2-12-2018 1:00 PM

The Major Histocompatibility Complex in Song Sparrows: Immunity, Signals, and Mate Choice

Joel W.G. Slade
The University of Western Ontario

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

Graduate Program in Biology
A thesis submitted in partial fulfillment of the requirements for the degree in Doctor of Philosophy

Abstract

In recent years, sexual selection theory has redefined genetic quality to consider not only additive genetic effects on fitness but also non-additive genetic effects, such as heterozygote advantage or disadvantage. In jawed vertebrates, the major histocompatibility complex (MHC) gene family has been shown to exhibit both additive and non-additive genetic effects on fitness. MHC gene products are involved in initiating adaptive immune responses, and MHC genotype determines the range of pathogens to which an individual can respond. Therefore, parasite-mediated selection at MHC may favour locally-adapted, rare, or particular combination of alleles. Because heterozygote advantage at MHC is widespread, sexual selection should favour mechanisms by which individuals assess the MHC genotypes of potential mates, and mate non-randomly. Studies exploring the role of MHC in immunity and sexual selection are widespread amongst mammals and fish, but in birds (especially songbirds) there is relatively scant evidence for MHC-mediated mating and the mechanism by which this might be accomplished remains unknown. First, I assessed differentiation at MHC class I and II that might underlie locally-good gene effects in two populations of song sparrows (*Melospiza melodia*) previously shown to exhibit higher resistance to sympatric malaria (*Plasmodium*) strains. I found no population differentiation, suggesting no locally-good gene effects at MHC, but individuals with higher class I diversity were less likely to be infected when experimentally inoculated with *Plasmodium*. Second, I explored whether song sparrows convey information on MHC class II genotype through chemical (preen oil) or auditory (birdsong) cues. Pairwise similarity at MHC was related to pairwise similarity of preen oil chemical composition, but not to pairwise similarity in song repertoire content. Song repertoire size, a sexually selected trait in this species, was nonlinearly related to MHC diversity, such that males with intermediate MHC diversity sang the most songs. Finally, to investigate MHC-mediated mate choice, I compared MHC similarity of socially mated pairs of free-living song sparrows to random expectations. Contrary to my prediction of MHC-disassortative mating, social pairs were more similar at MHC than expected by chance. This work

emphasizes the importance of considering mate choice in the context of fitness effects at MHC.

Keywords

Assortative mating, birdsong, chemical communication, disassortative mating, genetic diversity, heterozygote advantage, major histocompatibility complex, mate choice, *Plasmodium*, preen wax, positive selection, song sparrow

Co-Authorship Statement

All data chapters are co-authored with Dr. Elizabeth MacDougall-Shackleton as the last (i.e., overseeing) author. Dr. MacDougall-Shackleton provided funding for all projects, collected field data, contributed to study design, provided advice on statistical analysis, and helped edit all chapters.

A version of Chapter 2 was published in the *Journal of Heredity* with Dr. Yanina Sarquis-Adamson, Dr. Gregory Gloor, Dr. Marc-André Lachance, and Dr. Elizabeth MacDougall-Shackleton as co-authors. Dr. Sarquis-Adamson performed the infectivity experiment. Dr. Gloor aided with the next-generation sequencing bioinformatics analysis and gave editorial comments on the manuscript. Dr. Lachance performed the population genetic differentiation analysis and gave editorial comments.

A version of Chapter 3 was published in *Proceedings of the Royal Society B* with Matthew Watson, Tosha Kelly, Dr. Gregory Gloor, Dr. Mark Bernards, and Dr. Elizabeth MacDougall-Shackleton as co-authors. Matthew Watson helped with lab work and bioinformatics analysis. Tosha Kelly collected the preen wax and blood samples in the field. Dr. Gloor aided with the next-generation sequencing bioinformatics analysis and gave editorial comments on the manuscript. Dr. Bernards helped with the chemical analysis of preen wax and gave editorial comments on the manuscript.

