8-28-2023 10:00 AM

Random Mating in the Face of Balancing Selection at the Major Histocompatibility Complex Class I in Song Sparrows (Melospiza melodia)

Scout R.L Thompson, *Western University*

Supervisor: MacDougall-Shackleton, Elizabeth A., *The University of Western Ontario*

A thesis submitted in partial fulfillment of the requirements for the Master of Science degree in Biology

Abstract

The major histocompatibility complex (MHC) is a large group of genes encoding cell-surface proteins that recognize and bind pathogens to initiate an adaptive immune response. MHC loci experience intense pathogen-mediated selection which may be directional, where specific alleles provide the best disease protection, or balancing, where rare alleles or diverse combinations are most protective. However, balancing selection (specifically heterozygote advantage) is more common and often accompanied by disassortative mating. I sought to use genetic and behavioural information to evaluate whether balancing selection and disassortative mating occur at MHC class I (MHCI) in a population of song sparrows (*Melospiza melodia*). Despite evidence of balancing selection (high ratios of nonsynonymous to synonymous sequence variation and trans-species polymorphisms), the genetic distance between social mates was no different than expected under random mating. This may suggest a low impact of MHCI diversity on lifetime reproductive success, or an inability to discern MHCI genotype.

Keywords

Melospiza melodia, Major histocompatibility complex, balancing selection, sexual selection, random mating, trans-species polymorphisms, positive selection

Summary for Lay Audience

Throughout their evolutionary history, animals and other living things have had to defend themselves against bacteria, viruses, and other infectious diseases. A key part of this defense is a group of genes known as the major histocompatibility complex (MHC). MHC genes make receptors that cells use to alert the immune system to infections, either inside the cell such as viruses (class 1) or outside the cell such as bacteria (class 2). The kind of infections that a receptor can recognize depends on the versions of the genes (alleles) that the host carries. Animals often prefer mates with different MHC alleles from their own, probably because this allows them to produce offspring that have several different alleles and are protected against a wider range of infections. Surprisingly, however, a previous study of song sparrows found that these birds chose mates that were similar, not different, to themselves at class 2 MHC. I asked whether this unexpected pattern also happens at class 1 MHC. I collected DNA from blood samples, and identified which birds were mated pairs by providing them with unique combinations of coloured plastic leg bands and then watching them interact in the wild. I chose to use song sparrows for this study as they can assess MHCII similarity through smell – although whether this is true for MHCI is unknown. Surprisingly, I found that song sparrows seem to pick their partners at random with respect to similarity or difference at class 1 MHC. This may suggest that song sparrows prioritize other considerations in mate choice, such as if a particular mate will be a good parent, or that unlike class 2 MHC, song sparrows simply have no way to assess similarity at class 1. Identifying whether mate choice improves animals' ability to produce healthy offspring is important for conservation efforts. To predict a species' ability to evolve in response to a changing environment, we must know how much genetic diversity is needed, and whether mate choice helps to maintain high levels of diversity.

Co-Authorship Statement

The work contained within this thesis relied in part on DNA samples and pairing information collected by Dominique Potvin, Joel Slade, Leanne Grieves, and Elizabeth (Beth) MacDougall-Shackleton. Dominique and Joel collected pairing data and samples during the years 2007, and 2014 /2015, respectively. Leanne provided guidance and generated the barcodes used to create my custom Illumina primers. Beth provided all required funding, contributed to both data collection and study design, gave advice on analyses, and helped edit all sections. Accordingly, publications resulting from this thesis will include Dominique Potvin, Joel Slade, Leanne Grieves and Beth MacDougall-Shackleton as co-authors.

Acknowledgements

Where do I even start? This thesis wouldn't be possible without the help, support, and guidance of far too many people that I could not possibly fit within the confines of this thesis. To all friends, family, and anyone else who I did not specifically mention, I thank you immensely for all that you've done and hope that you will forgive me for refraining from making this section 200 pages long. In no particular order:

To Garth Casbourn; you are truly one of the coolest and most interesting people I have ever met. I cannot thank you enough for your guidance and expertise out in the field and for sharing with me your astonishing knowledge of birds. Thank you for all you have done for my lab, I expect fantastic things from you and can't wait to see what you do next.

To Leanne Grieves; though we have only met over zoom, I am a huge admirer and can only hope to be as great of a biologist as you one day. Thank you so much for your guidance, and for your help with troubleshooting and handling other lab matters.

To Joel Slade and Dominique Potvin; though we have not met, I of course must extend a huge thank you for your work and all that you have done. You truly laid the foundation for my thesis and without you, it wouldn't have been possible.

To my advisory and proposal committee: David Coltman, David Smith, and Nusha Keyghobadi; thank you so much for your insightful questions and helpful advice. You all helped challenge my thinking and allowed me to push myself to become a better and more effective scientist and communicator. I greatly appreciate the time and effort you took to read and engage with my work and presentations.

To Greg Gloor; thank you for providing the Perl script and server that were necessary to complete this project. Thank you especially for agreeing to meet with me to go over things when I couldn't help but doubt myself on account of having no idea what I was doing.

To David Carter and the team at LRGC: thank you for your help and hard work in conducting the Illumina sequencing for this project. None of my analysis would have been possible without your contribution.

To Sydney Doucette and Justin To; thank you for being my treasured colleagues. Thank you for your input and support, for putting up with my antics, and for being a blast to work with. There are no other people I'd rather trudge through the snow with at 6am than you.

To Ben Klassen, my best friend, and the only person to get me truly and deeply. I would not be who I am today without you and cannot thank you enough for sticking by my side for the last 12 years. Thank you for your support and laughter, and for being the person I can always text whatever weird thing is going through my head at 2am. I hope I can be blessed with your friendship for many, many more years to come.

To Nathaniel Spadafora, my partner, my Djungelskog. I love you more than words can express. You have been my biggest cheerleader and supporter throughout this entire process. Thank you for your love and encouragement, your dedication to understanding my work, for always telling me how proud you are of me, and for consoling me after my PCRs inevitably failed. I could not have achieved any of this without you.

To my beloved cats; Elliott Moose, Mabel, and my dearly departed Adeline. You may not be able to read this, but if you could I would thank you all for your much needed love, laughter, and comfort. Thank you for being there to cheer me up, and for always giving me a reason to get out of bed in the morning. I love you little weirdos so much, even if you're all very stinky.

To my supervisor, my captain, and the brains behind it all, Beth MacDougall-Shackleton; thank you so much. For everything. I don't think I would be where I am now if you hadn't approached me about joining your lab. Your love for your work is infectious, and I am nothing but awe-struck and inspired by your hard work and dedication to your craft. Thank you for your infinite wisdom and advice, your help and support, and for always being there for me. I cannot express my deep enough gratitude for all that you've done to allow this project to come together, it truly would not have happened without you. I may not always understand what exactly you saw in me during my honours-thesis presentation, but I am extremely happy that you took the chance on me anyway and helped me grow into the scientist and researcher I am today. Thank you.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

+G – With 5 Discrete Gamma Categories

+I – With Invariant Sites

+R - With Relaxed Assumptions of Gamma Distribution

AIC – Akaike Information Criterion

aRLT – Approximate Likelihood Ratio Test

BIC – Bayesian Information Criterion

BLAST – Basic Local Alignment Search Tool

bp – Base Pair

CD4+s – T Helper Cells

CHD - Chromo Helicase DNA-binding

CTLs – Cytotoxic T Lymphocytes

CWS – Canadian Wildlife Service

dN – Nonsynonymous Substitution Rate

DNA – Deoxyribonucleic Acid

dNTPs – Deoxynucleotide Triphosphate

dS – Synonymous Substitution Rate

EDTA - Ethylenediaminetetraacetic Acid

F1 – First Generation

F2 – Second Generation

FEL – Fixed Effects Likelihood

FUBAR – Fast Unconstrained Bayesian Approximation

GTR – General Time Reversible

HKY85 - Hasegawa-Kishino-Yano, 85

Llk – Log Likelihood

LRS – Likelihood Ratio Statistic; also, Lifetime Reproductive Success

LRT – Likelihood Ratio Test

MEME – Mixed Effects Maximum Likelihood

$MgCl_2$ – Magnesium Chloride

MHC – Major Histocompatibility Complex

MHC-B – The Major Histocompatibility B Complex

MHCI – Major Histocompatibility Complex Class I

MHCII – Major Histocompatibility Complex Class II

MHC-Y – The Major Histocompatibility Y Complex

MOE – Main Olfactory Epithelium

MUSCLE - Multiple Sequence Comparison by Log-Expectation

NGS – Next Generation Sequencing

NKs – Natural Killer Cells

PAML – Phylogenetic Analysis by Maximum Likelihood

PBR – Peptide Binding Region

PCR – Polymerase Chain Reaction

popART - Population Analysis with Reticulate Trees

SD – Standard Deviation

SDS - Sodium Dodecyl Sulfate

SEM – Standard Error of the Mean

SH – Shimodaira-Hasegawa

SMS – Smart Model Selection

Taq – Taq Polymerase

TE - Trisaminomethane-Ethylenediaminetetraacetic Acid

TN93 – Tamura-Nei, 93

Tris - Trisaminomethane

TSPs – Trans-Species Polymorphisms

UV – Ultraviolet

VNO – Vomeronasal Organ

1 *Introduction*

1.1 *Sexual selection and fitness*

Sexual selection refers to a special case of selection in which members of one sex compete, either directly or indirectly, for access to individuals of the opposite sex (Andersson 1994). This theory has been used to explain the stark and widespread sexual dimorphism that can be seen in many species, and why sexually dimorphic traits persist despite often involving costs to the survival of the bearer (Andersson 1994). Typically, but not always, males are the competitive sex and females are the choosy sex, such that females do not always mate with the first male they encounter, and males of many species are adorned with extensive colour and ornamentation (Andersson 1994).

In some systems, female preference for particular males (e.g. those with attractive ornaments) appears to be driven by direct fitness benefits, such as the provision of resources and/or protection, that are signaled by the ornaments in question. For example, in eastern bluebirds (*Sialia sialis*), males with larger and brighter breast patches provide significantly better parental care than their duller counterparts (Siefferman and Hill 2003). However, mate choice may also offer indirect (genetic) benefits that enhance offspring survival or quality independent of any material benefits offered. These genetic effects can be additive, where the fitness benefit of a given allele is independent of the rest of the genome, or non-additive, where the fitness benefit of a given allele is dependent on the rest of an individual's allelic repertoire (Neff and Pitcher 2005). Beneficial alleles that fall under these two categories have been dubbed "good genes" and "compatible genes", respectively (Mays and Hill 2004; Neff and Pitcher 2005).

1.1.1 *Additive effects (good genes)*

Additive effects on fitness, and the related concept of good genes, were proposed as one of the first explanations for the persistence of female mate choice in species where the direct benefits of mate choice are not evident (Mays and Hill 2004). For the effects of a gene to be considered additive, they must be both heritable and provide fitness benefits independently of the rest of the genome. Males who possess 'good' genes are often

universally sought after by females as these males will provide fitness benefits to their offspring regardless of their own genotype. This can be seen in pronghorn (*Antilocapra americana*) where female mate choice tends to converge on the same small proportion of demonstrably vigorous males (Byers and Waits 2006). The offspring of these males have significantly higher survivorship than those with less athletically vigorous (and hence less attractive) sires, supporting the idea of genes with population wide advantage (Byers and Waits 2006; Neff and Pitcher 2005).

It should be noted that because of geographic differences in environmental selection pressures, even genes that provide a population-wide advantage may not be beneficial across all populations. These genes, which are only beneficial to a single population within a specific environment, are known as "locally good genes". This can be seen in three-spined sticklebacks (*Gasterosteus acueatus*) where individuals with locally adapted genotypes at the major histocompatibility complex (MHC), an important family of immune genes, had significantly lower parasite loads and showed reduced infection intensities (measured as an index of parasite density and composition) when exposed to sympatric parasites (Eizaguirre et al. 2012). This may explain earlier findings of female sticklebacks preferring to mate with males possessing particular MHC haplotypes, which were also found to confer better parasite resistance (Eizaguirre et al. 2009). Still, the fact that such alleles enhance fitness in a particular population or environment regardless of the rest of an individual's genotype indicates additive, instead of non-additive, effects on fitness.

1.1.2 *Non-additive effects (compatible genes)*

A more recent explanation for the persistence of mate choice where there are no obvious direct benefits involves non-additive effects on fitness, and the related concept of compatible genes. Compatible genes only provide benefits to fitness in specific genotypic contexts (Neff and Pitcher 2005). For example, at loci subject to heterozygote advantage, the most compatible mate for a given female (i.e. the mate most likely to produce heterozygous offspring) depends on her own genotype. Thus, in systems dominated by compatible-gene effects on fitness, individual females will have their own idiosyncratic and unique preferences regarding mate selection.

Compatible genes may operate through dominance or overdominance at homologous sites, or by having epistatic effects at other loci (Neff and Pitcher 2005). Overdominance, otherwise known as heterozygote advantage, refers to when individuals heterozygous at specific loci have a greater fitness advantage than homozygous individuals (Brown 1997). This is evident in the classic example of humans heterozygous for the mutation causing sickle cell anemia having increased protection against severe malarial infection without succumbing to the severe vascular complications that afflict homozygous individuals (Archer et al. 2018). Under heterozygote advantage, individuals that choose mates with dissimilar genotypes, a phenomenon known as disassortative mating, should produce offspring with higher fitness (Tregenza and Wedell 2000). Indeed, a mathematical model based on mimetic color pattern in the Numata longwing (*Heliconius numata*) showed that heterozygote advantage favours the evolution of disassortative mating (Maisonneuve et al. 2021). This model might explain why rhesus macaques (*Macaca mulatta*) heterozygous at an immune system locus sire significantly more offspring than homozygotes (Sauermann et al. 2001).

However, the benefits of heterozygosity may not always be due to overdominance. In laboratory mice (*Mus musculus*), while heterozygous individuals were found to have higher resistance to multiple-strain salmonella infections than the average of their parental homozygotes, they were not more resistant than both (Penn et al. 2002). This implies a system of incomplete dominance, characterized by one allele having a stronger effect on the phenotype instead of it being an intermediate of the homozygotes (Penn et al. 2002). These results may provide another explanation for disassortative mating preferences.

Additionally, non-additive genetic effects may not always favour maximizing diversity. In some cases, disassortative mating between highly dissimilar individuals may result in outbreeding depression, where optimal allele combinations are broken up. An extreme case of this was seen when different subspecies of ibex (*Capra ibex aegagrus* and *C. ibex nubiuna*) from Turkey and the Sinai were introduced to help recover another population in Czechoslovakia (*C. ibex ibex*) (Turček and Hickey 1951). Despite the fertility of the F1 hybrids, the F2 generation had extremely low survival rates, and this eventually resulted

in population extinction (Turček and Hickey 1951). Cases such as this may partially explain why individuals of some species prefer to mate assortatively with individuals genetically similar to themselves.

1.2 *The major histocompatibility complex (MHC)*

Parasites and pathogens are everywhere and can often have catastrophic effects on host fitness, whether through tissue damage or competitive resource consumption (Ewald 1994; Nash et al. 2015). In response, multicellular organisms have evolved diverse systems of immune defense. One such defense, the major histocompatibility complex (MHC), is a large group of genes found in gnathostomes (jawed vertebrates). MHC loci encode cell-surface proteins (transmembrane receptors) that recognize and bind pathogens to initiate an adaptive immune response (Erickson 1987; Trowsdale 1995). The MHC is of high interest to behavioural and evolutionary biologists as evidence of both additive and non-additive effects on fitness and mate choice have been recorded at these loci. The gene products of the MHC can be classified into two categories: class I, whose receptors primarily deal with intracellular pathogens such as viruses, and class II, whose receptors primarily deal with extracellular pathogens, such as bacteria.

