease resistance of more versus less diverse genotypes. However, a cross-sectional analysis in the study population showed that adults were more diverse than nestlings (more alleles per individual) at MHC class II (Watson, 2017). One interpretation of this pattern is that MHC-diverse individuals are more likely to survive to adulthood than their less diverse counterparts. However, this pattern was observed in just one of two years of study, and MHC diversity did not predict overwinter return rates (interpreted as survivorship; Watson, 2017). Thus, if MHC diversity confers a survival advantage in this population, it does not do so consistently every year (Watson, 2017). Moreover, as noted above, assortative mating at MHC could also explain the difference in diversity between age cohorts. Song repertoire size, a sexually selected trait associated with early-life condition, is also highest in males with intermediate rather than maximal MHC class II diversity in this population (Slade, Watson, et al., 2017). Thus, the relationship between MHC class II diversity and fitness in our study population does not appear to be uniformly positive, perhaps because of high standing levels of genetic variation.

We interpret the observed pattern of MHC-assortative pairing as reflecting preferences, actively expressed by one or both sexes, for MHC-similar social mates. In theory, assortative mating could also be explained through passive mechanisms, that is, by highly restricted natal dispersal such that close relatives are more likely to interact than nonrelatives. However, song sparrows are highly mobile, with natal dispersal distances on the order of 6 km (Zink & Dittmann, 1993) and individuals in our study population routinely migrating hundreds of kilometers between breeding and wintering grounds (Kelly et al., 2016). Thus, we do not think it likely that

mobility constrains social mate choice at the geographic scale we investigated (<1 km²). Moreover, based on band-recapture analysis, natal philopatry is low (5%–15% of new recruits each year were banded as nestlings; Stewart & MacDougall-Shackleton, 2008) and strongly male-biased, further reducing the likelihood of close relatives encountering one another as potential mates. Similarly, we cannot exclude the possibility that the observed mating patterns reflect active inbreeding, such that individuals preferentially mate with close relatives, which would also result in mated pairs being relatively similar at MHC. However, we think this unlikely as song sparrows do not appear to either prefer or avoid close relatives as social mates (Keller & Arcese, 1998). Behavioral tests under controlled conditions, for example, testing preferences for preen oil odor from MHC-similar versus MHC-dissimilar individuals (Leclaire et al., 2017), represent a critical next step to confirm (a) whether assortative pairing observed in the field reflects preferences expressed in the laboratory, and (b) whether chemical cues provide a mechanism for assessing MHC similarity in songbirds, as they do in other vertebrates (Brennan & Zufall, 2006; Leclaire et al., 2017; Milinski et al., 2005). Comparing the pairing behavior of free-living animals to preferences expressed under standardized conditions will also shed light on potential trade-offs between direct and indirect benefits associated with social and genetic mating decisions, and whether MHC-related preferences are similar in both sexes.

Regardless of whether assortative mating stems from expressed mating preferences or from some other mechanism, its effects on individual genetic diversity (i.e., increasing homozygosity) oppose those of balancing selection (i.e., increasing divergence). In systems where strong and long-standing balancing selection, presumably mediated by past arms races with pathogens, has generated high variation at immune loci, the ability of mate choice to “balance” balancing selection (i.e., to avoid producing offspring with an excessive number of different alleles) may be advantageous. Whether the assortative mating we observed in this population at the hypervariable class II MHC also occurs at the less variable class I MHC remains an open, and important, question. In particular, whereas song repertoire size does not increase monotonically with class II diversity in this population (Slade, Watson, et al., 2017), resistance to hematozoan infection does increase monotonically with class I diversity (Slade, Sarquis-Adamson, et al., 2017).

Assortative mating does not necessarily reduce genetic diversity at the population level, at least under monogamous mating systems where the ability to attract a mate does not differ appreciably between genotypes. Assortative mating at MHC may thus balance the costs and benefits of genetic diversity within individuals, without constraining the evolutionary potential of populations to adapt to future changes in pathogen regime.

ACKNOWLEDGMENTS

This work was supported by the Natural Sciences and Engineering Research Council of Canada (Discovery Grant

award 293123-2012RGPIN to EAMS); the Society of Canadian Ornithologists (Taverner Award to JWGS); and the American Ornithologists' Union (Hesse Research Award to JWGS). JWGS was supported by an Ontario Graduate Scholarship. We thank Greg Gloor for assisting with bioinformatic analysis; the Queen's University Biological Station for logistic support; Shawn Garner, Nusha Keyghobadi, and Bryan Neff for advice on study design and interpretation; and Tosha Kelly, Alannah Lymburner, and Dominique Potvin for field assistance.