A version of Chapter 4 was published in *Biology Letters* with Matthew Watson and Dr. Elizabeth MacDougall-Shackleton. Matthew Watson helped with recording birdsong in the field, lab work, and bioinformatic analysis.

A version of Chapter 5 is being prepared for submission to *Molecular Ecology* with Matthew Watson and Dr. Elizabeth MacDougall-Shackleton. Matthew Watson helped with lab work and bioinformatic analysis.

Acknowledgments

First and foremost, I thank Dr. Elizabeth MacDougall-Shackleton for accepting me into her lab. Beth's insights, encouragement, and intellect helped guide me throughout my thesis. Not only did Beth help foster a positive work environment to accomplish my thesis research, but she also showed immense compassion towards my well-being. Ultimately, Beth has helped mold me to become proactive and compassionate in my future career.

I thank my advisory committee: Dr. Nusha Keyghobadi and Dr. Bryan Neff. Dr. Keyghobadi not only provided outstanding input on my research, but also served as an examiner for my research proposal and comprehensive assessment. Dr. Neff provided excellent advice on my research, and served as my thesis reader. Both of my advisors provided advice on research design, analyses, interpreting results, and ultimately helped me see the big picture in my work.

I greatly thank Matthew Watson for helping me throughout lab work, bioinformatics, and data collection in the field. Without Matthew, I would not have been able to process the genetic data efficiently. Matthew was more than just a colleague during his stay in the lab, he was (and still is) a friend. Thank you for your hard work and friendship, Matthew.

I thank other lab members and research assistants past and present, who not only provided advice, but helped foster a positive atmosphere in the lab. Erica Lovett helped me with lab work. Heather McGillivray and Dr. Dominque Potvin helped with field work in 2014. Dr. Yanina Sarquis-Adamson not only helped with lab work, but she provided the foundational work for chapter 2 of my thesis. Without Tosha Kelly, I would not have genetic or preen wax samples. Tosha is also a master at finding song sparrow nests, as is Alannah Lymburner – thank you both for help providing nesting pair data. Finally, I thank Leanne Grieves for her collegial support towards the end of my degree.

This thesis would also not have been completed without the help from Dr. Marc-André Lachance, Dr. Mark Bernards, Dr. Greg Gloor, Dr. Jeremy McNeil, Dr. Ben Rubin,

and Dr. Shawn Garner. Dr. Lachance helped me throughout my entire degree with evolutionary and phylogenetic analysis, and was an examiner for my comprehensive assessment. Dr. Bernards taught me how to prepare and analyze chemical data from preen wax. Dr. Greg Gloor helped with processing my next-generation sequencing data. Dr. McNeil helped me to understand concepts of chemical ecology and was an examiner in my comprehensive assessment. Dr. Rubin provided advice on statistical analysis of my thesis data. Finally, Dr. Shawn Garner helped me analyze the MHC-mediated mate choice data.

Funding for my doctoral degree came from Dr. Elizabeth MacDougall-Shackleton's National Science and Engineering Council of Canada (NSERC) discovery grant and my Ontario Graduate Scholarships (OGS) from 2015 to 2017. Additional research funds came from a Taverner Award from the Society of Canadian Ornithologists and a Hesse Award from the American Ornithologists' Union.

I thank those who provided emotional support and encouragement outside of my lab. Thank you to my parents, Brenda and Billy Slade, and my grandmother, Lois Slaunwhite for being there for me. I also thank my best friends, Sean and Brent Sweeney for providing me an endless supply of laughter. I also thank the Pride Men's Chorus London (PMCL) for helping me obtain so much happiness through comradery and music. Finally, to all of my fellow graduate friends outside of my lab, you know who you are, and you know how important you were (and still are) in my life.