1.2.1 *The MHC class I*

Class I receptors occur on the surface of all nucleated cells and function to display intracellular peptides to cytotoxic T lymphocytes (CTLs) (Hewitt 2003; Kulski et al. 2002). These peptides are fragments of cytosolic proteins that have been degraded by the proteosome and can sometimes include viral proteins following or during active infection (Hewitt 2003; Kulski et al. 2002). When an antigenic peptide is displayed, it triggers an immediate immune response where CTLs induce apoptosis in all cells presenting the same peptide to prevent the spread of infection (Hewitt 2003; Kulski et al. 2002). Class I receptors may also function as inhibitory ligands for natural killer cells (NKs) (Wang et al. 2008). A reduction in their numbers, such as what occurs during active Influenza A and B infection, may induce inappropriate apoptosis by NK cells (Wang et al. 2008). Due to the rapid detection and retaliation cycle between the receptors and their pathogens, MHCI loci are under intense pathogen-mediated selection (Doherty and Zinkernagel

1975; Jeffery and Bangham 2000; Bernatchez and Landry 2003). The results of this selection can be most clearly seen at exon 3, which encodes the peptide-binding region (PBR), and consequently, is the MHCI's most variable portion (Cloutier et al. 2011).

1.2.2 *The MHC class II*

Class II receptors occur only on the surface of antigen-presenting cells, such as dendritic or B cells, and function to display extracellular peptides to T helper cells (CD4+) (Cresswell 1994; Nelson and Fremont 1999). These peptides are fragments of proteins endocytosed by lysosomes following the phagocytosis of proteins from outside of the cell, which can include pathogenic proteins during active infection (Cresswell 1994; Nelson and Fremont 1999). The presentation of an antigenic peptide to a CD4+ cell may trigger several immune responses including inflammation, phagocyte recruitment, and/or B cell activation (Owen et al. 2013). Much like MHC class I, pathogens have also evolved ways to avoid class II receptor detection, such as during *Francisella tularensis* infection, which induces receptor degradation (Wilson et al. 2009). Interactions with this and other pathogens have driven host-pathogen arms races, the intense selective effects of which are particularly apparent at MHCII exon 2, the highly variable region encoding the class II PBR (Doherty and Zinkernagel 1975; Jeffery and Bangham 2000; Bernatchez and Landry 2003).

1.2.3 *MHC Diversity*

One of the most intriguing features of the MHC is its tremendous variability not only among species but also within species. MHC loci are among the most polymorphic in the entire vertebrate genome. For example, as of 2022, there were 24038 class I and 9182 class II alleles recorded at the DNA level in humans (Barker et al. 2023). This variation reflects both multiple alleles being maintained per locus, and a large number of MHC loci being formed through gene duplication (Sommer 2005).

The highly polymorphic nature of the MHC is crucial for host defense against the evolution of evasion or suppression mechanisms by pathogens (Janeway et al. 2001). In terms of population persistence, the more alleles there are present in a population, the lower the likelihood that a given pathogen has developed resistance against every single

allele. In terms of individual fitness, similarly, heterozygote advantage at MHC is widespread, presumably reflecting greater disease resistance in heterozygotes. However, this does not imply that individual fitness always increases with individual genetic diversity at MHC. For every additional MHC allele, a number of T-cells that recognize any self-peptides that bind to the new receptor must be removed in order to maintain self-tolerance (Janeway et al. 2001). Failure to maintain self-tolerance may result in a variety of autoimmune disorders, which are characterized by tissue and organ damage and have been associated with a significant reduction in fitness. For example, in a wild population of Soay sheep (*Ovis aries*) individuals with high levels of self-reacting antibodies show lower reproductive rates (Brinkworth and Barreiro 2014; Graham et al. 2010). Therefore, the optimal number of MHC alleles to balance the advantages of being able to recognize a wide array of pathogens and the detriments of a lower T cell repertoire or the fitness consequences of an autoimmune disorder if self-tolerance cannot be maintained (Janeway et al. 2001) is not necessarily the maximal number.

Much like the variation within the MHC itself, the mechanisms by which this diversity has been generated and maintained are extensive. Duplication events have been proposed as a main driver of diversity (Kondrashov and Koonin 2004; Magadum et al. 2013) as the duplication of a gene allows for the acquisition of new mutations that may lead to novel pathogen recognition without risking the loss of function of the original gene. The maintenance of this allelic diversity following duplication of loci is largely thought to reflect balancing selection, specifically heterozygote or rare allele advantage (Nei et al. 1997).

1.2.4 *Balancing selection at MHC*

Across most of the genome, positive selection, which increases the frequency of beneficial alleles, is much rarer than purifying selection, which reduces the frequency of harmful alleles (Cvijović et al. 2018). However, positive selection is relatively common within immune system genes (Shultz and Sackton 2019; Bernatchez and Landry 2003). Positive selection may involve either directional or balancing selection. However, while evidence exists of both operating at MHC, balancing selection has been found to be far more common (Bernatchez and Landry 2003). This is due in part to the ability of

balancing selection to actively maintain a large number of alleles within a population, which is crucial to enable a sufficient breadth of pathogen recognition. Balancing selection may take the form of heterozygote or rare allele advantage, though the former seems to be favoured as a means of explaining high levels of MHC polymorphism (Penn et al. 2002; Kubinak et al. 2012). Examples of heterozygote advantage, where fitness in the form of enhanced disease resistance increases linearly with individual MHC genetic diversity, has been reported in numerous species (e.g., mice, Penn et al. 2002; McClelland et al. 2003; Atlantic salmon *Salmo salar*, Evans and Neff 2009; and water voles *Arvicola terrestris*, Oliver et al. 2009). However, other cases where an intermediate level of MHC diversity is associated with better fitness have been found, such as in three-spined sticklebacks (Eizaguirre et al. 2009). Rare allele advantage, otherwise known as negative frequency-dependent selection, occurs when there is a negative relationship between the relative fitness of an allele and its prevalence within a population. Rare allele advantage at immune genes occurs because in evolutionary arms races, pathogens often evolve resistance to a host population's most common immune alleles. An example of this was seen in a wild population of Soay sheep where individuals with the rarest MHC haplotype had the highest lifetime breeding success (Huang et al. 2022). Additionally, the frequency of this haplotype over the study period increased significantly more than what would be expected under neutral evolution (Huang et al. 2022).

1.2.5 *MHC-mediated mate choice*

MHC loci, given their importance to individual fitness, represent a compelling avenue to investigate mating preferences. Indeed, mate selection based on MHC genotype has been well documented for both class I and class II loci. In some cases, individuals seem to prefer mates who are themselves optimally diverse, presumably because such individuals provide better direct benefits. For example, male three-spined sticklebacks with optimal MHC diversity were found to make the best nests for rearing young (Jäger et al. 2007).

MHC-mediated mating is most frequently examined in the context of indirect (generally compatible-gene) benefits. As reviewed above, disease resistance often increases with individual diversity at MHC, leading to the general expectation that organisms should mate disassortatively with respect to MHC in order to maximize offspring fitness through

heterozygote advantage. Evidence of this has been established for many mammals (e.g., bank voles *Clethrionomys glareolus*, Radwan et al. 2008), fish (e.g., Atlantic salmon, Consuegra and Garcia de Leaniz 2008), and birds (e.g., savannah sparrows *Passerculus sandwichensis*, Freeman-Gallant et al. 2003). As disassortative mating increases the individual genetic diversity of offspring, it is often expected to be commonplace to further enhance the diversifying effects of pathogen-mediated balancing natural selection.

However, as stated earlier, the optimal level of MHC diversity is not necessarily the maximal level. For example, in humans, a highly polymorphic MHC genotype has been linked to the development of several autoimmune disorders (Ishigaki et al. 2022). This may explain why assortative mating at MHC loci has also been documented for many species (e.g., European badgers *Meles meles*, Sin et al. 2015*;* three-spined sticklebacks, Demetra et al. 2017*;* and tiger salamanders *Ambystoma tigrine*, Bos et al. 2009). Assortative mating may also serve as a way to prevent exposure to novel pathogens that may be carried by a MHC dissimilar mate (Lewis 1998). Finally, there exist species that mate randomly with respect to MHC genotype (e.g., brown bears *Ursus arctos*, Kuduk et al. 2014; great tits *Parus major*, Sepil et al. 2015; and common yellowthroats *Geothlypis trichas*, Bollmer et al. 2010). These cases are believed to be the result of these species being unable to discern MHC genotype or possessing more critical concerns when selecting a mate.

Considering that organisms are not born equipped with the ability to sequence DNA, evidence of preferences for particular MHC genotypes, particularly for compatibility, must imply the existence of another method of discernment. For many species, this has been identified as occurring through olfactory cues, with extensive documentation in mammals and fish (Milinski 2022). Studies of MHC-based mate selection in mice identified both the vomeronasal organ (VNO), a sensory organ located in the soft tissue of the nasal septum, and the main olfactory epithelia (MOE) as responsive to MHCI peptide-ligands (Leinders-Zufall et al. 2004; Spehr et al. 2006). Although not all vertebrates possess a functioning or present VNO, all have at least some form of MOE (Muller-Schwarze 2006), positioning olfactory communication as the leading hypothesis for MHC genotype assessment.

Other potential mechanisms for MHC genotype assessment have been proposed as well. In birds, ornamentation and overall body condition have been proposed as a means by which MHC genotype may be communicated. For example, greater body size in great snipes (*Gallinago media*, Ekblom et al. 2004), and longer spur length in ring-necked pheasants (*Phasianus colchius*, Von Schantz et al. 1997) correlate both with female preference and specific MHC genotype. Additionally, for passerine birds, birdsong has been suggested as another possible mechanism. Birdsong varies geographically (Podos and Warren 2007) and is a trait for which females exhibit behavioural preference (Collins 2004). As females across passerine species often prefer songs of local origin, birdsong has been proposed as a way for birds to identify mates with locally-good MHC genotypes (Stewart and MacDougall-Shackleton 2008), or to avoid inbreeding (Grant and Grant 1996). Overall, however, olfactory assessment remains the leading hypothesis for how MHC-compatible mates could be identified.

1.3 *Study species*

1.3.1 *Song sparrows (Melospiza melodia)*

Melospiza melodia, or song sparrows, are a medium-sized sparrow native to North America. Northern populations, such as those inhabiting Canada, typically migrate to the southern United States during the winter and return for the breeding season (Bent and Austin 1968). They are one of the first passerine species to arrive back in the spring in Ontario and have high adult site fidelity (Bent and Austin 1968), making them ideal for long-term studies and those that require capture. Song sparrows are one of the most morphologically diverse songbirds in North America, both in physical conformation and ecosystem occupation (Pruett and Winker 2010; Greenberg et al. 2012; Arcese et al. 2002) and are extensively used as a model species in field studies on behaviour and ecology.

Mate choice is likely tightly linked to fitness for both males and females in this species as both sexes provide direct benefits that are critical for offspring survival. These benefits include nest construction (females), egg laying and incubation (females), territory defense (males), provisioning and defending nestlings (both sexes) (Arcese et al. 2002). Captive

studies of breeding-condition song sparrows exposed to odour cues of opposite-sex conspecifics revealed that males, as well as females, prefer particular characteristics (Grieves et al., 2019), suggesting that both sexes likely exercise mate choice.

1.3.2 *Passerine MHC*

Across taxa, the MHC of birds, and more specifically that of passerines (perching birds, order Passeriformes) is distinct in both its evolutionary history and genomic composition. The avian MHC is characterized by recent duplication and pseudogene formation, in contrast to mammalian loci which are thought to be ancient in origin and maintained independently (Hess and Edwards 2002). Within birds, non-passerines generally exhibit stronger selection at MHCII loci, while passerines show stronger selective forces at MHCI (Minias et al. 2018). The chicken (*Gallus gallus*) was the first bird for which the genomic locations of MHC were known, posessing only two MHC regions (MHC-B and MHC-Y) located closely together on the q-arm of its 16th chromosome (Miller and Taylor 2016) with low rates of recombination between them. While our understanding of passerine MHC genomic locations is still in its infancy, genetic mapping in the zebra finch (*Taeniopygia guttata*) revealed both MHCI and MHCII loci located on chromosome 16, but a considerable distance and a high rate of recombination between them (Ekblom et al. 2011).

1.3.3 *Indirect benefits and MHC genotype signaling*

As with many other species, while the provision of direct benefits is important in song sparrows, it is not the only point of concern during mate selection. Throughout their lives, song sparrows come into contact with a wide array of pathogens, including haemosporidian parasites of the genus *Plasmodium* (associated with avian malaria), which can have major consequences for fitness and survival (Dunn et al. 2021). Song sparrows with more MHCI alleles are more resistant to infection following experimental exposure to *Plasmodium*, suggesting an important role for compatible gene effects on fitness. However, this pattern was not observed for MHCII alleles (Slade et al. 2017a). Additionally, Slade et al. (2017a) found no evidence of population differences in MHCI allele frequencies across Ontario populations that would suggest a role for local

adaptation and locally-good gene effects. While no relationship was found between MHCII diversity (individual allele count) and *Plasmodium* resistance (Slade et al. 2017a), song complexity (a predictor of lifetime reproductive success in song sparrows, Reid et al. 2004; Searcy 1984) varied nonlinearly as a function of MHCII diversity (Slade et al. 2017b). Song sparrows with intermediate numbers of MHCII alleles have the largest song repertoires and presumably the highest fitness (Slade et al. 2017b).

The discrepancy between measures of fitness associating with maximal diversity of MHCI alleles but intermediate diversity of MHCII alleles may reflect differences in the genomic composition of the two loci. In passerines, the MHCII is far more duplicated and polygenic than MHCI (Kaufman 2022). In our study population of song sparrows currently 192 MHCII alleles have been described (10-26 per individual) as opposed to 27 MHCI alleles (1-4 per individual) (Slade et al. 2017a,b). Thus, song sparrows may be at a higher risk of developing autoimmune disorders or experiencing dilution effects (too few copies of the most protective alleles) at MHCII relative to MHCI. If so, the difference between MHCI and II in allelic diversity and number of loci may explain these different relationships between individual genetic diversity and measures of fitness.

While historically passerines, including song sparrows, were thought to be anosmic (i.e. lack the ability to smell) (Bang and Cobb 1968), a growing body of research has shown that passerines possess a robust olfactory system and use scent in a variety of behavioural contexts, including reproduction (Caro et al. 2015). Captive song sparrows exhibit behavioural preferences for body odour from MHCII dissimilar conspecifics of the opposite sex, implying a role for odour in MHCII genotype signaling (Grieves et al. 2019). Avian body odour is largely a result of preen oil (Grieves et al. 2022), a waxy but partially volatile substance secreted from the preen gland and spread throughout the feathers by grooming. Pairwise similarity in preen oil chemical profiles is positively related to pairwise similarity at MHCII in song sparrows (Slade et al. 2016, Grieves et al. 2019). MHCII loci are thought to influence the chemical profile of preen oil both indirectly (through mediating the composition of the microbiome inhabiting the preen gland, including odour-producing and -altering microbes that in turn affect how the

volatile compounds in preen oil are processed and expressed), and directly, through the release of MHCII ligand-peptides (Archie and Theis 2011, Grieves et al. 2019).

1.3.4 *MHCII molecular evolution and evidence of historic balancing selection in song sparrows*

Song sparrows show molecular patterns at MHCII exon 2 that are consistent with the effects of pathogen-mediated balancing selection (Slade et al. 2019). Specifically, the rates of nonsynonymous to synonymous sequence variations are elevated in specific codons, and trans-species polymorphisms occur such that some song sparrow alleles are more similar to heterospecific sequences than to conspecific sequences (Slade et al. 2019). Sexual selection theory predicts that balancing natural selection may be reinforced by disassortative mating, and indeed, disassortative preferences at MHCII were found in odour-based preference experiments in song sparrows from western Ontario (Grieves et al. 2019). In contrast, however, free-living song sparrows in eastern Ontario show patterns of assortative mating such that socially mated pairs are more similar, not less similar, than expected under random mating at MHCII (Slade et al. 2019). Slade et al. (2019) speculated that the eastern Ontario population may harbour such high MHCII diversity that autoimmune disorders or dilution effects represent important costs of excessive diversity, suggesting that assortative mating may be a way to counteract the diversifying effects of balancing selection and produce offspring with optimal levels of individual variation at MHCII (Slade et al. 2019). Such costs are likely much less important at the less duplicated, less diverse MHCI.