CONFLICT OF INTEREST

The authors declared that they have no conflict of interests.

AUTHOR CONTRIBUTIONS


JWGS and EAM-S designed the experiment. JWGS, EAM-S, and MJW collected field data. JWGS and MJW performed DNA extractions, PCR, and bioinformatics. JWGS performed statistical analyses. JWGS and EAM-S prepared the manuscript.

DATA ACCESSIBILITY

Data available from the Dryad Digital Repository: <https://doi.org/10.5061/dryad.9v2362d>. DNA sequences have been submitted to GenBank (accession numbers KX263957-KX264148; KX375230-KX375341; MF197785-MF197843; MH670952-MH671105).

ORCID

Joel W. G. Slade  <https://orcid.org/0000-0002-4587-5033>

Elizabeth A. MacDougall-Shackleton  <https://orcid.org/0000-0002-4138-6398>

REFERENCES

- Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. *Journal of Molecular Biology*, 215(3), 403–410. [https://doi.org/10.1016/S0022-2836\(05\)80360-2](https://doi.org/10.1016/S0022-2836(05)80360-2)
- Apanius, V., Penn, D., Slev, P. R., Ruff, L. R., & Potts, W. K. (1997). The nature of selection on the major histocompatibility complex. *Critical Reviews in Immunology*, 17(2), 179–224. <https://doi.org/10.1615/CritRevImmunol.v17.i2.40>
- 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(3), 363–377. <https://doi.org/10.1046/j.1420-9101.2003.00531.x>
- 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 of London B: Biological Sciences*, 279(1748), 4778–4785. <https://doi.org/10.1098/rspb.2012.1885>
- Bollmer, J. L., Dunn, P. O., Whittingham, L. A., & Wimpee, C. (2010). Extensive MHC class II B gene duplication in a Passerine, the common yellowthroat (*Geothlypis trichas*). *Journal of Heredity*, 101(4), 448–460. <https://doi.org/10.1093/jhered/esq018>