Table of Contents

List of Tables

List of Figures

List of Abbreviations and Symbols

α2	Alpha two domain of major histocompatibility complex class I
β	Beta
ΔAIC	Disparity in Akaike information criterion
ΔAICc	Disparity in corrected Akaike information criterion
ω	Omega (dN/dS)
$[M]^+$	Protonated molecule concentration
A	Adenine
AFLP	Amplified fragment length polymorphism
AIC	Akaike information criterion
AICc	Corrected Akaike information criterion
BIOENV	Statistical method that identifies candidate environmental variables (e.g., GC peaks of preen wax) based on community structure (e.g., MHC UniFrac distance)
BLAST	Basic local alignment search tool
BO	Bird origin
BOPO	Bird origin by parasite origin interaction
bp	Base pair
C	Cytosine or Carbon
cDNA	Complementary DNA
CI	Confidence interval
DNA	Deoxyribonucleic acid
dNTPs	Deoxynucleoside triphosphates
dN	Nonsynonymous base substitution
dS	Synonymous base substitution
D_S	Mean within-site genetic distance
D_T	Between-site genetic distance
EDTA	Ethylenediaminetetraacetic acid
Exo	Exonuclease
ExPASy	Expert protein analysis system
FastAP	Novel thermosensitive alkaline phosphatase by Thermo Scientific
FID	Flame ionization detection
F_{ST}	Fixation index (genetic population differentiation)
G	Guanine
GC	Gas chromatography
gDNA	Genomic DNA
H	Test statistic for a non-parametric one-way analysis of variance
IIβ	Major histocompatibility complex class II exon 2
IS	Internal standard
ISU	Individual sequencing unit
K	Number of model parameters

K	Number of genetic population clusters
k	Number of environmental parameters used in BIOENV
KCl	Potassium chloride
lnL	Natural logarithm-likelihood
logLik	Log-likelihood
MCMC	Markov chain Monte Carlo
MEGA	Molecular evolutionary genetic analysis
$MgCl_2$	Magnesium chloride
MHC	Major histocompatibility complex
MHCI	Major histocompatibility complex class I
MHCII	Major histocompatibility complex class II
MOE	Main olfactory epithelia
MS	Mass spectrometry
n	Sample size
p	Probability
PAML	Phylogenetic analysis by maximum likelihood
PBR	Peptide binding region (for example, of the major histocompatibility complex)
PCR	Polymerase chain reaction
pGEM-T	*Escherichia coli* plasmid containing an ampicillin resistant gene with a rapid thymine overhanging cloning site
pH	Potential of hydrogen
PO	Parasite origin
PREV	Previous infection status
P-SOSP	Lineage of *Plasmodium* found in song sparrows
R^2	Coefficient of determination
r	Correlation coefficient
RNA	Ribonucleic acid
SA	Superallele
Sosp-DAB*	Putative functional MHC class II exon 2 allele in song sparrows
T	Thymine
t	Test statistic following a Student's t-distribution
Taq	Thermus aquaticus
Tris-HCl	Tromethamine – hydrochloric acid
TRIzol	Solution used to isolate RNA
U	Units
UniFrac	Unique fraction metric, a measure of phylogenetic distance between taxa or between alleles
VNO	Vomeronasal organ
WAG	Whelan and Goldman amino acid substitution model
Y	Pyrimidine (Cytosine or Thymine)

List of Appendices

Chapter 1

1 General Introduction

1.1 Genetic effects on fitness and mate choice

In many animal species, females do not mate with the first male they encounter, but instead invest time and energy into mate choice (Andersson 1994; Andersson and Simmons 2006; Clutton-Brock 2007). The field of sexual selection attempts to explain the persistence of mate choice, as well as the widespread existence of conspicuous, costly, and sexually dimorphic traits (Zahavi 1975; Andersson 1986, 1994; Andersson and Simmons 2006). Originally, female choice and male ornamentation were explained solely on the basis of high-quality males providing resources (direct benefits) for females (Emlen and Oring 1977). Eventually additive genetic effects were investigated to explain female choice and male ornamentation (Mays and Hill 2004). That is, the fitness effect of a given allele is independent of the rest of the genome (Neff and Pitcher 2005). Recently, however, sexual selection theory has expanded to recognize the importance of non-additive genetic effects on fitness and mate choice. Under non-additive genetic effects, the fitness effect of a given allele depends on the rest of the genome (the allele's homologue at the same locus, and/or alleles at other loci; Neff and Pitcher 2005). Alleles that enhance fitness through additive versus non-additive effects have been termed "good genes" and "compatible genes", respectively (Mays and Hill 2004; Neff and Pitcher 2005).