1.4 *Objectives and hypothesis*

I sought to characterize patterns of selection and test for non-random patterns of mating at exon 3 of the MHCI (roughly corresponding to its major antigen binding site, or PBR (Trowsdale 1995), in the same free-living population of song sparrows studied by Slade and colleagues (2019). I used sequence information from song sparrows and other passerines to evaluate the occurrence of past balancing selection on the molecular level, reflected by (a) signatures of positive selection in song sparrows and (b) trans-species polymorphisms (TSPs) between song sparrows and other species; TSPs can indicate

ancient polymorphisms that arose prior to speciation events being maintained by balancing selection (Klein et al. 1998). I used behavioural data identifying socially mated pairs together with individual sequence information to determine if mated pairs were more or less similar at MHCI exon 3 than what would be expected under random mating.

To assess historical balancing selection, I tested for signatures of positive selection by evaluating ω, defined as the ratio of nonsynonymous substitutions per nonsynonymous site (dN) to synonymous substitutions per synonymous site (dS) within specific codons. A high dN to dS ratio (ω >1) indicates an increase in the frequency of beneficial alleles, which can either be the result of an increase in the frequency of rare alleles/increasing population differentiation (balancing selection) or represent a snapshot of an allele sweeping to fixation (directional selection) (Kimura 1977; Miyata and Yasunaga 1980). While directional selection has been found to occur (Hughes et al. 1996; Ou et al. 1998; Cohen 2002), for genes with extensive variation, like MHC, positive selection almost always represents balancing selection (Bernatchez and Landry 2003). Alternatively, a dN to dS ratio that is equal to or below 1 indicates neutral or purifying (the reduction in the frequency of harmful alleles) selection, respectively.

I tested for evidence of TSPs, characterized by alleles within my population being more similar to heterospecific alleles than to other song sparrow alleles at these loci, by evaluating sequence similarity between song sparrows and other passerines through phylogenetic analysis. To distinguish between TSPs maintained through ancient balancing selection (pre-dating speciation events; similarity by descent) versus different species having similar alleles because of evolutionary convergence, I constructed separate phylogenetic trees modeling positive or neutral selection following Mancilla-Morales et al. (2022). Similarity by descent can be characterized by allele sharing between phylogenetic families at neutral sites, indicating a retained ancestral signature, while convergence shows a stronger signal at positively selected sites, indicating similar selection pressures (Klein et al. 1998). While TSPs may be expected to appear under neutral evolution between recently diverged species, TSPs present after longer periods of time (i.e the millions of years that may divide species of the same family) can only arise through active maintenance by balancing selection (Klein et al. 1998).

I tested for nonrandom mating by quantifying MHCI similarity between members of all identified socially mated pairs. I determined the observed genetic distance between their allelic repertoires at exon 3 (PBR) and compared this to the expected genetic distance that would result from random mating, as determined by a Monte Carlo simulation. Nonrandom mating can be inferred by the average genetic distance between observed social pairs being lower or higher than that determined by the simulation, suggesting assortative or disassortative pairing, respectively. Assortative mating would be consistent with findings from free-living birds in the study population at MHCII, implying that sexual selection may work to counter the diversifying effects of balancing selection (Slade et al. 2019). In contrast, disassortative mating would imply an enhancing diversifying effect in tandem with the balancing natural selection that is likely to occur at MHCI through heterozygote advantage in malaria resistance (Slade et al. 2016). Alternatively, if the average genetic distance between the social pairs is found not to be significantly different from the simulation, then mating is random, and sexual selection is neither opposing nor enhancing effects of natural selection on the maintenance of diversity at MHCI.

I hypothesized that in my study population, MHCI exon 3 experiences strong balancing selection, further amplified by nonrandom mating. I predicted evidence of positive selection through elevated dN/dS ratios within specific codons, and evidence of TSPs by alleles displaying reciprocal monophyly rather than convergence, as observed in studies of MHC in song sparrows and other taxa (Slade et al. 2019, Jamie and Meier 2020). I predicted that unlike patterns at MHCII in my study population (Slade et al. 2019), I would find evidence of disassortative mating in song sparrows at MHCI, because *Plasmodium* resistance (a proxy for fitness) increases linearly with MHCI diversity (Slade et al. 2017a), and because lower numbers of loci at MHCI imply a lower potential for negative effects resulting from excessive allelic diversity than at MHCII.

2 *Methods*

2.1 *Study population and sample collection*

I used genetic and behavioural information collected from a population of song sparrows breeding near Newboro, Ontario from April to May in the years 2007, 2014, 2015, and 2022. In total, genetic samples from 218 individuals including 42 breeding pairs were available (Table 2.1.1). Previous work genotyping MHCI in 36 individuals from this and another Ontario population characterized 27 unique MHCI sequences with 1-4 alleles observed per individual (Slade et al. 2017a) (Genbank accession numbers KX264149-KX264175).

During April 12th – May 11th of 2022, I worked with my colleagues to capture adult song sparrows in seed-baited potter traps. Newly captured individuals were outfitted with a Canadian Wildlife Service (CWS) uniquely numbered aluminum leg band, and a unique combination of three colored plastic leg bands to enable field identification. A preliminary estimate of sex was determined through the presence (male) or absence (female) of a cloacal protuberance and a measurement of un-flattened wing length. A small volume of whole blood (~40μL) was collected from each individual through brachial venipuncture and blotted onto high wet-strength filter paper for genetic analysis (Fig. 2.1.1).

I confirmed field estimates of sex through PCR amplification of the CHD-Z (~350bp) and CHD-W (~450bp) regions of the sex chromosomes using the primers P2 (forward) and P8 (reverse) (Griffiths et al. 1998). PCRs were conducted in a total volume of 10μl, including 1.0μl of 5X colourless GoTaq® FlexiBuffer (Promega), 0.2μM of each primer, and ~20ng of template DNA. Cycling conditions were 1 min at 94.0°C; 39 cycles of 30s at 94°C, 45s at 48.0°C, and 45s at 72°C; and a final extension step of 10 min at 72°C. I ran 5μl of each PCR product on a 2% agarose gel run at 96V for 1hr, using 6X loading dye, stained with RedSafe and exposed to UV light for visualization. Males (ZZ) show one band and females (ZW) show two bands. Our initial field estimates of sex matched PCR-confirmed sex 83% of the time.

In 2007, members of the lab identified socially mated pairs through observations of behaviours from birds with known nests. Researchers observed colour-banded song sparrows through binoculars, noting interactions such as co-foraging, close following, copulations and solicitation displays, and provisioning the same nest (Potvin and MacDougall-Shackleton 2009). In 2014 and 2015, my colleagues identified socially mated pairs mainly through capture records that met one or both of the following criteria: (a) two individuals of the opposite sex were trapped at the same time, or (b) two individuals were captured in the same trap within 48 hr of each other, and no other song sparrows had been trapped at the same location for the entire duration of the field season (Slade et al. 2019). Behavioural observations as described above were used as supplemental information. In 2022, I identified socially mated pairs initially by using the same capture criteria used in 2014 and 2015, and then confirmed their status by observing, using binoculars and color bands, whether both birds still inhabited the same territory later in the season. In total, we identified 53 social pairs across the four years. However, only 42 pairs had DNA available for analysis (Table 2.1.1).

All animal work was approved by the University of Western Ontario Animal Use Subcommittee (protocol #2020-054 to EAM-S) and conducted under the required federal permit.

Table 2.1.1-1 Numbers of individual song sparrows genotyped, and number of socially mated pairs identified and genotyped, in each study year.

Year	# individuals genotyped	# pairs genotyped
2007	23	9
2014	63	9
2015	72	8
2022	60	16
Total	218	42

2.2 *DNA amplification and sequencing*

I extracted DNA from whole blood using 400µL of cell lysis buffer (100 mM Tris pH 8.0, 10 mM EDTA pH 8.0, 1% SDS) and 3µL of 20mg/mL proteinase K. Blood blots were cut up into small squares, added to the buffer/proteinase K solution, and digested at 60^{o}C for two hours, inverting tubes every 15 min. Proteins were precipitated out of samples using 200 µL of 7.5 M ammonium acetate, spun down in a tabletop centrifuge and discarded. DNA was precipitated from the supernatant using 900 µL of isopropanol, centrifuged into a pellet, then rinsed with 200 µL of 70% ethanol. I resuspended each DNA pellet in 1X TE buffer, estimated concentration using Nanodrop Spectrophotometry (ThermoScientific) and diluted DNA to a working concentration of ~20ng using 1x TE buffer (10 mM Tris-8, 1 mM EDTA).

I used polymerase chain reaction (PCR) to amplify a 213bp region within MHCI exon 3, the expected length of which in passerines is approximately 276bp (Minias *et al*. 2018). I used primers GCA21M (forward) and fA23M (reverse) (Loiseau et al. 2009) known to work in song sparrows (Slade et al. 2017a) and designed from cDNA to preferentially amplify transcribed alleles (Loiseau et al. 2009). I used next generation sequencing on an Illumina MiSeq platform, so that each forward and reverse primer contained an Illumina adaptor sequence (32bp), four 'wobble' bases, and a unique barcode sequence (12bp). In total, I had 16 uniquely tagged forward, and 16 uniquely tagged reverse primers, sufficient to generate 256 individual combinations.

PCRs were conducted in a total volume of 30µl, including 3.6µl of GoTaq® 10X FlexiBuffer (Promega), 0.2µM of each primer, 1mM of additional $MgCl_2$, 0.2U GoTaq polymerase (Promega), 0.2mM of dNTPs, and ~20ng of template DNA. Cycling conditions were 2 min at 94.0°C; 32 cycles of 30s at 94°C, 30s at 57.0°C, and 30s at 72°C; and a final extension step of 10 min at 72°C. I confirmed amplification by running 5µl of each PCR product on a 2% agarose gel stained with RedSafe and run at 96V for 1hr, using UV light for visualization.

I assigned each individual a unique combination of barcoded primers for identification. I sent amplified products to the London Regional Genomics Centre for high throughput

sequencing on an Illumina MiSeq flow cell. Following sequencing, I used a custom perl script (Gloor et al. 2010) to demultiplex the sequence reads and assign them back to individuals. I filtered the demultiplexed reads (i.e removed presumed sequencing errors and/or chimeras) using the DADA2 package (Callahan et al. 2016) in R 4.1.3 (R Core Team, 2022) and a previously determined error rate of 1% as described in detail in Slade et al. (2017b). Essentially, Slade et al. (2017b) used bacterial cloning (Promega pGEM-T Easy Vector System) to generate colonies that each contained a single allele. Colonies were PCR-amplified then sequenced on an Illumina Miseq flow cell to determine the frequency of secondary reads (i.e., presumed PCR or sequencing errors) to establish the error rate associated with PCR and sequencing. Slade et al. (2017b) observed a secondary sequence incidence of under 1% and conservatively used a value of 1% to filter out rare sequences that could represent PCR errors; I used this same value to filter my reads. I exported the sequences generated from the reads to Geneious Prime v. 2021.2.2 (Biomatter Ltd., Auckland, New Zealand) to align to other passerine MHCI sequences using BLAST (Altschul et al. 1990) and MUSCLE (Edgar 2004).

Slade et al. (2017a) found a maximum of four alleles per individual song sparrow at MHCI, implying at least two loci, but their sample size (36 individuals) was considerably lower than mine. To further investigate the number of loci at MHCI, I created a minimum spanning network of all unique sequences in the Population Analysis with Reticulate Trees software (popART) (Leigh and Bryant 2015). A minimum spanning network attempts to visualize the relationships among nodes (in this case the unique sequences) by creating the shortest possible paths between them (Bandelt et al. 1999). I re-sized the nodal points according to the number of individuals the respective sequence was retrieved from to qualitatively visualize the frequencies and distribution of alleles at the population level.

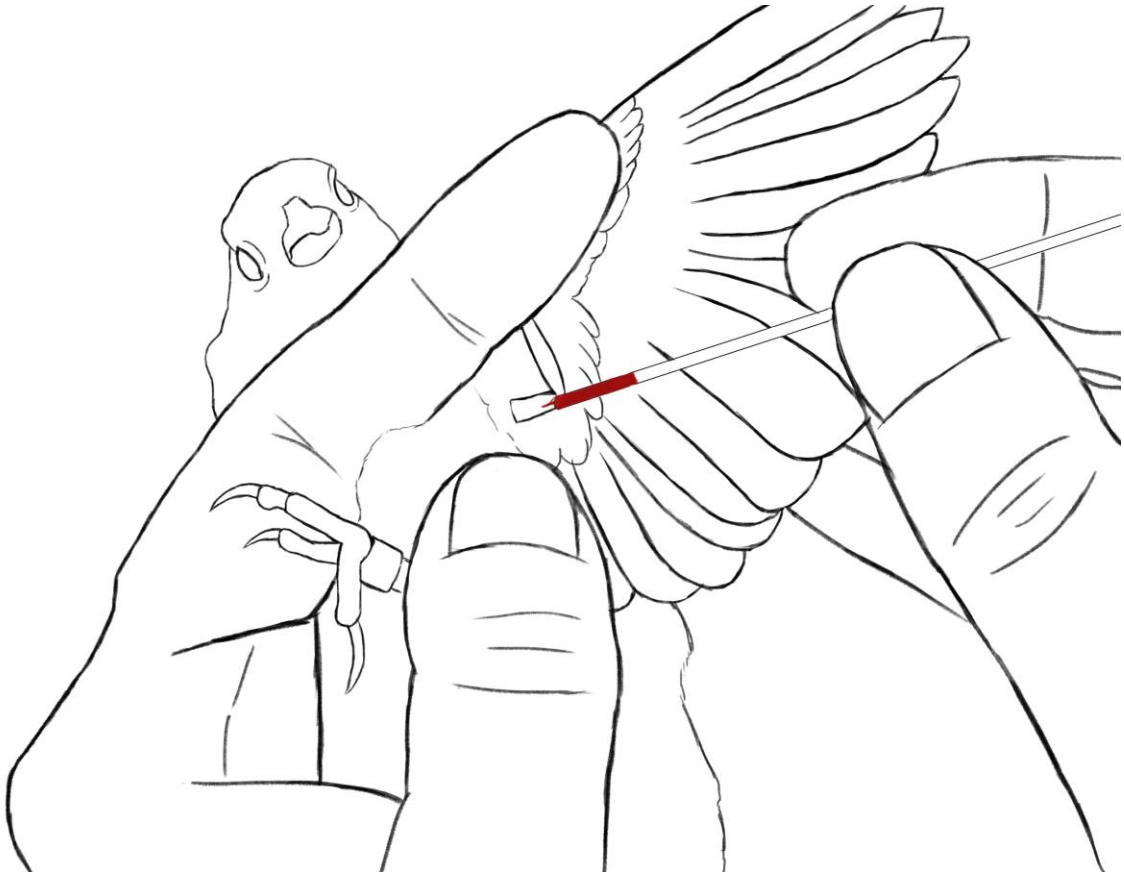

Figure 2.2.1 Blood collection via brachial venipuncture of an adult song sparrow

~40ul of blood is collected by capillary action in a heparinized tube from a small puncture in the brachial vein.

2.3 *Evaluating balancing selection*

2.3.1 *Contemporary balancing selection*

I inferred the effects of contemporary or recent balancing selection by assessing positive selection both sequence-wide and within specific codons using PAML-4.9j (Yang 2007) and the Datamonkey server (https://www.datamonkey.org; Weaver et al. 2018).