- Bonneaud, C., Chastel, O., Federici, P., Westerdahl, H., & Sorci, G. (2006). Complex Mhc-based mate choice in a wild passerine. *Proceedings of the Royal Society of London B: Biological Sciences*, 273(1590), 1111–1116. <https://doi.org/10.1098/rspb.2005.3325>
- Bos, D. H., Williams, R. N., Gopurenko, D., Bulut, Z., & Dewoody, J. A. (2009). Condition-dependent mate choice and a reproductive disadvantage for MHC-divergent male tiger salamanders. *Molecular Ecology*, 18(15), 3307–3315. <https://doi.org/10.1111/j.1365-294X.2009.04242.x>
- Bottazzo, G. F., Todd, I., Mirakian, R., Belfiore, A., & Pujol-Borrell, R. (1986). Organ-specific autoimmunity: A 1986 overview. *Immunological Reviews*, 94(1), 137–169. <https://doi.org/10.1111/j.1600-065X.1986.tb01168.x>
- Brambilla, A., Keller, L., Bassano, B., & Grossen, C. (2018). Heterozygosity–fitness correlation at the major histocompatibility complex despite low variation in Alpine ibex (*Capra ibex*). *Evolutionary Applications*, 11(5), 631–644. <https://doi.org/10.1111/eva.12575>
- Brennan, P. A., & Zufall, F. (2006). Pheromonal communication in vertebrates. *Nature*, 444(7117), 308–315.
- 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.
- Burnham, K. P., & Anderson, D. R. (2002). *Model selection and multimodel inference: A practical information-theoretic approach*. Berlin, Germany: Springer Science & Business Media.
- Chen, J., Bittinger, K., Charlson, E. S., Hoffmann, C., Lewis, J., Wu, G. D., ... Li, H. (2012). Associating microbiome composition with environmental covariates using generalized UniFrac distances. *Bioinformatics*, 28, 2106–2113. <https://doi.org/10.1093/bioinformatics/bts342>
- deBoer, R. J., & Perelson, A. S. (1993). How diverse should the immune system be? *Proceedings of the Royal Society of London B: Biological Sciences*, 252(1335), 171–175. <https://doi.org/10.1098/rspb.1993.0062>
- Doherty, P. C., & Zinkernagel, R. M. (1975). H-2 compatibility is required for T-cell-mediated lysis of target cells infected with lymphocytic choriomeningitis virus. *Journal of Experimental Medicine*, 141(2), 502–507. <https://doi.org/10.1084/jem.141.2.502>
- Edgar, R. C., Haas, B. J., Clemente, J. C., Quince, C., & Knight, R. (2011). UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics*, 27(16), 2194–2200. <https://doi.org/10.1093/bioinformatics/btr381>
- Edwards, S. V., Gasper, J., & March, M. (1998). Genomics and polymorphism of *Agph-DAB1*, an Mhc class II B gene in red-winged blackbirds (*Agelaius phoeniceus*). *Molecular Biology and Evolution*, 15(3), 236–250. <https://doi.org/10.1093/oxfordjournals.molbev.a025921>
- Forsberg, L. A., Dannewitz, J., Petersson, E., & Grahn, M. (2007). Influence of genetic dissimilarity in the reproductive success and mate choice of brown trout – females fishing for optimal MHC dissimilarity. *Journal of Evolutionary Biology*, 20(5), 1859–1869. <https://doi.org/10.1111/j.1420-9101.2007.01380.x>
- Freeman-Gallant, C. R., Johnson, E. M., Saponara, F., & Stanger, M. (2002). Variation at the major histocompatibility complex in savannah sparrows. *Molecular Ecology*, 11(6), 1125–1130. <https://doi.org/10.1046/j.1365-294X.2002.01508.x>
- Galaverni, M., Caniglia, R., Milanese, P., Lapalombella, S., Fabbri, E., & Randi, E. (2016). Choosy wolves? Heterozygote advantage but no evidence of MHC-based disassortative mating. *Journal of Heredity*, 107(2), 134–142. <https://doi.org/10.1093/jhered/esv090>
- Gloor, G. B., Hummelen, R., Macklaim, J. M., Dickson, R. J., Fernandes, A. D., MacPhee, R., & Reid, G. (2010). Microbiome profiling by Illumina sequencing of combinatorial sequence-tagged PCR products. *PLoS ONE*, 5, e15406–e15406.
- Hedrick, P. W. (1998). Balancing selection and MHC. *Genetica*, 104(3), 207–214. <https://doi.org/10.1023/A:1026494212540>
- Hedrick, P. W. (2007). Balancing selection. *Current Biology*, 17(7), R230–R231. <https://doi.org/10.1016/j.cub.2007.01.012>