1.1.1 Additive (good gene) effects

Additive, or good gene, effects on fitness have historically been the main explanation for high quality ornaments in animals where mate choice exists (Mays and Hill 2004). The co-evolutionary relationship between female choice and male ornamentation is posited as an honest signals of quality, whereby males provide direct (material resources, Emlen and Oring 1977) and indirect (genetic) benefits, and that the sexually selected characteristic and preference for said characteristic is heritable (e.g., Zuk et al. 1990; Andersson 1994). An explanation of additive good gene models of fitness and mate choice is the Hamilton-

Zuk hypothesis, whereby animals that possess bright and elaborate ornamentation possess heritable good genes to fight off parasites, because resources can be allocated to produce extravagant ornamentation rather than in immunity (Hamilton and Zuk 1982). Thus, these traits act as honest indicators of genetic quality whereby selection acts on the heritable variance of these traits.

Because environments, and thus selective pressures, vary spatially, additive genetic effects on fitness and mate choice can include locally-good gene effects (Reinhold 2004). Geographically variable selection pressures may result in individuals originating from the local population having locally-good genes that confer heritable resistance to the local pathogens or other aspects of the local environment. Common garden experiments have provided evidence of local adaptation, such as salinity in frogs (*Bufo calamita*, Gomez-Mestre et al. 2003), ecotype differentiation in anoles (*Anole oculatus*, Thorpe et al. 2005), and basal metabolic rates in great tits (*Parus major*, Broggi et al. 2005), whereby genetically local individuals hold a fitness advantage over those from other populations. For example, in song sparrows (*Melospiza melodia*), males that sing local song types had lower parasite load over birds with less local songs (Stewart and MacDougall-Shackleton 2008). This may help explain preference for local traits, such as birdsong (Searcy et al. 2002) since females may acquire locally-good genes for their offspring.

1.1.2 Non-additive (compatible gene) effects

Non-additive, or compatible gene, effects on fitness occur when the fitness effect of an allele depends on the rest of the genome. Thus, compatible alleles may increase the fitness of the bearer through dominance or overdominance effects at homologous alleles in diploids or through epistatic effects at other loci (Neff and Pitcher 2005). Overdominance, also known as heterozygote advantage (Brown 1997), is widely invoked as a benefit of individual genetic diversity; in particular, genetically diverse individuals may have superior pathogen resistance (e.g., Penn et al. 2002; McClelland et al. 2003b; MacDougall-Shackleton et al. 2005; Oliver et al. 2009). Individuals that select mates with dissimilar genotypes should produce offspring with higher fitness, for example due to heterozygote advantage, than individuals choosing mates with similar genotypes (Tregenza and Wedell

2000). For example, in white-crowned sparrows (*Zonotrichia leucophrys*), individuals that had higher levels of genome-wide heterozygosity were not infected with bloodborne parasites when compared to homozygous individuals (MacDougall-Shackleton et al. 2005). Also, in common bluethroats (*Luscinia svecica*), extra-pair young were more immunocompetent than within-pair young, implying a compatible gene effect obtained through extra-pair copulations (Jonsen et al. 2000). However, non-additive genetic effects are not always associated with maximally heterozygous genotypes having the highest fitness. For example, individuals selecting mates with maximally divergent genotypes may produce offspring of low fitness due to outbreeding depression (Tregenza and Wedell 2000).