PAML is a package of programs that can be used for phylogenetic analyses of DNA and protein sequences using maximum likelihood methods (Yang 2007). I used the *CODEML* program within PAML-4.9j to estimate ω within my sequences to generate a series of candidate evolutionary models. PAML treats ω for any codon in a sequence as a random variable from a statistical distribution, allowing it to vary at each site (Yang 2007). PAML generates evolutionary models in pairs, the first being a null model that does not allow for positively selected codons ($\omega > 1$) and the second a more general model that does (Yang 2007). The explanatory power of the paired models relative to one another can be determined by the LRT statistic, defined as twice the difference in log-likelihood between the null and alternative hypotheses, $2\Delta\ell = 2(\ell 1 - \ell 0)$, where $\ell 0$ is the log-likelihood score for the null model, and $\ell 1$ is the log-likelihood under the alternative model (Álvarez-Carretero et al. 2023). The LRT statistic is compared against the $\chi 2$ distribution with the degree of freedom equal to the difference in the number of free parameters between the two models (Álvarez-Carretero et al. 2023). The paired models of interest to my study were M1a (neutral) and M2a (positive), and their beta distribution counterparts, M7 and M8, respectively (Yang 2007). Provided that M2a and M8 are found to have more explanatory power than their paired model, they can be further used to estimate which particular codons are under positive selection using a Bayes empirical Bayes (BEB) approach (Yang 2007).

I further tested which codons have undergone positive selection using three different approximation approaches on the Datamonkey server (Weaver et al. 2018). These approaches included Fixed Effects Likelihood (FEL), which uses a maximum likelihood model to infer dN/dS rates on a per-site basis (Kosakovsky-Pond and Frost 2005); Fast Unconstrained Bayesian Approximation (FUBAR), which uses a Bayesian approach to

infer dN/dS rates on a per-site basis (Murrell et al. 2013); and Mixed-Effects Maximum Likelihood (MEME) which uses a mixed-effects maximum likelihood approach to test individual sites for episodic or diversifying selection (Murrell et al. 2012). FEL and FUBAR assume that the selection pressure remains constant at all sites (Kosakovsky-Pond and Frost 2005; Murrell et al. 2013) while MEME infers two ω rate classes at each site with corresponding weights (Murrell et al. 2012).

For further confirmation of these methods, I compared all codons that I detected as having undergone positive selection to all positively selected codons identified across passerines by Minias *et al.*, (2018).

2.3.2 *Ancient balancing selection*

Traces of ancient balancing selection may be inferred from the presence of trans-species polymorphisms. To distinguish TSPs resulting from longstanding balancing selection from those generated by evolutionary convergence, I constructed two separate phylogenetic trees: one for sequences trimmed to include only codons determined to be under positive selection and one for sequences trimmed to include only codons determined to be under neutral selection (as explained above). Maximum likelihood trees were constructed using the ten most common alleles within my study population and a subset of passerine alleles identified by BLAST with at least >95% nucleotide sequence similarity. I originally identified 210 such heterospecific sequences, but to avoid sampling bias, I did not include more than 10 alleles from the same species and did not include more than one species from the same genus – resulting in a total of 96 sequences used. Decisions on which alleles and species to retain were made arbitrarily.

I identified positive and neutral codons using FEL (Kosakovsky-Pond and Frost 2005), FUBAR (Murrell et al. 2013), and MEME (Murrell et al. 2012) on the Datamonkey server. I constructed trees in PhyML 3.0 (Guindon et al. 2010), based on sequence alignments of only neutral or positively selected sites from 96 passerine species, using the automatic selection of substitution model feature (Lefort et al. 2017) and aLRT SH-like branch supports (Anisimova and Gascuel 2006). I used an MHCI sequence from *Falco*

tinnunculus (common kestrel) (Genbank accession number: EU120696) as an outgroup to build the phylogenies.

2.4 *Evaluating non-random mating*

To evaluate MHCI-based mate selection, I compared the average genetic distance between observed social pairs to that expected under random mating. I built a maximum-likelihood protein phylogeny of all recovered alleles using the Whelan and Goldman substitution model (Whelan and Goldman 2001) with five discrete gamma categories in Geneious Prime v. 2021.2.2 (Biomatter Ltd., Auckland, New Zealand). I used Whelan and Goldman substitution to take into consideration that some non-identical allelic pairs may be more similar than others due to physical attributes such as polarity, which may influence the suite of pathogens to which the encoded receptor can bind.

Using the WAG phylogeny, I calculated the pairwise amino acid distances between the allelic repertoires of all potential pairs (male-female combinations of individuals caught in the same year) and all observed social pairs. Specifically, I calculated UniFrac distances using the package GUniFrac (Chen et al. 2012) implemented in R 4.2.1 (R Core Team, 2022). UniFrac distances range between zero to one, zero indicating that two individuals have identical sets of alleles, and one indicating that their allelic repertoires are completely different and have zero branch sharing. Because my sequence data could only confirm whether an individual had a particular allele (but not how many copies of that allele), I used unweighted Unifrac distances, which do not consider abundance. Finally, I generated an expected distribution of genetic distances under random mating using Monte Carlo Simulation (Bryan 2020) with 10,000 permutations in RStudio 2022.12.0.353 (Posit Team, 2022) running base R (R Core Team 2022) and compared this to the distribution of observed distances between socially mated pairs.

3 *Results*

3.1 *Variation at MHCI*

Overall, 81% of reads were retained following the filtering process. The average (± SEM) final number of reads per individual was 10599 ± 691. I found no correlation between the number of DNA alleles identified in an individual and their number of retained reads (r_{216} = -0.01, p = 0.56), suggesting that individual variation in sequencing depth was not a concern.

Sequences were 174-177bp long, reflecting an in-frame (3 bp) deletion at codon 49 occurring in 11 sequences (Fig. 3.1.1). Nonsynonymous variants were found at 28 codon sites and synonymous variants at 31 sites (Fig. 3.1.2).

I identified 26 unique DNA sequences (different at one or more base pairs) from the 218 individuals genotyped, with an average (± SEM) of 2.84 ± 0.06 alleles per individual. I did not find that allele numbers varied significantly among the years (ANOVA: $F_{4,218}$ = 1.81, p = 0.15). The number of alleles found in each individual ranged from 1 to 5, indicating a minimum of three loci. The minimum spanning network revealed two highly prevalent alleles (MHCI-1 and MHC-2), another less prevalent allele (MHCI-6), and many rarer alleles joined by one or more mutational differences (Fig 3.1.3).

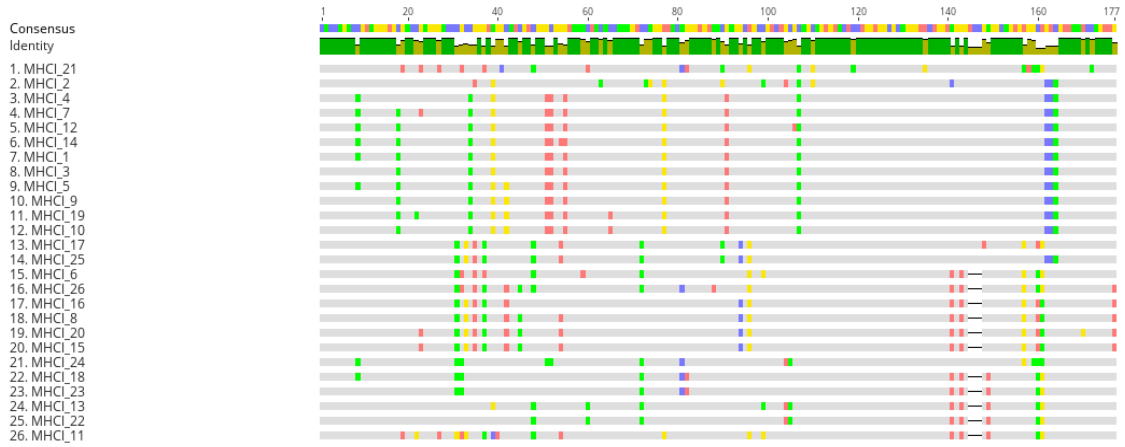

Figure 3.0.1.1 Nucleotide sequence alignment

Nucleotide sequence alignment of all 26 retained alleles. Sequences were 174-177bp in length reflecting a 3bp deletion at positions 146-148. Average pairwise sequence identity highlighted in green below the consensus sequence. (Geneious version 2022.0 created by Biomatters. Available from https://www.geneious.com)

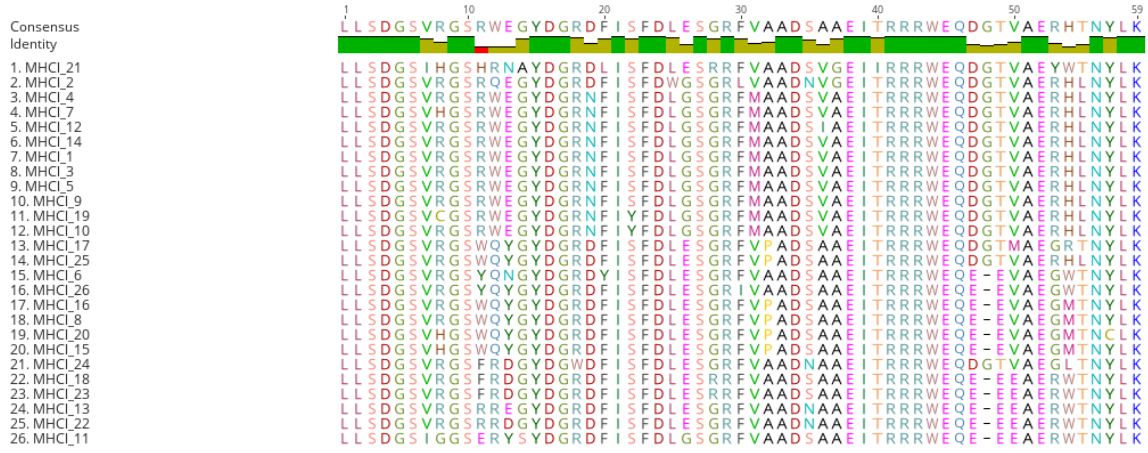

Figure 3.1.2 Amino acid sequence alignment

Amino acid sequence alignment of all 26 retained alleles. Sequences were 58-59 codons in length reflecting a deletion at the 49th position. Average pairwise sequence identity highlighted in green below the consensus sequence. (Geneious version 2022.0 created by Biomatters. Available from https://www.geneious.com)

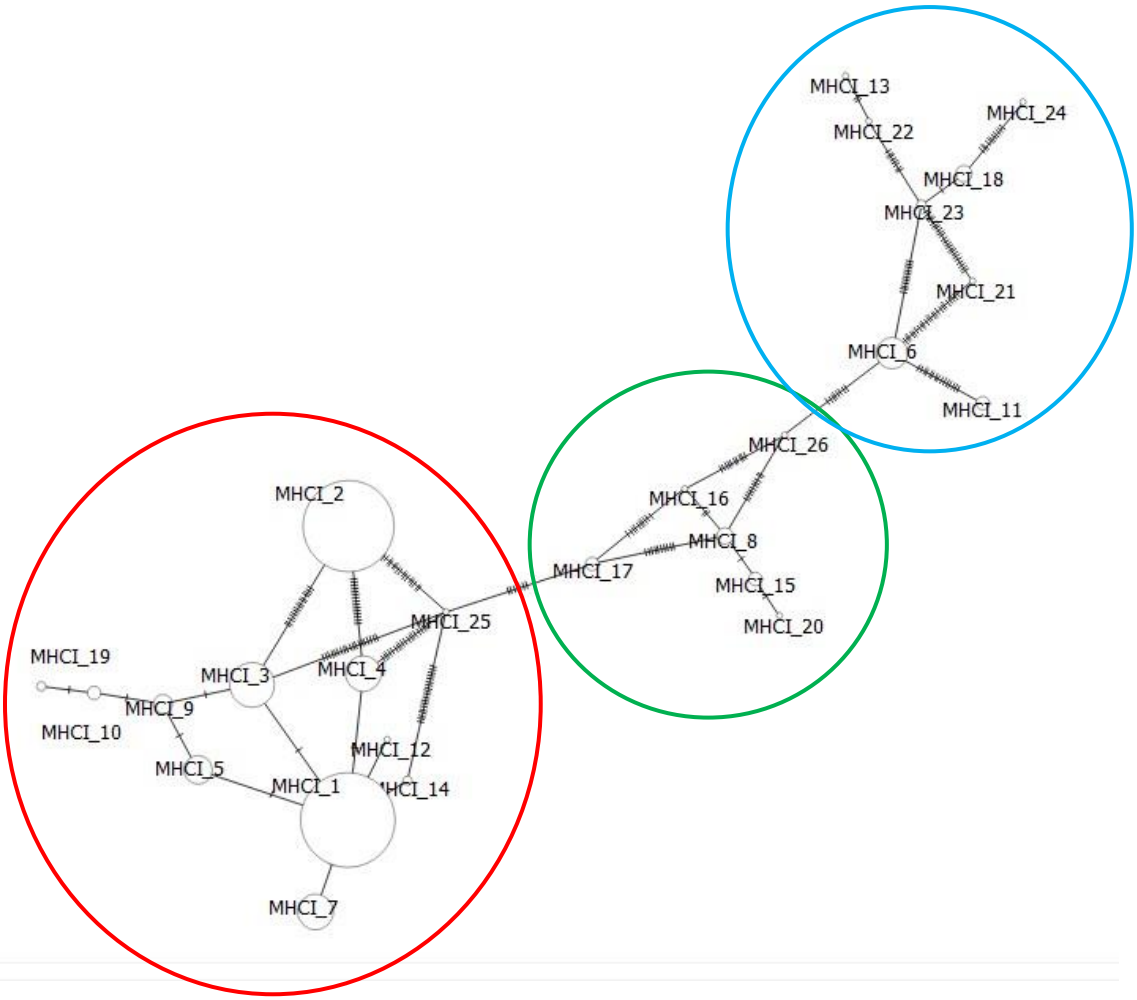

Figure 3.1.3 Minimum spanning network of recovered MHCI alleles

Minimum spanning network describing relationships between the 26 retained song sparrow alleles. Hatch marks indicate mutational distance. Nodal points have been resized according to the prevalence of the allele in the total population. Putative allelic clusters outlined in red, green, and blue circles.

3.2 *Selection at MHCI*

3.2.1 *Contemporary balancing selection*

I observed that the patterns of molecular evolution within my sequences were better explained by models that allowed for positive selection. In PAML, I found that model M2a fit the data better than its neutral counterpart M1 (ΔAIC = 50) (Table 3.2.1). I saw similar results between models M8 and M7 (ΔAIC = 50) (Table 3.2.1). Under model M2a, 0.77 sites were estimated to be under purifying selection, 0.22 under neutral drift, and 0.01 under positive selection (Table 3.2.1). Under model M8, 0.99 sites were estimated to be under purifying selection, and 0.01 under positive selection (Table 3.2.1). Both models (M2a and M8) identified the same 10 codons as having experienced positive selection, 4 of which had been previously identified by Minias *et al.* (2018) (Fig. 3.2.1). In Datamonkey, 3, 4, and 7 codons were identified as having undergone positive selection by FEL, FUBAR, and MEME, respectively (Fig. 3.2.1). Six of these codons had been implicated in the M2a and M8 models, and 7 had been previously identified by Minias *et al.* (2018) (Fig. 3.2.1). Overall, I consistently found 7 positively selected codons, with 6 matching previous passerine studies (Minias et al. 2018) (Fig. 3.2.1).

Table 3.2.1 Parameter estimates of PAML model simulations

Models	lnL	AIC	ΔAIC	Parameter estimates	Positively selected codons
Model 2a	-970	1952	---	p0 = 0.77, p1 = 0.22, p2 = 0.01	8, 11, 12, 13, 53, 54, 55
Model 8	-970	1952	0	p0 = 0.99, p1 = 0.01, p = 0.05, q = 0.02	8, 11, 12, 13, 53, 54, 55
Model 1	-997	2002	50	p0 = 0.82, p1 = 0.18	Not allowed
Model 7	-997	2002	50	*p* = 0.01, q = 0.02	Not allowed

Site models as estimated by CODEML in descending order from best to least explanatory power using ΔAIC values between paired models. Parameters include: p0 (proportion of sites under purifying selection), p1 (proportion of sites under neutral selection), p2 (proportion of sites under positive selection), p and q (beta distribution ω variables for sites within the range of $0 \leq \omega \leq 1$).