- Hess, C. M., & Edwards, S. V. (2002). The evolution of the major histocompatibility complex in birds. *BioScience*, 52(5), 423–431. [https://doi.org/10.1641/0006-3568\(2002\)052\[0423:TEOTMH\]2.0.CO;2](https://doi.org/10.1641/0006-3568(2002)052[0423:TEOTMH]2.0.CO;2)
- Hughes, A. L., & Hughes, M. K. (1995). Natural selection on the peptide-binding regions of major histocompatibility complex molecules. *Immunogenetics*, 42(4), 233–243. <https://doi.org/10.1007/BF00176440>
- Hughes, A. L., & Nei, M. (1989). Evolution of the major histocompatibility complex: Independent origin of nonclassical class I genes in different groups of mammals. *Molecular Biology and Evolution*, 6(6), 559–579. <https://doi.org/10.1093/oxfordjournals.molbev.a040573>
- Kamiya, T., O'Dwyer, K., Westerdahl, H., Senior, A., & Nakagawa, S. (2014). A quantitative review of MHC-based mating preference: The role of diversity and dissimilarity. *Molecular Ecology*, 23(21), 5151–5163. <https://doi.org/10.1111/mec.12934>
- Keller, L. F., & Arcese, P. (1998). No evidence for inbreeding avoidance in a natural population of song sparrows (*Melospiza melodia*). *American Naturalist*, 152(3), 380–392.
- Kelly, T. R., MacGillivray, H. L., Sarquis-Adamson, Y., Watson, M. J., Hobson, K. A., & MacDougall-Shackleton, E. A. (2016). Seasonal migration distance varies with natal dispersal and predicts parasitic infection in song sparrows. *Behavioral Ecology and Sociobiology*, 70, 1857–1866. <https://doi.org/10.1007/s00265-016-2191-2>
- Klein, J. (1980). Generation of diversity at MHC loci: Implications for T cell receptor repertoires. In: *Immunology 80*. Academic, London, UK.
- Klein, J. (1986). *Natural history of the major histocompatibility complex*. New York, NY: John Wiley and Sons.
- Kosakovsky, S. L., & Frost, S. D. W. (2005). Not so different after all: A comparison of methods for detecting amino acid sites under selection. *Molecular Biology and Evolution*, 22(5), 1208–1222. <https://doi.org/10.1093/molbev/msi105>
- Kubinak, J. L., Nelson, A. C., Ruff, J. S., & Potts, W. K. (2012). Tradeoffs limiting MHC heterozygosity. In G. E. Demas & R. J. Nelson (Eds.), *Ecoimmunology* (pp. 225–258). New York, NY, USA: Oxford University Press Inc.
- Kuduk, K., Babik, W., Bellemain, E., Valentini, A., Zedrosser, A., Taberlet, P., ... Radwan, J. (2014). No evidence for the effect of MHC on male mating success in the brown bear. *PLoS One*, 9(12), e113414. <https://doi.org/10.1371/journal.pone.0113414>
- Kumar, S., Stecher, G., & Tamura, K. (2016). MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Molecular Biology and Evolution*, 33, 1870–1874.
- Leclaire, S., Strandh, M., Mardon, J., Westerdahl, H., & Bonadonna, F. (2017). Odour-based discrimination of similarity at the major histocompatibility complex in birds. *Proceedings of the Royal Society of London B: Biological Sciences*, 284(1846), 20162466. <https://doi.org/10.1098/rspb.2016.2466>
- Lewis, K. (1998). Pathogen resistance as the origin of kin altruism. *Journal of Theoretical Biology*, 193(2), 359–363. <https://doi.org/10.1006/jtbi.1998.0725>
- Lozupone, C., & Knight, R. (2005). UniFrac: A new phylogenetic method for comparing microbial communities. *Applied and Environmental Microbiology*, 71(12), 8228–8235. <https://doi.org/10.1128/AEM.71.12.8228-8235.2005>
- Manly, B. (1997). *Randomization, bootstrapping and Monte Carlo methods in biology*, 2nd ed. London, UK: Chapman and Hall.
- McClelland, E. E., Penn, D. J., & Potts, W. K. (2003). Major histocompatibility complex heterozygote superiority during coinfection. *Infection and Immunity*, 71(4), 2079–2086. <https://doi.org/10.1128/IAI.71.4.2079-2086.2003>
- Milinski, M., Griffiths, S., Wegner, K. M., Reusch, T. B. H., Haas-Assenbaum, A., & Boehm, T. (2005). Mate choice decisions of stickleback females predictably modified by MHC peptide ligands.