Although extravagant ornaments are frequently explained by invoking indirect (genetic) benefits (Zahavi 1975; Kodric-Brown and Brown 1984; Andersson 1986, 1994; Andersson and Simmons 2006), these ornaments can also provide information on direct benefits, whereby the choosier sex obtains material resources, such as territory quality (e.g., Keyser and Hill 2000). In particular, ornaments may signal direct benefits in species where males provide substantial parental care, such as in birds. For example, male eastern bluebirds (*Sialia sialis*) with high quality plumage (larger and brighter breast patches) provisioned their offspring more than males of lesser quality (Siefferman and Hill 2003). Likewise, the complexity of birdsong (an acoustic ornament) predicted parental effort in sedge warblers (*Arcocephalus schoenobaenus*, Buchanan and Catchpole 2000).

1.2 The major histocompatibility complex (MHC)

The major histocompatibility complex (hereafter, MHC) is a gene family of particular interest to behavioural and evolutionary ecology, and to the field of sexual selection, because both additive and non-additive genetic effects on fitness and mate choice have been well-documented at these loci (Bernatchez and Landry 2003; Milinski 2006; Piertney and Oliver 2006; Trowsdale 2011). The MHC is the cornerstone of the adaptive immune system in jawed vertebrates. MHC genes encode cell-surface receptors that present both self and non-self proteins to T cells to elicit an immune response (Trowsdale 2011). MHC

gene products can be categorized into class I and class II molecules, which interact primarily with intracellular and extracellular pathogens, respectively.

1.2.1 MHC class I

MHC class I genes are expressed in all nucleated cells and are involved with presenting intracellular peptides to the immune system (Yewdell and Bennink 1992). Peptides from intracellular pathogens, such as viruses and protozoan parasites including malaria (*Plasmodium* spp.) bind to the peptide-binding region (PBR) of MHC class I, and are then transported to the outside of the cell. Cytotoxic T cells bind to the MHC class I molecule, and through interacting with the antigen presented, recognize that the cell is infected. This prompts cytotoxic T cells to destroy the infected cell to prevent the spread of the pathogen (Kulski et al. 2002). In contrast, MHC class I molecules from healthy cells present self-derived peptides to cytotoxic T cells and are not destroyed. The most variable region of MHC class I is exon 3, which encodes part of the peptide-binding region (PBR; $\alpha 2$ domain). Both class I and class II are under high selective pressure due to the co-evolutionary arms race between pathogens and their host (Kubinak et al. 2012).

1.2.2 MHC class II

MHC class II genes are expressed only in antigen-presenting cells (e.g., dendritic cells, macrophages, and B cells; Klein 1986). Such cells phagocytize extracellular pathogens such as bacteria and nematodes, and present their peptides to helper T cells. This prompts an immune response to the site of infection. The highly variable exon 2 of MHC class II encodes the PBR (β domain), which is the main focus for studies on this class of MHC (Bernatchez and Landry 2003).

1.3 Processes driving MHC diversity

MHC is of biological interest due not only to its important role in disease resistance, but also to its tremendous diversity: the majority of jawed vertebrates express multiple MHC alleles at both classes. Moreover, levels of MHC diversity vary tremendously between taxa, as well as between class I and II molecules. For example, the entire human (*Homo sapiens*)

population contains over 6000 MHC class I alleles and over 1000 class II alleles (Robinson et al. 2013), while a population of lake whitefish (*Coregonus* sp.) possess 15 class I alleles and 20 class II alleles (Binz et al. 2001). Some of this diversity reflects duplication events: over evolutionary time, MHC genes in many vertebrate lineages have become duplicated and eventually acquire new mutations that may allow individuals to recognize additional new pathogens (Kulski et al. 2002). Within birds, the passerines have undergone high levels of duplication at MHC (O'Connor et al. 2016). For example, at MHC class I, at least 38 different gDNA alleles (implying a minimum of 19 genomic loci) have been confirmed within a single individual in willow warblers (*Phylloscopus trochilus*, O'Connor et al. 2016). Similarly, at MHC class II, 39 different gDNA alleles and 16 different cDNA alleles have been confirmed within a single individual in common yellowthroats (*Geothlypis trichas*, Bollmer et al. 2010), implying a minimum of 20 genomic loci, and at least eight transcribed loci. Over shorter time-scales, however, MHC diversity is also thought to be maintained in natural populations through balancing selection. Specifically, heterozygote advantage at MHC, negative frequency-dependent selection favouring rare MHC alleles, and MHC-mediated disassortative mating have all been proposed to contribute to high levels of MHC diversity found in most natural populations (Kubinak et al. 2012).