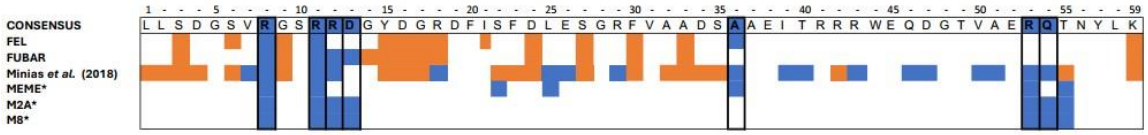

Figure 3.2.1 Consensus sequence of recovered MHCI alleles highlighting positively selected codons

Consensus sequence generated from all 26 identified song sparrow alleles along with approximation approaches describing positively and negatively selected sites. Sites highlighted in blue are those that the respective program identified as undergoing positive selection, and sites outlined in orange are those that were identified as undergoing purifying selection. Positively selected sites that were agreed upon by at least 3/6 test methods were highlighted. Results from the extensive passerine study conducted by Minias *et al*. (2018) were included as a point of comparison.

3.2.2 *Ancient balancing selection (trans-species polymorphisms)*

My 26 recovered alleles showed >95% nucleotide similarity with 210 other passerine MHCI sequences. My final list, filtered as described earlier to reduce sampling bias, contained sequences from 15 species within 8 different families. FEL, FUBAR, and MEME were used to determine positively and negatively selected codons from the alignment (Fig. 3.2.2). I considered a codon to be under positive or negative selection if the conclusion was agreed upon by at least two of these methods. In PhyML 3.0, the generalized time reversible model with 4 discrete gamma distribution categories (GTR+G) was identified as the best model for the positive codon tree (Table 3.2.2), and the Hasegawa-Kishino-Yano 1985 model with relaxed assumption of gamma distribution rates (HKY85+R) was identified for the neutral tree (Table 3.2.3). The GTR model assumes different rates of substitution between nucleotides and that nucleotides occur at different frequencies (Tavaré 1986), while the HKY85 model assumes that nucleotides occur at different frequencies and that transitions and transversions occur at different rates (Hasegawa et al. 1985). While both trees showed intermingling of alleles, the neutral codon tree exhibited better internal branch support (aRLT > 90) (Fig. 3.2.3) (Fig. 3.2.4).

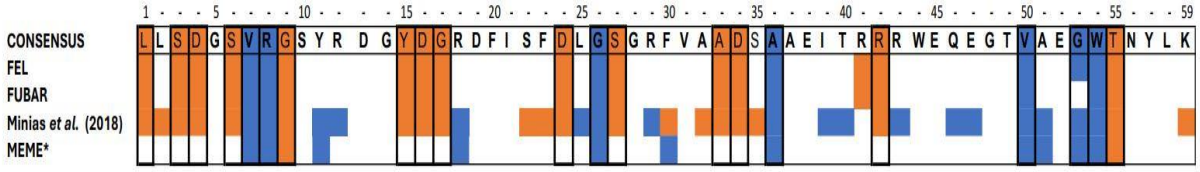

Figure 3.2.2 Passerine consensus sequence highlighting positively and negatively selected codons

Consensus sequence generated from all 107 alleles used to build maximum likelihood phylogenies along with approximation approaches describing positive and negatively selected sites. Sites outlined in blue are those that the respective program identified as undergoing positive selection, and sites outlined in orange are those that were identified as undergoing negative selection. Sites that were agreed upon by at least 2/3 test methods were highlighted. Results from the extensive passerine study conducted by Minias et al. (2018) were included as a point of comparison.

Table 3.2.2 **PhyML smart model selection (SMS) results for construction of the neutrally selected tree.** Parameters include: model (substitution matrix), site-to-site rate heterogeneity, K (number of free parameters), Llk (log likelihood), AIC (Akaike's information criterion value), BIC (Bayesian information criterion value). SMS selects the best model using AIC values.

Model	Site-to-site rate Heterogeneity	K	Llk	AIC	BIC
HKY85	+R	198	-1512	3419	3950
TN93	+R	199	-1511	3420	3954
GTR	+R	202	-1509	3422	3964
K80	+R	195	-1519	3427	3950
HKY85	+G	193	-1525	3436	3954
GTR	+G	197	-1521	3437	3965
HKY85	+G +I	194	-1525	3438	3959
GTR	+G +I	198	-1521	3439	3970
GTR	+I	197	-1571	3535	4064
HKY85	+I	193	-1578	3543	4060
GTR		196	-1600	3592	4117
HKY85		192	-1606	3596	4111

Table 3.2.3 **PhyML smart model selection (SMS) results for construction of the postiviely selected tree.** Parameters include: model (substitution matrix), site-to-site rate heterogeneity, K (number of free parameters), Llk (log likelihood), AIC (Akaike's information criterion value), BIC (Bayesian information criterion value). SMS selects the best model using AIC values.

Model	Site-to-site rate heterogeneity	K	Llk	AIC	BIC
GTR	+G	197	-306	1006	1212
GTR	+G +I	198	-306	1008	1215
TN93	+G	194	-312	1011	1214
GTR	+R	202	-306	1015	1226
GTR	+I	197	-321	1036	1242
GTR		196	-325	1041	1246

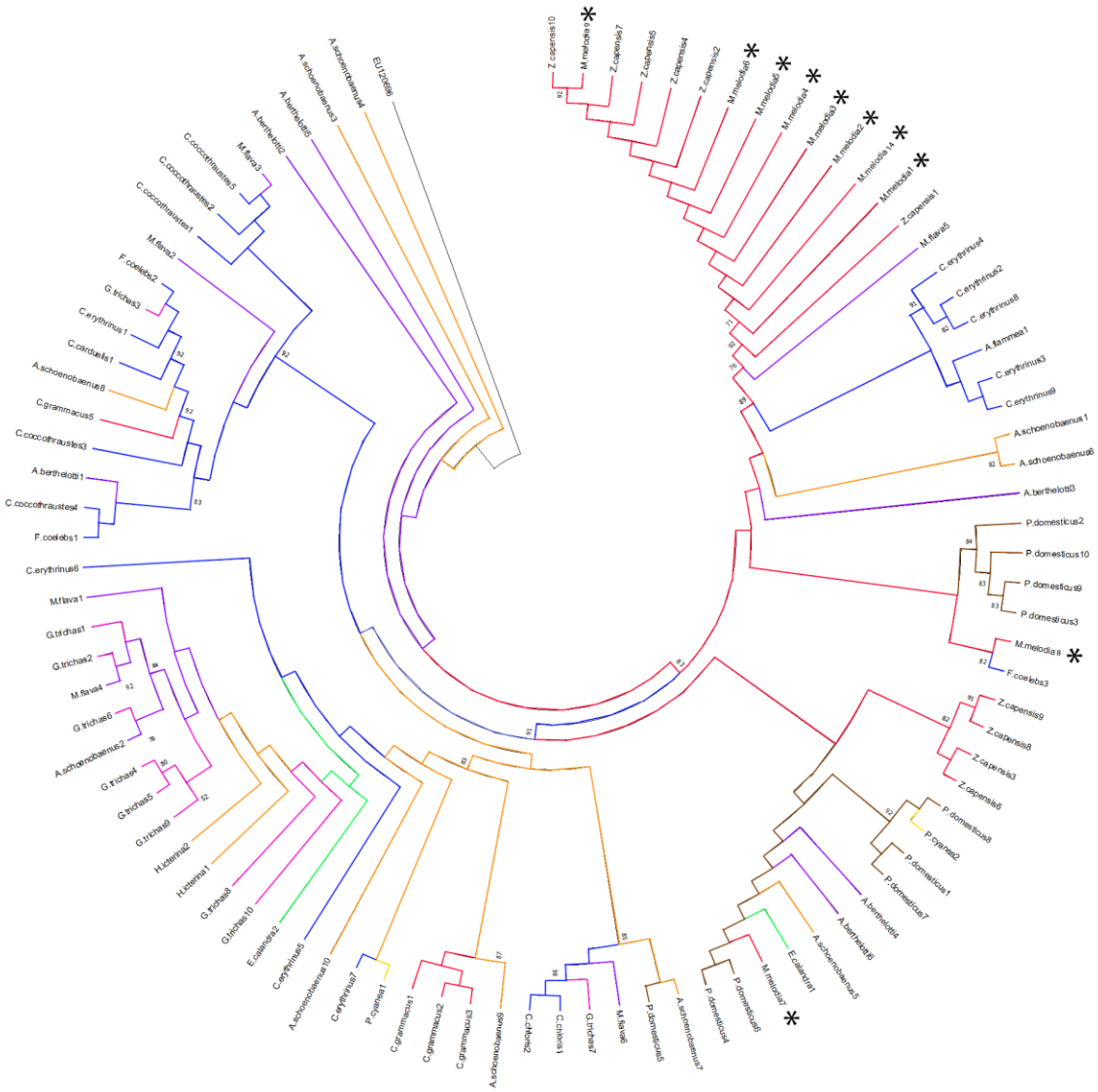

Figure 3.2.3 Phylogenetic reconstruction of positively selected codons

Maximum likelihood phylogeny of sites that were determined to be under positive selection, which consisted of 7 different codons. Branch support values >90 are not shown. The GTR+G model was used with aRLT-like branch support. Alleles from 107 individuals were included. Species within the same taxonomic family are highlighted in the same colour: *Passerellidae* (red, includes song sparrows), *Acrocephalidae* (orange), *Cardinalidae* (yellow), *Emberizidae* (green), *Fringillidae* (blue), *Motacillidae* (purple), *Parulidae* (pink), *Passeridae* (brown). Song sparrow alleles are identified by (*).

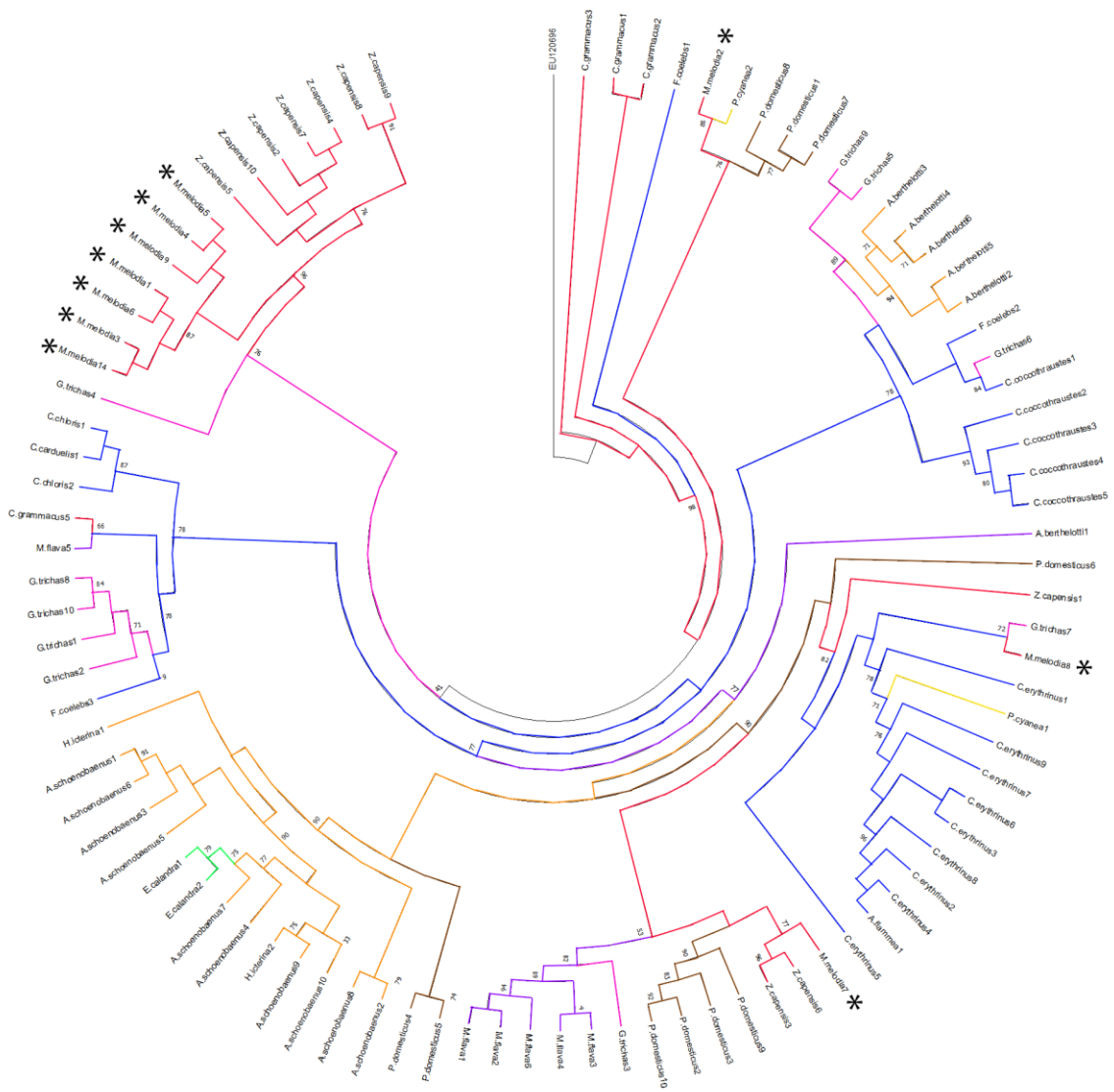

Figure 3.2.4 Phylogenetic reconstruction of neutrally selected codons

Maximum likelihood phylogeny of sites that were determined to be under neutral selection, which consisted of 36 different codons. Branch support values >90 are not shown. The HKY85+R model was used with aRLT-like branch support. Alleles from 107 individuals were included. Species within the same taxonomic family are highlighted in the same colour: *Passerellidae* (red, includes song sparrows), *Acrocephalidae* (orange), *Cardinalidae* (yellow), *Emberizidae* (green), *Fringillidae* (blue), *Motacillidae* (purple), *Parulidae* (pink), *Passeridae* (brown). Song sparrow alleles are identified by (*).

3.3 *Nonrandom Mating*

The average pairwise genetic distance for observed social pairs (mean ± SD = 0.34 ± 0.21) was not significantly different from that of the simulated randomized pairs (0.33 ± 0.20; two-tailed p = 0.76) (Fig. 3.3.1)(Fig. 3.3.2). Thus, I found that observed social pairs were neither more or less similar at MHCI exon 3 than expected under random mating.

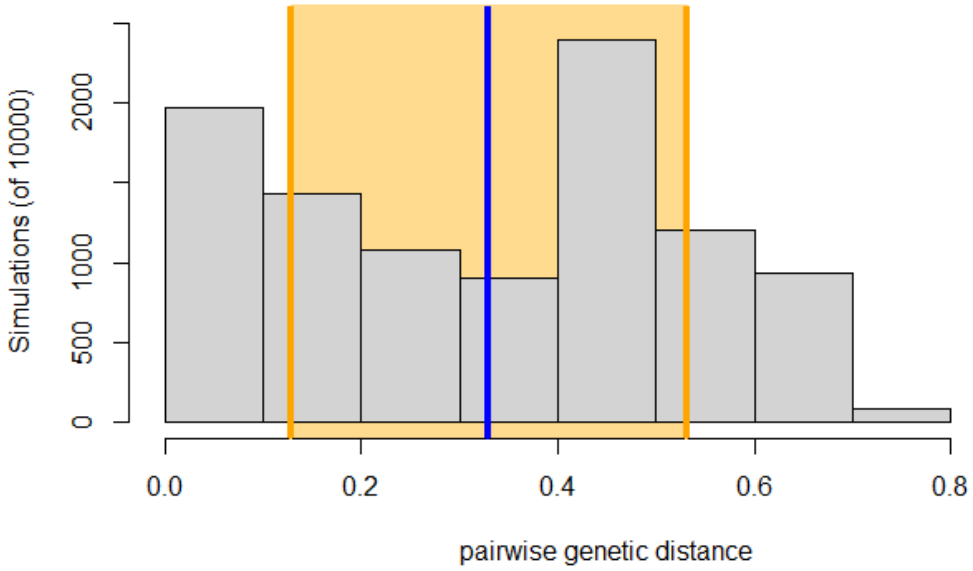

Figure 3.3.1 Monte Carlo frequency distribution of pairwise genetic distances at MHCI exon 3

Frequency distribution of pairwise genetic distances (unweighted UniFrac; gray bars) at exon 3 of class I MHC were generated by a Monte Carlo simulation of 10,000 randomized male-female pairings. The vertical blue line at 0.33 represents the average pairwise genetic distance under random mating. The vertical orange lines represent ±1 SD around this average.