- Proceedings of the National Academy of Sciences of the United States of America*, 102(12), 4414–4418. <https://doi.org/10.1073/pnas.0408264102>
- Miller, H. C., Moore, J. A., Nelson, N. J., & Daugherty, C. H. (2009). Influence of major histocompatibility complex genotype on mating success in a free-ranging reptile population. *Proceedings of the Royal Society of London B: Biological Sciences*, 276(1662), 1695–1704.
- Minias, P., Pikus, E., Whittingham, L. A., & Dunn, P. O. (2018). A global analysis of selection at the avian MHC. *Evolution*, 72(6), 1278–1293. <https://doi.org/10.1111/evo.13490>
- Murrell, B., Moola, S., Mabona, A., Weighill, T., Sheward, D., Pond, S. L. K., & Scheffler, K. (2013). FUBAR: A fast, unconstrained Bayesian approximation for inferring selection. *Molecular Biology and Evolution*, 30(5), 1196–1205. <https://doi.org/10.1093/molbev/mst030>
- Murrell, B., Wertheim, J. O., Moola, S., Weighill, T., Scheffler, K., & Pond, S. L. K. (2012). Detecting individual sites subject to episodic diversifying selection. *PLoS Genetics*, 8(7), e1002764. <https://doi.org/10.1371/journal.pgen.1002764>
- Neff, B. D. (2004). Stabilizing selection on genomic divergence in a wild fish population. *Proceedings of the National Academy of Sciences of the United States of America*, 101(8), 2381–2385. <https://doi.org/10.1073/pnas.0307522100>
- Neff, B. D., Garner, S. R., Heath, J. W., & Heath, D. D. (2008). The MHC and non-random mating in a captive population of Chinook salmon. *Heredity*, 101(2), 175–185. <https://doi.org/10.1038/hdy.2008.43>
- Niskanen, A. K., Kennedy, L. J., Ruokonen, M., Kojola, I., Lohi, H., Isomursu, M., ... Aspi, J. (2014). Balancing selection and heterozygote advantage in major histocompatibility complex loci of the bottlenecked Finnish wolf population. *Molecular Ecology*, 23(4), 875–889. <https://doi.org/10.1111/mec.12647>
- Penn, D. J., Damjanovich, K., & Potts, W. K. (2002). MHC heterozygosity confers a selective advantage against multiple-strain infections. *Proceedings of the National Academy of Sciences of the United States of America*, 99(17), 11260–11264. <https://doi.org/10.1073/pnas.162006499>
- Phillips, K. P., Cable, J., Mohammed, R. S., Herdegen-Radwan, M., Raubic, J., Przesmycka, K. J., ... Radwan, J. (2018). Immunogenetic novelty confers a selective advantage in host–pathogen coevolution. *Proceedings of the National Academy of Sciences of the United States of America*, 115(7), 1552–1557. <https://doi.org/10.1073/pnas.1708597115>
- Piertney, S. B., & Oliver, M. K. (2006). The evolutionary ecology of the major histocompatibility complex. *Heredity*, 96(1), 7–21. <https://doi.org/10.1038/sj.hdy.6800724>
- Potvin, D. A., & MacDougall-Shackleton, E. A. (2009). Parental investment amplifies effects of genetic complementarity on growth rates in song sparrows. *Melospiza Melodia. Animal Behaviour*, 78(4), 943–948. <https://doi.org/10.1016/j.anbehav.2009.07.023>
- R. Core Team (2017). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.
- Roberts, S. C. (2009). Complexity and context of MHC-correlated mating preferences in wild populations. *Molecular Ecology*, 18(15), 3121–3123. <https://doi.org/10.1111/j.1365-294X.2009.04244.x>
- Robinson, J. (2003). IMGT/HLA and IMGT/MHC: Sequence databases for the study of the major histocompatibility complex. *Nucleic Acids Research*, 31(1), 311–314. <https://doi.org/10.1093/nar/gkg070>
- Selvatti, A. P., Gonzaga, L. P., & de Moraes Russo, C. A. (2015). A Paleogene origin for crown passerines and the diversification of the Oscines in the New World. *Molecular Phylogenetics and Evolution*, 88, 1–15. <https://doi.org/10.1016/j.ympev.2015.03.018>
- Sepil, I., Radersma, R., Santure, A. W., De Cauwer, I., Slate, J., & Sheldon, B. C. (2015). No evidence for MHC class I-based disassortative mating in a wild population of great tits. *Journal of Evolutionary Biology*, 28(3), 642–654. <https://doi.org/10.1111/jeb.12600>
- Sin, Y. W., Annavi, G., Newman, C., Buesching, C., Burke, T., Macdonald, D. W., & Dugdale, H. L. (2015). MHC class II-assortative mate choice in European badgers (*Meles meles*). *Molecular Ecology*, 24(12), 3138–3150. <https://doi.org/10.1111/mec.13217>
- Slade, J. W. G. (2018). The major histocompatibility complex in song sparrows: Immunity, signals, and mate choice. PhD thesis. University of Western Ontario, London, ON, Canada.