1.3.1 Heterozygote advantage, antagonistic co-evolution, and positive selection at MHC

Heterozygote advantage is at the forefront of research to explain polymorphisms at the PBR of both MHC class I and II (e.g., Penn et al. 2002; Oliver et al. 2009; Kubinak et al. 2012; Niskanen et al. 2014; Galaverni et al. 2015). Originally, models of pathogen-mediated selection assumed that increased diversity at MHC should always be selectively favoured (Kubinak et al. 2012; Ruff et al. 2012), as individuals with a greater diversity of alleles (particularly at the PBR, involved in antigen recognition and binding) should be able to recognize and respond to a greater diversity of pathogens. Supporting this model, many studies examining the role of MHC diversity in surviving multi-strain infections and/or effects of low MHC allelic diversity (e.g., inbred strains) on disease resistance do support the idea that increased individual genetic diversity at MHC is associated with enhanced disease resistance (e.g., mice *Mus musculus domesticus*, 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 studies have found fitness advantages associated with intermediate rather than maximal MHC diversity (e.g., three-spined sticklebacks *Gasterosteus aculeatus*, Wegner et al. 2003, 2004, Jäger et al. 2007). This phenomenon may be explained by the immune system only being able to respond to a finite number of antigens (Yewdell and Bennink 1999). Alternatively, individuals presenting too many peptides to T cells could be more susceptible to autoimmune responses (Bottazzo et al. 1986; Apanius et al. 1997) or a dilution effect may occur, whereby individuals only have limited copies of protective alleles (Wegner et al. 2003; Kubinak et al. 2012). Therefore, even though some taxa show evidence of heterozygote advantage at MHC, in other taxa the optimal level of MHC diversity may be intermediate rather than maximal.

Antagonistic co-evolution with pathogens is also thought to be a driver of MHC diversity (Sutton et al. 2011). Through an evolutionary arms race, pathogens evolve to evade the host population's most common MHC alleles, particularly at the PBR. Negative frequency-dependent selection then favours hosts that possess rare alleles (Clarke and Kirby 1966; Slade and McCallum 1992). This process is thought to increase MHC diversity because new mutations add rare alleles to the host population, which then become prevalent through antagonistic co-evolution and negative frequency-dependent selection (Potts and Wakeland 1990; Jeffery and Bangham 2000).

Geographic variation in parasite assemblages (Pagenkopp et al. 2008) may generate geographic variation in selection pressures at MHC, which may drive population differentiation at these loci (e.g., Loiseau et al. 2009; Bichet et al. 2015). Sarquis-Adamson and MacDougall-Shackleton (2016) showed experimentally that song sparrows were more resistant to a sympatric *Plasmodium* strain over an allopatric strain. They hypothesized that population differences at MHC may explain resistance to local parasites through (additive) locally-good gene effects. Alternatively, however, such effects might reflect immune memory resulting from individuals having prior exposure to the local *Plasmodium* strains. In Chapter 2 of this thesis, I test for additive (locally-good genes) effects and non-additive

(heterozygote advantage) effects at class I and class II of MHC, on the resistance of song sparrows to *Plasmodium* parasites.

Considering pathogen-mediated selection on MHC, codons within the PBR often show signatures of positive selection (i.e., higher levels of nonsynonymous base substitutions in comparison to synonymous ones). Because pathogens are constantly evolving, and the PBR makes direct contact with their antigens, mutations that give rise to functional change that recognize new pathogens are likely to be maintained in the population (Hughes and Nei 1988; Yang and Swanson 2002). To shine light on selection at the PBR, in chapter 5 of this thesis, I test for signatures of positive selection at the PBR in class II MHC of song sparrows.