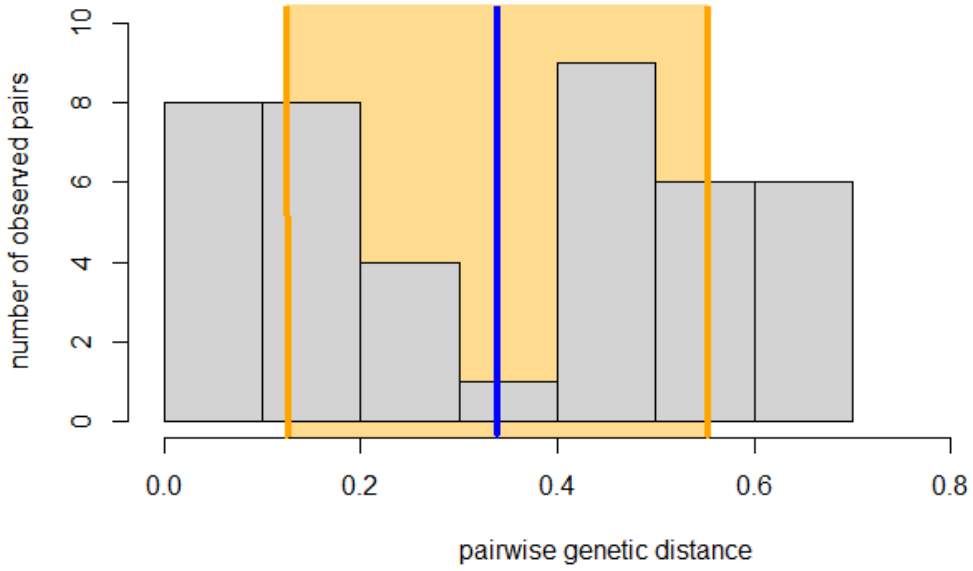

Figure 3.3.2 Observed pairs frequency distribution of pairwise genetic distances at MHCI exon 3

Frequency distribution of pairwise genetic distances (unweighted UniFrac; gray bars) at exon 3 of class I MHC of observed social pairs. The vertical blue line at 0.34 represents the average pairwise genetic distance found between observed social pairs. The vertical orange lines represent ±1 SD around this average.

4 *Discussion*

4.1 *Overview*

As predicted, I found evidence of ancient and contemporary balancing selection acting within my study population through the presence of positively selected codons and trans-species polymorphisms. Although I cannot definitively rule out directional selection, balancing selection is far more common across MHC loci (Bernatchez and Landry 2003). Within my sequences, 10 codons were implicated as having undergone positive selection, 6 of which had been previously referenced in other passerine studies (Minias et al. 2018). This may reflect the structural composition of the peptide binding groove and how it evolves in response to pathogen evasion (Furlong and Yang 2008). Furthermore, phylogenetic reconstruction showed clustering of neutrally selected codons between species, which possibly indicates that their similarity is due to descent from a common ancestor (Klein et al. 1998). Long-standing TSPs can only arise via balancing selection due to its ability to actively maintain multiple alleles in a population over long periods of time through negative frequency-dependent selection and/or heterozygote advantage (Klein et al. 1998).

Despite the above evidence for balancing selection at MHCI, presumably reflecting some degree of parasite-mediated heterozygote advantage, I found that the average genetic distance between observed social pairs was not significantly different from randomly generated pairs. This implies a system of random mating and that MHCI does not play a detectable role in song sparrow mate selection in the wild. While initially surprising, my findings are consistent with the literature. Random mating at MHCI has been exhibited in many species of passerines (Westerdahl 2004; Sepil et al. 2015; Wright et al. 2016; Stervander et al. 2020; Strandh et al. 2012; Luisa et al. 2022). In great reed warblers (*Acrocephalus arundinaceus*), great tits, Seychelles warblers (*Acrocephalus sechellensis*), Raso larks (*Alauda razae*), and blue petrels (*Halobaena caerulea*), mated pairs are neither more similar nor less similar than expected under random mating at MHCI (Westerdahl 2004; Sepil et al. 2015; Stervander et al. 2020; Strandh et al. 2012). Additionally, house sparrows (*Passer domesticus*) do not show any behavioural preference for MHCI genotype when presented with preen oil from potential mates

(Luisa et al. 2022). These findings could reflect a lack of compatible-genes effects of MHCI on fitness, or a general inability for passerines to detect MHCI genotype. MHCI mate choice in other taxa is thought to be facilitated by the vomeronasal organ (Singer et al. 1997), which has not been found to be present or functioning in birds.

4.2 *Evidence of balancing selection*

In agreement with my predictions, I found evidence of both ancient and contemporary balancing selection within my study population, potentially reflecting long-standing, and persistent evolutionary relationships with pathogens.

A high dN/dS ratio was found within several codon positions within my sequences – a molecular hallmark of Darwinian positive selection. While I cannot rule out the possibility that this may be the result of directional selection, the linear relationship between MHCI allele number and malarial resistance points towards balancing selection (Slade et al. 2017a) – which is universally more common across MHC loci (Bernatchez and Landry 2003). This indicates that exon 3 of MHCI, which roughly corresponds to the antigen binding region of the cell surface receptor, has been allowed to extensively undergo sequence changes that generate new protective alleles pathogens have not yet evolved evasion to. Passerines in general have stronger selective forces at MHCI than other avian species, possibly due to higher exposure rates to intracellular pathogens and differences in MHC architecture (Minias et al. 2018).

I identified several codons under either positive or purifying selection. While all examined codons were contained within the peptide binding groove, its structural conformation may explain these findings. Codons under positive selection may be those that bind directly to the antigen side chains – thus needing to be able to change over evolutionary time to recognize different antigenic peptides (Furlong and Yang 2008). Codons under purifying selection may be those involved with keeping the antigen in place, T-cell binding, or those required for structural integrity (Furlong and Yang 2008). Four of my positively selected sites match previous passerine pseudo-sequence models, which estimate the specific codons that facilitate peptide binding, and to the peptide-binding cleft in previous models of homology (Follin et al. 2013).

My phylogenetic trees, while showing allelic mixing between families for both positive and neutral codons, showed higher internal branch support for the neutrally selected positions. This is indicative of TSPs being maintained over long periods of evolutionary time, rather than arising by convergence (Klein et al. 1998). While TSPs may be expected to appear under neutral evolution between recently diverged species, TSPs present after longer periods of time (i.e. the millions of years that may divide species of the same family) can only be maintained by balancing selection (Klein et al. 1998). It should be noted that direct comparison of my two phylogenies may be an inappropriate measure given that the neutral phylogeny, having more codons, would inherently have more statistical power and thus greater ability to detect phylogenetic signal. However, the majority of MHC studies still suggest that ancient TSPs are the most likely explanation for shared allelic lineages among species (Klein, Sato, and Nikolaidis 2007). Additionally, both phylogenies showed allele sharing between song sparrows and the rufous-collared sparrow (*Zonotrichia capensis*), which inhabits a distinctly different geographical range in South America (Chapman 1940). Given that the distribution, prevalence, and diversity of avian haemosporidian parasites is known to vary with geographical regions (Valkiunas 2005), it is unlikely that the two species would experience similar enough parasite-mediated selection pressures to explain allele sharing by convergence (Lenz et al. 2013; Schwensow et al. 2010).

4.3 *Random mating at MHCI*

Contrary to my prediction of disassortative mating, I found evidence of random mating at MHCI in my study population of song sparrows. While disassortative mating at these loci would likely confer fitness benefits, considering the linear relationship between MHCI diversity (number of alleles) and disease resistance (Slade et al. 2017a), my result agrees with much of the literature on the subject. Across passerines, non-random mating has not been reported at MHCI (Westerdahl 2004; Sepil et al. 2015; Wright et al. 2016; Stervander et al. 2020; Strandh et al. 2012). This is in opposition to MHCII, at which song sparrows show assortative mate choice in the wild (Slade et al. 2019) and disassortative odour preferences in captivity (Grieves et al. 2019), possibly representing differences between mate preference and choice. The number of observed pairs in my

study was higher than that for the wild MHCII based study, in which a significant deviation from random mating was detected (Slade et al. 2019). Given that MHCII has a higher degree of polymorphism, it is likely that my study had greater statistical power and thus poor sampling is not a likely explanation for my not detecting deviation from random mating. Instead, the very small difference in means between observed (UniFrac distance = 0.34) and expected genetic distances (UniFrac distance = 0.33) suggests that any deviation from random mating that might be detectable through greatly increased sample sizes is not likely to be biologically significant.

It is possible that random mating at MHCI loci may be the result of females and males having other more pressing matters to consider when selecting a mate. Factors such as condition, territory quality, parental investment, and/or MHCII genotype of a potential partner could have higher impacts on breeding success than MHCI compatibility. While intermediate diversity at MHCII in song sparrows is associated with higher song repertoire size (Slade et al. 2017b) and thus presumably greater lifetime reproductive success (Reid et al. 2004), the relationship between MHCI diversity and susceptibility to haematozoan parasites (Slade et al. 2016) may have little effect on fitness. In mountain white-crowned sparrows (*Zonotrichia leucophrys oriantha*), infection with haematozoa was found to have no effect on or even to enhance LRS (Zylberberg et al. 2015). If so, MHCI genotype, which mediates haematozoan infections, may not strongly affect fitness.

Additionally, as song sparrows are socially monogamous, the lack of nonrandom mating I observed at MHCI may be a matter of constraints on social mate choice. I examined mate choice, which is not synonymous with mate preference. Females and males may not be able to make choices based on MHCI genotype due to being constrained by ecological conditions, territory preferences, arrival dates, the availability of potential mates, and competition within each sex for preferred mates. Mate selection in the wild may be more of a matter of who is “good enough” out of who is left, possibly based on more direct benefits, rather than an unconstrained choice of the most compatible individual. My findings do not exclude the possibility that MHCI similarity mediates choice of extra-pair mates, although less common in song sparrows than in other passerines (Hill et al. 2010), it remains a context which is likely much less constrained than that of social mate choice.

Alternatively, random mating with respect to MHCI genotype may reflect a sensory constraint; i.e passerines are unable to determine the MHCI allelic repertoires of potential mates. In mammals, MHCI-based mate choice is thought to be facilitated by the binding of MHCI peptides, which are secreted in various bodily fluids such as sweat and urine, to the vomeronasal organ, a sensory organ located in the soft tissue of the nasal septum containing various receptors for specific volatile and non-volatile compounds (Hawkes 2001; Leinders-Zufall et al. 2004). To date, this organ has not been found to be functionally present within any species of passerine (Silva and Antunes 2017). MHCII non-random mating on the other hand has been found to be facilitated by olfaction in passerines (Grieves et al. 2019). However, this is thought to be partially due to the influence of MHCII genotype on the microbial community surrounding the preen gland, which may affect how the volatile compounds released by the gland are processed (Grieves et al. 2019). MHCI would not have a similar effect on body odour as its receptors only deal with intracellular pathogens and would not influence the microbiome (generally, symbiotic bacteria living within the preen gland but outside of cells). Supporting this, house sparrows show no preference for the body odour of MHCI similar or dissimilar potential mates (Luisa et al. 2022), a finding consistent with an inability to assess MHCI similarity.

4.4 *Future Studies*

Future studies may benefit from including behavioural trials to further investigate whether song sparrows exhibit behavioural preferences for the odour of potential mates based on MHCI similarity. While the results of my study indicate that song sparrows show no behavioural preference, suggesting an inability to detect MHCI genotype, a controlled captive study may show that the ability to detect genotype is there, but it is simply obfuscated in the wild. Such a finding could suggest that MHCI has little impact on lifetime reproductive success, or at least less impact than other salient considerations such as direct benefits or MHCII compatibility. This possibility could be ideally addressed by a long-term field study in a relatively closed and isolated system such as Mandarte Island (Smith 2006) where all breeding attempts can be monitored and lifetime reproductive success estimated with confidence.

Additionally, examining whether MHCI genotype can be predicted by preen oil chemical composition, as has already been done for MHCII (Grieves et al. 2019), may help clarify whether another approach to conspecific genotype detection should be considered. It should be noted that most studies that investigate MHC based odour discrimination rely on odours sampled directly from the preen gland, which may not fully or accurately represent the chemosignals interpreted by the opposite sex. Blue petrel preen oil secretions showed significant qualitative and quantitative variations once spread onto the plumage (Mardon et al. 2011), possibly stemming from the partial breakdown of secretions via enzymatic lipases (Jacob and Pomeroy 1979; Jacob and Ziswiler 1982), microbial activity (Hagelin et al. 2003), and/or photolysis and oxidation (Wisthaler and Weschler 2010). Thus, characterizing whole-bird odour using a "headspace approach" (Saito et al. 2021) represents a promising approach for behavioural analyses.

5 *Conclusion*

My results further highlight the differences in the evolutionary trajectories of the two MHC classes and the need for more consideration for the less-studied class 1. In contrast to findings of non-random mating at MHCII (Grieves et al. 2019), including in my study population (Slade et al. 2019), I found that song sparrows paired randomly according to MHCI. Interestingly, I also found evidence of ancient and contemporary balancing selection at MHCI in the form of positively selected codons and TSPs. While balancing selection is common at MHC, these findings challenge the idea that sexual selection necessarily operates in concert with natural selection as inferred from molecular evolution and serve as a reminder that multiple factors must be considered. Based on the literature, although disassortative mating at MHCI would likely increase allelic diversity and thus pathogen resistance, I propose that song sparrows mate randomly at these loci due to perceptual constraints. That is, song sparrows and other birds may simply be unable to detect MHCI genotype, due to the lack of a functioning vomeronasal organ (Leinders-Zufall et al. 2004; Kelliher et al. 2006) combined with the fact that the community structure of symbiotic odour-producing bacteria appears sensitive to MHCII genotype (Grieves et al. 2021) but is presumably independent of MHCI genotype.

Overall, my findings help characterize past and present selection at MHCI and its potential evolutionary trajectory, while underlining the importance of perceptual constraint on the ability of sexual selection to amplify the diversifying effects of pathogen-mediated natural selection. Further exploring the evolution of genes involved in disease resistance is important because this work can help predict the capacity of wild populations to resist and adapt to new pathogens, a critical task in light of current trends of emerging and further geographical spread of infectious diseases.

Bibliography

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): 403–10. https://doi.org/10.1016/S0022-2836(05)80360-2

Álvarez-Carretero, S., Paschalia, K., Yang, Z. 2023. Beginner's guide on the use of PAML to detect positive selection. *Molecular Biology and Evolution,* 40(4). https://doi.org/10.1093/molbev/msad041

Andersson, M. 1994. *Sexual Selection*. Princeton University Press, Princeton, NJ.

Anisimova, M., and Gascuel, O., 2006. Approximate likelihood-ratio test for branches: a fast, accurate, and powerful alternative. *Systematic Biology,* 55(4): 539–52. https://doi.org/10.1080/10635150600755453

Arcese, P., Sogge, M.K., Marr, A.B., Patten. M.A. 2002. *Song Sparrow (Melospiza melodia), Version 2.0.* In The Birds of North America, No. 704 (Poole, A and Gill, F eds.). The Birds of North America, Inc., Philadelphia, PA.