- Slade, J. W. G., Sarquis-Adamson, Y., Gloor, G. B., Lachance, M.-A., & MacDougall-Shackleton, E. A. (2017). Population differences at MHC do not explain enhanced resistance of song sparrows to local parasites. *Journal of Heredity*, 108(2), 127–134.
- Slade, J. W. G., Watson, M. J., Kelly, T. R., Gloor, G. B., Bernards, M. A., & MacDougall-Shackleton, E. A. (2016). Chemical composition of preen wax reflects major histocompatibility complex similarity in songbirds. *Proceedings of the Royal Society of London B: Biological Sciences*, 283(1842), 20161966.
- Slade, J. W. G., Watson, M. J., & MacDougall-Shackleton, E. A. (2017). Birdsong signals individual diversity at the major histocompatibility complex. *Biology Letters*, 13(11), 20170430. <https://doi.org/10.1098/rsbl.2017.0430>
- Slade, R. W., & McCallum, H. I. (1992). Overdominant vs. frequency-dependent selection at MHC loci. *Genetics*, 132(3), 861–862.
- Stewart, K. A., & MacDougall-Shackleton, E. A. (2008). Local song elements indicate local genotypes and predict physiological condition in song sparrows *Melospiza melodia*. *Biology Letters*, 4(3), 240–242.
- Tregenza, T., & Wedell, N. (2000). Genetic compatibility, mate choice and patterns of parentage: Invited review. *Molecular Ecology*, 9(8), 1013–1027. <https://doi.org/10.1046/j.1365-294x.2000.00964.x>
- Trowsdale, J. (1995). “Both man & bird & beast”: Comparative organization of MHC genes. *Immunogenetics*, 41(1), 1–17.
- Watson, M. (2017). MHC class IIβ diversity as a correlate of neutral-locus similarity and diversity, and a predictor of overwinter return, in song sparrows (*Melospiza melodia*). MSc thesis. University of Western Ontario, London, ON, Canada.
- 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(3), 773–777. <https://doi.org/10.1093/molbev/msx335>
- Wegner, K. M., Kalbe, M., Kurtz, J., Reusch, T. B. H., & Milinski, M. (2003). Parasite selection for immunogenetic optimality. *Science*, 301, 1343. <https://doi.org/10.1126/science.1088293>
- Westerdahl, H. (2004). No evidence of an MHC-based female mating preference in great reed warblers. *Molecular Ecology*, 13(8), 2465–2470. <https://doi.org/10.1111/j.1365-294X.2004.02238.x>
- Whelan, S., & Goldman, N. (2001). A general empirical model of protein evolution derived from multiple protein families using a maximum-likelihood approach. *Molecular Biology and Evolution*, 18(5), 691–699. <https://doi.org/10.1093/oxfordjournals.molbev.a003851>
- Xu, B., & Yang, Z. (2013). PAMLx: A graphical user interface for PAML. *Molecular Biology and Evolution*, 30(12), 2723–2724. <https://doi.org/10.1093/molbev/mst179>
- Yamazaki, K., Yamaguchi, M., Baranoski, L., Bard, J., Boyse, E. A., & Thomas, L. (1979). Recognition among mice. Evidence from the use of a Y-maze differentially scented by congenic mice of different major histocompatibility types. *Journal of Experimental Medicine*, 150(4), 755–760.
- Yang, Z. (2000). Maximum likelihood estimation on large phylogenies and analysis of adaptive evolution in human influenza virus A. *Journal of Molecular Evolution*, 51(5), 423–432. <https://doi.org/10.1007/s002390010105>
- Yang, Z. (2007). PAML 4: Phylogenetic analysis by maximum likelihood. *Molecular Biology and Evolution*, 24(8), 1586–1591. <https://doi.org/10.1093/molbev/msm088>

- Yang, Z., & Swanson, W. J. (2002). Codon-substitution models to detect adaptive evolution that account for heterogeneous selective pressures among site classes. *Molecular Biology and Evolution*, 19(1), 49–57. <https://doi.org/10.1093/oxfordjournals.molbev.a003981>
- Yang, Z., Wong, W. S. W., & Nielsen, R. (2005). Bayes empirical Bayes inference of amino acid sites under positive selection. *Molecular Biology and Evolution*, 22(4), 1107–1118. <https://doi.org/10.1093/molbev/msi097>
- Zink, R. M., & Dittmann, D. L. (1993). Gene flow, refugia, and evolution of geographic variation in the song sparrow (*Melospiza melodia*). *Evolution*, 47(3), 717–729. <https://doi.org/10.1111/j.1558-5646.1993.tb01228.x>

How to cite this article: Slade JWG, Watson MJ, MacDougall-Shackleton EA. “Balancing” balancing selection? Assortative mating at the major histocompatibility complex despite molecular signatures of balancing selection. *Ecol Evol.* 2019;9:5146–5157. <https://doi.org/10.1002/ece3.5087>