1.4 MHC-mediated mate choice

1.4.1 Mate choice based on MHC genotype (compatibility)

MHC-mediated mate choice often implies the choosier sex mates with an individual that has an optimally-dissimilar MHC genotype (Milinski 2006). MHC-disassortative mate choice is normally explained using the previously described models of maximizing offspring fitness through heterozygote advantage, negative-frequency dependent selection, and/or inbreeding avoidance (Kubinak et al. 2012). However, MHC-assortative mate choice does occur, and is posited to reduce outbreeding depression in populations and/or the maintenance of co-adapted gene complexes (Tregenza and Wedell 2000; Bos et al. 2009; Sin et al. 2015).

Studies have revealed extensive MHC-mediated mate preference in mammals. Yamazaki et al. (1976) first discovered MHC-disassortative mate preference when they noted male mice chose females in oestrus that were dissimilar at MHC. The majority of the research in mammals documents MHC-disassortative over MHC-assortative mate preference, such as in bank voles (*Clethrionomys glareolus*, Radwan et al. 2008), humans (Wedekind et al. 1995; Havlicek and Roberts 2009), mandrills (*Mandrillus sphinx*, Setchell et al. 2010), fat-tailed lemurs (*Cheirogaleus medius*, Schwensow et al. 2008b), and grey mouse lemurs (*Microcebus murinus*, Schwensow et al. 2008a). However, MHC-assortative

mating has been documented in the Malagasy jumping rat (*Hypogeomys antimena*, Sommer 2005), and European badgers (*Meles meles*, Sin et al. 2015). Regardless of the direction of MHC genotype-based mate choice in mammalian systems, these studies showcase the adaptive importance of MHC-mediated mating in mammalian systems, such as heterozygote advantage (disassortative) (e.g., Radwan et al. 2008) or outbreeding avoidance (assortative) (e.g., Sin et al. 2015).

Further research has identified MHC-mediated mating in other jawed vertebrates (Ruff et al. 2012). In fish, MHC-disassortative mate preference has been documented in, for example, Atlantic salmon (Landry et al. 2001; Rajakaruna et al. 2006; Consuegra and Garcia de Leaniz 2008), Chinook salmon (*Oncorhyncus tshawytscha*, Neff et al. 2008), and the rose bitterling (*Rhodeus ocellatus*, Agbali et al. 2010). In reptiles, MHC-disassortative mate preference has been documented in sand lizards (*Lacerta agilis*, Olsson et al. 2003), and tuatara (*Sphenodon punctatus*, Miller et al. 2009). Contrary to other trends in MHC-mediated mate preference, the only documented case in any amphibians is MHC-assortative preference in tiger salamanders (*Ambystoma tigrinum*, Bos et al. 2009). Finally, research in birds has revealed both directions of MHC-mediated mate preferences. MHC-disassortative mating has been documented in savannah sparrows (*Passerculus sandwichensis*, Freeman-Galant et al. 2003), and in red jungle fowl (*Gallus gallus*, Gillingham et al. 2009). Complex MHC-assortative mating has been documented in house sparrows (*Passer domesticus*, Bonneaud et al. 2006), whereby males with high MHC diversity were preferred as mates but only if they shared MHC alleles with females. Overall, studies examining MHC-mediated mating preferences do not follow a single trend of MHC-disassortative or –assortative mate choice. This highlights how selection may favour either direction of preference depending on the adaptive importance of MHC in these various mating systems. In chapter 5 of this thesis, I investigate the hypothesis that song sparrows choose social mates that are dissimilar to themselves at MHC class IIβ.