Archer, N.M., Petersen, N., Clark, M.A., Buckee, C.O., Childs, L.M., and Duraisingh, M.T. 2018. Resistance to *Plasmodium falciparum* in sickle cell trait erythrocytes is driven by oxygen-dependent growth inhibition. *Proceedings of the National Academy of Sciences,* 115(28): 7350–55. https://doi.org/10.1073/pnas.1804388115.

Archie EA & Theis KR. 2011. Animal behaviour meets microbial ecology. *Animal Behaviour,* 82: 425-436. https://doi.org/10.1016/j.anbehav.2011.05.029

Bandelt, H. J., Forster, P., Röhl, A. 1999. Median-joining networks for inferring intraspecific phylogenies. *Molecular Biology and Evolution,* 16(1): 37–48. https://doi.org/10.1093/oxfordjournals.molbev.a026036

Barker, D.J., Maccari, G., Georgiou, X., Cooper, M.A., Flicek, P., Robinson, J., Marsh, S.G.E. 2023. The IPD-IMGT/HLA database. *Nucleic Acids Research,* 51(D1): D1053–60. https://doi.org/10.1093/nar/gkac1011

Bent, A.C., Austin, O.L. 1968. Life histories of North American cardinals, grosbeaks, buntings, towhees, finches, sparrows, and allies. *Bulletin of the United States National Museum,* 1-1889 http://repository.si.edu/xmlui/handle/10088/10027

Bernatchez, L., and C. Landry. 2003. MHC studies in non-model vertebrates: what have we learned about natural selection in 15 years?. *Journal of Evolutionary Biology*, 16(3): 363–77. https://doi.org/10.1046/j.1420-9101.2003.00531.x

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). *The Journal of Heredity*, 101(4): 448–60. https://doi.org/10.1093/jhered/esq018

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): 3307–15. https://doi.org/10.1111/j.1365-294X.2009.04242.x

Brown, J.L. 1997. A theory of mate choice based on heterozygosity. *Behavioral Ecology*, 8(1): 60–65. https://doi.org/10.1093/beheco/8.1.60.

Bryan, F.J., Manly, J.A. Navarro, A. 2020. *Randomization, bootstrap and Monte Carlo methods in biology*. 4th ed. CRC Press, Boca Raton, FL.

Byers, J.A, and Waits, L. 2006. Good genes sexual selection in nature. *Proceedings of the National Academy of Sciences*, 103(44): 16343–45. https://doi.org/10.1073/pnas.0608184103

Callahan, B.J., McMurdie, P.J., Rosen, M.J., Han, A.W., Johnson, A.J.A., Holmes, S.P. 2016. DADA2: high-resolution sample inference from illumina amplicon data. *Nature Methods*, 13(7): 581–83. https://doi.org/10.1038/nmeth.3869

Caro, S.P., Balthazart, J., Bonadonna, F. 2015. The perfume of reproduction in birds: chemosignaling in avian social life. *Hormones and Behavior*, 68: 25–42. https://doi.org/10.1016/j.yhbeh.2014.06.001.

Cloutier, A., Mills, J.A., Baker, A.J. 2011. Characterization and locus-specific typing of MHC class I genes in the red-billed gull (*Larus scopulinus*) provides evidence for major, minor, and nonclassical loci. *Immunogenetics*, 63: 377-394. https://doi.org/10.1007/s00251-011-0516-x

Chapman, F.M. 1940. *The post-glacial history of Zonotrichia capensis*. American Museum of Natural History, New York, NY.

Cohen S. (2002). Strong positive selection and habitat-specific amino acid substitution patterns in MHC from an estuarine fish under intense pollution stress. *Molecular Biology and Evolution*, 19(11):1870-1880 https://doi.org/10.1093/oxfordjournals.molbev.a004011

Collins, S. 2004. Vocal fighting and flirting: the functions of birdsong. *Nature's Music*, 2:39–79. https://doi.org/10.1016/B978-012473070-0/50005-0

Consuegra, S., and de Leaniz, C.G. 2008. MHC-mediated mate choice increases parasite resistance in salmon. *Proceedings of the Royal Society B*, 275(1641): 1397–1403. https://doi.org/10.1098/rspb.2008.0066

Cresswell, P. 1994. Assembly, transport, and function of MHC Class II molecules. *Annual Review of Immunology*, 12(1): 259–91. https://doi.org/10.1146/annurev.iy.12.040194.001355.

Cvijović, I., Good, B.H., Desai, M.M. 2018. The effect of strong purifying selection on genetic diversity. *Genetics*. 209(4): 1235-1278. https://doi.org/10.1534%2Fgenetics.118.301058

Demetra, A., Eizaguirre, C., Boehm, T., and Milinski, M. 2017. Mate choice in sticklebacks reveals that immunogenes can drive ecological speciation. *Behavioral Ecology*, 28(4): 953–61. https://doi.org/10.1093/beheco/arx074

Doherty, P.C., and Zinkernagel, R.M. 1975. Enhanced immunological surveillance in mice heterozygous at the H-2 gene complex. *Nature*, 256(5512): 50–52. https://doi.org/10.1038/256050a0

Dunn, J.C., Hawley, D.M., Huyvaert, K.P., Owen, J.C. 2021. *Infectious Disease Ecology of Wild Birds*. Oxford University Press, Oxford, UK. https://doi.org/10.1093/oso/9780198746249.003.0006

Edgar, R.C. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research*, 32(5): 1792–97. https://doi.org/10.1093/nar/gkh340

Eizaguirre, C., Yeates, S.E., Lenz, T.L., Kalbe, M., Milinski, M. 2009. MHC-based mate choice combines good genes and maintenance of MHC polymorphism. *Molecular Ecology*, 18(15): 3316–29. https://doi.org/10.1111/j.1365-294X.2009.04243.x

Eizaguirre, C., Lenz, T.L., Kalbe, M., Milinski, M. 2012. Divergent selection on locally adapted major histocompatibility complex immune genes experimentally proven in the field. *Ecology Letters*, 15(7): 723–31. https://doi.org/10.1111/j.1461-0248.2012.01791.x

Ekblom, R., Sæther, S.A., Grahn, M., Fiske, P., Kålås, J.A., Höglund, J. 2004. Major histocompatibility complex variation and mate choice in a lekking bird, the great snipe (*Gallinago media*). *Molecular Ecology*, 13(12): 3821–28. https://doi.org/10.1111/j.1365-294X.2004.02361.x

Ekblom, R., Stapley, J., Ball, A.D., Birkhead, T., Burke, T., Slate, J. 2011. Genetic mapping of the major histocompatibility complex in the zebra finch (*Taeniopygia guttata*). *Immunogenetics*, 63(8): 523–30. https://doi.org/10.1007/s00251-011-0525-9

Erickson, R.P. 1987. Natural history of the major histocompatibility complex. *American Journal of Human Genetics*, 40(5): 468–69. PMC1684141

Ewald, P.W. 1994. *Evolution of Infectious Disease*. Oxford University Press, Oxford, UK.

Follin, E., Karlsson, M., Lundegaard, C., Nielsen, M., Wallin, S., Paulsson, K., Westerdahl, H. 2013. In silico peptide-binding predictions of passerine MHC class I reveal similarities across distantly related species, suggesting convergence on the level of protein function. *Immunogenetics*, 65(4): 299–311. https://doi.org/10.1007/s00251-012-0676-3

Freeman-Gallant, C., Meguerdichian, M., Wheelwright, N., Sollecito, S. 2003. Social pairing and female mating fidelity predicted by RFLP similarity at the major histocompatibility complex in a songbird. *Biology*, 12(11). https://creativematter.skidmore.edu/bio_fac_schol/31

Furlong, R.F., and Yang, Z. 2008. Diversifying and purifying selection in the peptide binding region of DRB in mammals. *Journal of Molecular Evolution*, 66(4): 384–94. https://doi.org/10.1007/s00239-008-9092-6

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(10): e15406. https://doi.org/10.1371/journal.pone.0015406

Grant, B.R., and Grant, P.R. 1996. Cultural inheritance of song and its role in the evolution of Darwin's finches. *Evolution*, 50(6): 2471–87. https://doi.org/10.2307/2410714

Greenberg, R., Cadena, V., Danner, R.M., Tattersall, G. 2012. Heat loss may explain bill size differences between birds occupying different habitats. *PLoS ONE*, 7(7): e40933. https://doi.org/10.1371/journal.pone.0040933

Grieves, L.A., Gloor, G.B., Bernards, M.A., MacDougall-Shackleton, E.A. 2019. Songbirds show odour-based discrimination of similarity and diversity at the major histocompatibility complex. *Animal Behaviour*, 158:131–38. https://doi.org/10.1016/j.anbehav.2019.10.005.

Grieves, L.A., Gloor, G.B., Bernards, M.A., MacDougall-Shackleton, E.A. 2021. Preen gland microbiota covary with major histocompatibility complex genotype in a songbird. *Royal Society Open Science*, 8(10). https://doi.org/10.1098/rsos.210936

Grieves, L.A., Gilles, M., Cuthill, I.C., Székely, T., MacDougall-Shackleton, E.A. 2022. Olfactory camouflage and communication in birds. *Biological Reviews*, 97(3): 1193-1209. https://doi.org/10.1111/brv.12837

Griffiths, R., Double, M.C., Orr, K., Dawson, R.J. 1998. A DNA test to sex most birds. *Molecular Ecology*, 7(8): 1071–75. https://doi.org/10.1046/j.1365-294x.1998.00389.x

Guindon, S., Dufayard, J-F., Lefort, V., Anisimova, M., Hordijk, W., Gascuel, O. 2010. New algorithms and methods to estimate maximum-likelihood phylogenies:

assessing the performance of PhyML 3.0. *Systematic Biology*, 59(3): 307–21. https://doi.org/10.1093/sysbio/syq010

Hagelin, J.C., Jones, I.L., and Rasmussen, L.E.L. 2003. A tangerine-scented social odour in a monogamous seabird. *Proceedings of the Royal Society B: Biological Sciences*, 270(1522): 1323–29. https://doi.org/10.1098/rspb.2003.2379

Hawkes, C. 2001. The neurobiology of taste and smell. 2nd Edition.: Edited by Thomas E. Finger, Diego Restrepo and Wayne L. Silver. 2000. Chichester: John Wiley and Sons Price £71.50. Pp. 432. ISBN 0-47125-721-5." *Brain* 124 (7): 1468–69. https://doi.org/10.1093/brain/124.7.1468.

Hedrick, P. W. (2007). Balancing selection. *Current Biology*, 17(7), R230–R231. https://doi.org/10.1016/j.cub.2007.01.01

Hess, C.M., and Edwards, S.V. 2002. The evolution of the major histocompatibility complex in birds: scaling up and taking a genomic approach to the major histocompatibility 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(5): 423–31. https://doi.org/10.1641/0006-3568(2002)052[0423:TEOTMH]2.0.CO;2

Hewitt, E.W. 2003. The MHC class I antigen presentation pathway: strategies for viral immune evasion. *Immunology*, 110(2): 163–69. https://doi.org/10.1046/j.1365-2567.2003.01738.x

Hill, C.E., Akçay, C., Campbell, S.E., Beecher, M.D. 2010. Extrapair paternity, song, and genetic quality in song sparrows. *Behavioral Ecology*, 22(1): 73-81. https://doi-org.proxy1.lib.uwo.ca/10.1093/beheco/arq171

Huang, W., Dicks, K.L., Hadfield, J.D., Johnston, S.E., Ballingall, K.T., Pemberton, J.M. 2022. Contemporary selection on MHC genes in a free-living ruminant population. *Ecology Letters*, 25(4): 828–38. https://doi.org/10.1111/ele.13957.

Hughes, A.L., Yeager, M., Carrington, M. 1996. Peptide binding function and the paradox of HLA disease associations. *Immunology & Cell Biology*, 74(5):444-448. https://doi.org/10.1038/icb.1996.74

Ishigaki, K., Lagattuta, K.A., Luo, Y., James, E.A., Buckner, J.H., Raychaudhuri, S. 2022. HLA autoimmune risk alleles restrict the hypervariable region of T cell receptors. *Nature Genetics.* 54(4): 393–402. https://doi.org/10.1038/s41588-022-01032-z

Jacob, J., and Pomeroy, D.E. 1979. The feather lipids of the marabou stork (*Leptoptilos crumeniferus*). *Comparative Biochemistry and Physiology Part B: Comparative Biochemistry*, 64(3): 301–3. https://doi.org/10.1016/0305-0491(79)90148-2

Jacob, J., and Ziswiler, V. 1982. *The uropygial gland*. Chapter 4 in *Avian biology*. Academic Press, Amsterdam, NL.

Jäger, I., Eizaguirre, C., Griffiths, S.W., Kalbe, M., Krobbach, C.K., Reusch, T.B.H., Schaschl, H., Milinski, M. 2007. Individual MHC class I and MHC class IIB diversities are associated with male and female reproductive traits in the three-spined stickleback. *Journal of Evolutionary Biology*, 20(5): 2005–15. https://doi.org/10.1111/j.1420-9101.2007.01366.x

Jeffery, K.J.M., and Bangham, C.R.M. 2000. Do infectious diseases drive MHC diversity?. *Microbes and Infection*, 2(11): 1335–41. https://doi.org/10.1016/S1286-4579(00)01287-9.

Jamie, G.A., Meier, J.I. The persistence of polymorphisms across species radiations. *Trends in Ecology and Evolution*, 35: 795–808. https://doi.org/10.1016/j.tree.2020.04.007

Janeway, C.A. Jr., Travers, P., Walport, M., Shlomchik, M.J. 2001. *The major histocompatibility complex and its functions*, pp. 5.9-5.18. In *Immunobiology: the immune system in health and disease*: *fifth edition.* Garland Science, New York, NY.

Kaufman, J. 2022. *The avian major histocompatibility complex*, pp. 135-61. Chapter 7 in *Avian immunology (third edition)*. Academic Press, Boston, MA.

Kelliher, K.R., Spehr, M., Xiao-Hong, L., Zufall, F., Leinders-Zufall, T. 2006. pheromonal recognition memory induced by TRPC2-independent vomeronasal sensing. *European Journal of Neuroscience*, 23(12): 3385–90. https://doi.org/10.1111/j.1460-9568.2006.04866.x

Klein, J, Sato, A., Nagl, S., O'hUigín, C. 1998. Molecular trans-species polymorphism. *Annual Review of Ecology and Systematics*, 29(1): 1–21. https://doi.org/10.1146/annurev.ecolsys.29.1.1

Klein, J., A.Sato, and N.Nikolaidis. 2007. MHC, TSP, and the origin of species: from immunogenetics to evolutionary genetics. *Annual Review of Genetics*, 41:281–304. https://doi-org.proxy1.lib.uwo.ca/10.1146/annurev.genet.41.110306.130137

Kimura M. 1977. Prepondence of synonymous changes as evidence for the neutral theory of molecular evolution. *Nature*, 267: 275–276. https://doi.org/10.1038/267275a0

Kondrashov, F.A., and Koonin, E.V. 2004. A common framework for understanding the origin of genetic dominance and evolutionary fates of gene duplications. *Trends in Genetics*, 20(7): 287–90. https://doi.org/10.1016/j.tig.2004.05.001

Kosakovsky Pond, S.L., and 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., Ruff, J.S., Hyzer, C.W., Slev, P.R., Potts, W.K. 2012. Experimental viral evolution to specific host MHC genotypes reveals fitness and virulence trade-offs in alternative MHC types. *Proceedings of the National Academy of Sciences,* 109(9): 3422–27. https://doi.org/10.1073/pnas.1112633109

Kuduk, K., Babik, W., Bellemain, E., Valentini, A., Zedrosser, Z., Taberlet, P., Kindberg, J., Swenson, J.E., 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

Kulski, J.K., Shiina, T., Anzai, T., Kohara, S., Inoko, H. 2002. Comparative genomic analysis of the MHC: the evolution of class I duplication blocks, diversity and complexity from shark to man. *Immunological Reviews,* 190(1): 95–122. https://doi.org/10.1034/j.1600-065X.2002.19008.x

Lefort, V., Longueville, J-E., and Gascuel, O. 2017. SMS: smart model selection in PhyML. *Molecular Biology and Evolution,* 34(9): 2422–24. https://doi.org/10.1093/molbev/msx149

Leigh, J.W., and Bryant, D. 2015. Popart: full-feature software for haplotype network construction. *Methods in Ecology and Evolution,* 6(9): 1110–16. https://doi.org/10.1111/2041-210X.12410

Leinders-Zufall, T., Brennan, P., Widmayer, P., Chandramani S, P., Maul-Pavicic, A., Jäger, M., Xiao-Hong, L., Breer, H., Zufall, F., Boehm, T. 2004. MHC class I peptides as chemosensory signals in the vomeronasal organ. *Science,* 306(5698): 1033–37. https://doi.org/10.1126/science.1102818

Lenz, T.L. Eizaguirre, C., Kalbe, M., Milinski, M. 2013. Evaluating patterns of convergent evolution and trans-species polymorphism at MHC immunogens in two sympatric stickleback species. *Evolution,* 67(8): 2400-2412. https://doi.org/10.1111/evo.12124

Lewis, K. 1998. Pathogen resistance as the origin of kin altruism. *Journal of Theoretical Biology,* 193(2): 359–63. https://doi.org/10.1006/jtbi.1998.0725

Luisa, A., Amo de Paz, G., Kabbert, J., and Machordom, A. 2022. House sparrows do not exhibit a preference for the scent of potential partners with different MHC-I diversity and genetic distances. PLOS ONE, 17(12): e0278892. https://doi.org/10.1371/journal.pone.0278892

Loiseau, C., Richard, M., Garnier, S., Chastel, O., Julliard, R., Zoorob, R., Sorci, G. 2009. Diversifying selection on MHC class I in the house sparrow (*Passer domesticus). Molecular Ecology,* 18(7): 1331–40. https://doi.org/10.1111/j.1365-294X.2009.04105.x

Magadum, S., Banerjee, U., Murugan, P., Gangapur, D., Ravikesavan, R. 2013. Gene duplication as a major force in evolution. *Journal of Genetics,* 92(1): 155–61. https://doi.org/10.1007/s12041-013-0212-8

Maisonneuve, L., Chouteau, M., Joron, M., Llaurens, V. 2021. Evolution and genetic architecture of disassortative mating at a locus under heterozygote advantage. *Evolution,* 75(1): 149–65. https://doi.org/10.1111/evo.14129

Mancilla-Morales, M.D., Velarde, E., Contreras-Rodrigeuz, A., Gómez-Lunar, Z., Rosas-Rodriguez, J.A., Heras, J., Soñanez-Organis, J.G., Ruiz, E.A. 2022. Characterization, selection, and trans-species polymorphism in the MHC class II of heerman's gull (charadriiformes). *Genes,* 13(5): 917. https://doi.org/10.3390/genes13050917

Mardon, J., Saunders, S.M., and Bonadonna, F. 2011. From preen secretions to plumage: the chemical trajectory of blue petrels' *Halobaena caerulea* social scent. *Journal of Avian Biology*, 42(1): 29–38. https://doi.org/10.1111/j.1600-048X.2010.05113.x

Mays, H.L., and Hill, G.E. 2004. Choosing mates: good genes versus genes that are a good fit. *Trends in Ecology & Evolution,* 19(10): 554–59. https://doi.org/10.1016/j.tree.2004.07.018

Miller, M.M., and Taylor, R.L. 2016. Brief review of the chicken major histocompatibility complex: the genes, their distribution on chromosome 16, and their contributions to disease resistance. *Poultry Science*, 95(2): 375–92. https://doi.org/10.3382/ps/pev379

Milinski, M. 2022. A review of suggested mechanisms of MHC odor signaling. *Biology*, 11(1187). https://doi.org/10.3390/biology11081187

Minias, P., Pikus, E., Whittingham, L.A., Dunn, P.O. 2018. A global analysis of selection at the avian MHC. *Evolution*, 72(6): 1278–93. https://doi.org/10.1111/evo.13490.

Miyata T, and Yasunaga T. 1980. Molecular evolution of mRNA: a method for estimating evolutionary rates of synonymous and amino acid substitutions from homologous nucleotide sequences and its application. *Journal of Molecular Evolution*, 16(1): 23–36. https://doi.org/10.1007/bf01732067

Møller, A. P., and T. Szép. 2011. The role of parasites in ecology and evolution of migration and migratory connectivity. *Journal of Ornithology,* 152(1): 141–50. https://doi.org/10.1007/s10336-010-0621-x

Muller-Schwarze, Dietland. 2006. *Chemical Ecology of Vertebrates*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511607233

Murrell, B., Moola, S., Mabona, A., Weighill, T., Sheward, D., Kosakovsky Pond, S.L., Scheffler, K. 2013. FUBAR: a fast, unconstrained Bayesian approximation for

inferring selection. *Molecular Biology and Evolution*, 30(5): 1196–1205. https://doi.org/10.1093/molbev/mst030

Murrell, B., Wertheim, J.O., Moola, S., Weighill, T., Scheffler, K., Kosakovsky Pond, S.L. 2012. Detecting individual sites subject to episodic diversifying selection. *PLoS Genetics*, 8(7): e1002764. https://doi.org/10.1371/journal.pgen.1002764

Nash, A.A., Dalziel, R.G., and Fitzgerald, J.R. 2015. *Mechanisms of cell and tissue damage,* pp. 171-231. In *Mims' pathogenesis of infectious disease: sixth edition*. Academic Press, Cambridge, MA.

Neff, B.D., and Pitcher, T.E. 2005. Genetic quality and sexual selection: an integrated framework for good genes and compatible genes. *Molecular Ecology*, 14(1): 19–38. https://doi.org/10.1111/j.1365-294X.2004.02395.x

Nei, M., Gu, X., and Sitnikova, T. 1997. Evolution by the birth-and-death process in multigene families of the vertebrate immune system. *Proceedings of the National Academy of Sciences,* 94(15): 7799–7806. https://doi.org/10.1073/pnas.94.15.7799

Nelson, C.A., and Fremont, D.H. 1999. Structural principles of MHC class II antigen presentation. *Reviews in Immunogenetics,* 1(1): 47–59. https://pubmed.ncbi.nlm.nih.gov/11256572/

Owen, J.A., Punt, J., Stranford, S.A., Jones, P.P., Kuby, J. 2013. *Kuby Immunology*: *seventh edition*. W.H. Freeman and Company, New York, NY.

Ou, D., Mitchell, L.A., and Tingle, A.J. (1998). A new categorization of HLA DR alleles on a functional basis. *Human Immunology,* 59(10): 665-76. https://doi.org/10.1016/s0198-8859(98)00067-6

Penn, D.J., Damjanovich, K., and Potts, W.K. 2002. MHC heterozygosity confers a selective advantage against multiple-strain infections. *Proceedings of the National Academy of Sciences,* 99(17): 11260–64. https://doi.org/10.1073/pnas.162006499

Podos, J., and Warren, P.S. 2007. The evolution of geographic variation in birdsong. *Advances in the Study of Behavior*, 37: 403–58. https://doi.org/10.1016/S0065-3454(07)37009-5

Pruett, C.L., and Winker, K. 2010. Alaska song sparrows (Melospiza melodia) demonstrate that genetic marker and method of analysis matter in subspecies assessments, pp. 162-71. In *ornithological monographs:* 67(1). American Ornithological Society, Chicago, IL.

Radwan, J., Tkacz, A., and Kloch, A. 2008. MHC and preferences for male odour in the bank vole. *Ethology,* 114(9): 827–33. https://doi.org/10.1111/j.1439-0310.2008.01528.x

Reid, J.M., Arcese, P., Cassidy, A.L.E.V., Hiebert, S.M., Smith, J.N.M., Stoddard, P.K., Marr, A.B., Keller, L.F. 2004. Song repertoire size predicts initial mating success in male song sparrows, *Melospiza melodia*. *Animal Behaviour,* 68(5): 1055–63. https://doi.org/10.1016/j.anbehav.2004.07.003

Saito, K., Tokorodani, Y., Sakamoto, C., Kataoka, H. 2021. Headspace solid-phase microextraction/gas chromatography–mass spectrometry for the determination of 2-nonenal and its application to body odor analysis. *Molecules,* 26(19): 5739. https://doi.org/10.3390/molecules26195739

Sauermann, U., Nürnberg, P., Bercovitch, F., Berard, J., Trefilov, A., Widdig, A., Kessler, M., Schmidtke, J., Krawczak, M. 2001. Increased reproductive success of MHC class II heterozygous males among free-ranging rhesus macaques. *Human Genetics,* 108(3): 249–54. https://doi.org/10.1007/s004390100485

Schwensow, N., Dausmann, K., Eberle, M., Fietz, J., Sommer, S. 2010. Functional associations of similar MHC alleles and shared parasite species in two sympatric lemurs. *Infection, Genetics and Evolution*, 10(5): 662-668. https://doi.org/10.1016/j.meegid.2010.03.012

Searcy, W.A. 1984. Song repertoire size and female preferences in song sparrows. *Behavioral Ecology and Sociobiology,* 14(4): 281–86. https://doi.org/10.1007/BF00299499

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): 642–54. https://doi.org/10.1111/jeb.12600

Siefferman, L., and Hill, G. 2003. Structural and melanin coloration indicate parental effort and reproductive success in male eastern bluebirds. *Behavioral Ecology*, 14(6): 855–61. https://doi.org/10.1093/beheco/arg063.

Silva, L., and Antunes, A. 2017. Vomeronasal receptors in vertebrates and the evolution of pheromone detection. *Annual Review of Animal Biosciences,* 5: 353-370. https://doi.org/10.1146/annurev-animal-022516-022801

Sin, Y.W., Annavi, G., Newman, C., Buesching, C., Burke, T., Macdonald, D.W., Dugdale, H.W. 2015. MHC class II-assortative mate choice in European badgers (*Meles meles*). *Molecular Ecology*, 24(12): 3138–50. https://doi.org/10.1111/mec.13217

Singer, A.G., Beauchamp, G.K., and Yamazaki, K. 1997. Volatile signals of the major histocompatibility complex in male mouse urine. *Proceedings of the National Academy of Sciences,* 94(6): 2210–14. https://doi.org/10.1073/pnas.94.6.2210

Slade, J., Sarquis-Adamson, Y., Gloor, G.B., Lachance, M-A., MacDougall-Shackleton, E.A. 2017a. Population differences at MHC do not explain enhanced resistance of

song sparrows to local parasites. *Journal of Heredity*, 108(2): 127–34. https://doi.org/10.1093/jhered/esw082

Slade, J., Watson, M.J., MacDougall-Shackleton, E.A. 2017b. Birdsong signals individual diversity at the major histocompatibility complex. *Biology Letters*, 13(11). https://doi.org/10.1098/rsbl.2017.0430

Slade, J., Watson, M.J., and MacDougall-Shackleton, E.A. 2019. 'Balancing' balancing selection? Assortative mating at the major histocompatibility complex despite molecular signatures of balancing selection. *Ecology and Evolution*, 9(9): 5146–57. https://doi.org/10.1002/ece3.5087

Smith, J.N.M. 2006. *Conservation and biology of small populations: the song sparrows of Mandarte Island*. Oxford University Press, Oxford, UK.

Sommer, S. 2005. The importance of immune gene variability (MHC) in evolutionary ecology and conservation. *Frontiers in Zoology*, 2(1): 16. https://doi.org/10.1186/1742-9994-2-16

Spehr, M., J. Spehr, K. Ukhanov, K. R. Kelliher, T. Leinders-Zufall, and F. Zufall. 2006. Signaling in the chemosensory systems. *Cellular and Molecular Life Sciences*, 63(13): 1476–84. https://doi.org/10.1007/s00018-006-6109-4

Stervander, M., Dierickx, E.G., Thorley, J., Brooke, M. de L., 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. *Molecular Ecology*, 29(19): 3578–92. https://doi.org/10.1111/mec.15471

Stewart, K., and MacDougall-Shackleton, E.A. 2008. Local song elements indicate local genotypes and predict physiological condition in song sparrows *Melospiza melodia*. *Biology Letters*, 4(3): 240–42. https://doi.org/10.1098/rsbl.2008.0010

Strandh, M, Westerdahl, H., Pontarp, M., Canbäck, B., Dubois, M-P., Miquel, C., Taberlet, P., Bonadonna, F. 2012. Major histocompatibility complex class II compatibility, but not class I, predicts mate choice in a bird with highly developed olfaction. *Proceedings of the Royal Society B: Biological Sciences*, 279(1746): 4457–63. https://doi.org/10.1098/rspb.2012.1562

Tregenza, T., and N. Wedell. 2000. Genetic compatibility, mate choice and patterns of parentage: invited review. *Molecular Ecology*, 9(8): 1013–27. 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. https://doi.org/10.1007/BF00188427

Turček, F. J., and Hickey, J.J. 1951. Effect of introductions on two game populations in czechoslovakia. *The Journal of Wildlife Management*, 15(1): 113–14. https://doi.org/10.2307/3796784

Valkiunas, G. 2004. *Avian malaria parasites and other haemosporidia*. CRC Press, Boca Raton, FL.

Von Schantz, T., Wittzell, H., Göransson, G., Grahn, M. 1997. Mate choice, male condition-dependent ornamentation and MHC in the pheasant. *Hereditas*, 127(1–2): 133–40. https://doi.org/10.1111/j.1601-5223.1997.t01-1-00133.x

Wang, Z., Zhang, L., Qiao, A., Watson, K., Zhang, J., Fan, G-H. 2008. Activation of CXCR4 triggers ubiquitination and down-regulation of major histocompatibility complex class I (MHC-I) on epithelioid carcinoma HeLa cells. *Journal of Biological Chemistry*, 283(7): 3951–59. https://doi.org/10.1074/jbc.M706848200

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): 773–77. https://doi.org/10.1093/molbev/msx335

Westerdahl, H. 2004. No evidence of an MHC-based female mating preference in great reed warblers. *Molecular Ecology*, 13(8): 2465–70. https://doi.org/10.1111/j.1365-294X.2004.02238.x

Whelan, S., and 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): 691–99. https://doi.org/10.1093/oxfordjournals.molbev.a003851

Wilson, J.E., Katkere, B., and Drake, J.R. 2009. *Francisella tularensis* induces ubiquitin-dependent major histocompatibility complex class II degradation in activated macrophages. *Infection and Immunity*, 77(11): 4953–65. https://doi.org/10.1128/iai.00844-09

Wisthaler, A., and Weschler, C.J. 2010. Reactions of ozone with human skin lipids: sources of carbonyls, dicarbonyls, and hydroxycarbonyls in indoor air. *Proceedings of the National Academy of Sciences*, 107(15): 6568–75. https://doi.org/10.1073/pnas.0904498106

Wright, D.J., Brouwer, L., Mannarelli, M-E., Burke, T., Komdeur, J., Richardson, D.S. 2016. Social pairing of Seychelles warblers under reduced constraints: MHC, neutral heterozygosity, and age. *Behavioral Ecology*, 27(1): 295–303. https://doi.org/10.1093/beheco/arv150

Yang, Z. 2007. PAML 4: phylogenetic analysis by maximum likelihood. *Molecular Biology and Evolution*, 24(8): 1586–91. https://doi.org/10.1093/molbev/msm088

Curriculum Vitae

Name:	Scout Thompson
Post-secondary Education and Degrees:	University of Western Ontario London, Ontario, Canada 2016-2021 H.B.A. Genetics
Honours and Awards:	Western Graduate Research Scholarship 2021-2023 Biology Graduate Student Travel Award 2023 Biology Day Award of Excellence 2021 The Western Scholarship of Excellence 2016
Related Work Experience	Teaching Assistant The University of Western Ontario 2021-2023 Graduate Research Assistant The University of Western Ontario 2023

Publications:

Thompson, S.R.L., Lee, D.K., Lachance, M.A., Smith, D.R. (2021). Mutational effects of mobile introns on the mitochondrial genomes of *Metschnikowia* yeasts. *Frontiers in Genetics*, 12(785218). https://doi.org/10.3389/fgene.2021.785218