1.4.2 Mate choice based on MHC diversity

In contrast to choosing a mate based on its MHC similarity or dissimilarity, thus influencing MHC diversity of the resultant offspring, some taxa show evidence for mate

choice based on MHC diversity of the potential mate itself. That is, individuals may prefer a mate with optimal (whether this involves maximal or intermediate) levels of MHC diversity. Particularly in species where direct benefits are important to mate choice and fitness, the choosier sex may benefit through pairing with an optimally MHC-diverse partner because such individuals are expected to be in superior condition and thus better able to provide direct or material benefits. As noted above, optimal MHC diversity may correspond to either maximal or intermediate levels of diversity, and is likely to vary among taxa.

Preference for maximally-diverse individuals at MHC has been documented in several species. For example, in fat-tailed dwarf lemurs, females preferred MHC-dissimilar males, but also preferred males with maximal MHC diversity (Schwensow et al. 2008b). In house sparrows, males that successfully acquired a mate had higher MHC diversity than those that did not (Bonneaud et al. 2006). Meanwhile, female Seychelle warblers (*Acrocephalus sechellensis*) tend to produce extra-pair young with males that are more diverse at MHC than their social mates, implying that this preference for maximally diverse males may confer indirect benefits, despite the fact that females in this species do not appear to prefer maximally dissimilar social males with whom they would produce maximally diverse offspring (Richardson et al. 2005). Similarly, peahens (*Pavo cristatus*) paired with maximally MHC-diverse males lay larger than average eggs and larger than average clutches (Hale et al. 2009).

In other species, individuals choose mates based on MHC diversity but prefer intermediate rather than maximal MHC diversity (e.g., house sparrows, Bonneaud et al. 2006; brown trout *Salmo trutta*, Forsberg et al. 2007). An extensive example of this phenomenon is in three-spined sticklebacks, where females prefer mates with intermediate MHC diversity (e.g., Reusch et al. 2001; Milinksi 2006). These intermediately-diverse males also made the best nests (Jäger et al. 2007) and had the lowest parasite loads (Wegner et al. 2003), traits which presumably reflect direct benefits.

1.5 Signals of MHC

1.5.1 Signals of MHC genotype

The discovery of MHC-mediated mate choice in lab mice by Yamazaki et al. (1976) prompted research into understanding how vertebrates may evaluate MHC genotypes of potential mates. Mammals, and in particular, rodents, have an olfactory structure called the vomeronasal organ (VNO), which is used to detect nonvolatile molecules used in conspecific social and sexual status signals (Ruff et al. 2012). In mice, specific VNO neurons respond to synthetic MHC class I peptide-ligands that have identical amino acid sequences to naturally derived MHC peptides (Leinders-Zufall et al. 2004). This provides a direct physiological mechanism to detect and differentiate MHC genotypes based on nonvolatile cues in mammals. Although not all vertebrates possess a VNO, all have main olfactory epithelia (MOE) (Müller-Schwarze 2006). Indeed, olfactory sensory neurons in the MOE in mice also respond to MHC class I peptide-ligands, highlighting that MHC genotypes can be detected via both physiological pathways (Spehr et al. 2006). This implies that jawed vertebrates lacking a VNO may still process signals of MHC genotypes.

Similar to studies in mice, altering concentration of synthetic MHC peptides directly influences mating decisions in three-spined sticklebacks (Milinski et al. 2005), but the direct mechanism for peptide-detection in fishes is still unknown (Boehm and Zufall 2006). Additionally, a study on female humans revealed that they preferred their own scent when presented with various synthetic MHC class I peptides, highlighting discrimination between self and non-self MHC (Milinski et al. 2013). Overall, these studies indicate that mammals and fish have the capability to directly detect MHC genotypes of conspecifics, which has mate choice implications.

Discriminating MHC genotypes directly through olfactory cues dominates the research performed in mammals and fish. In most bird species, however (with the exception of seabirds, Nevitt 2000) olfactory ability has traditionally been assumed to be weak, because birds have poorly developed olfactory bulbs (Bang and Cobb 1968) and lack a vomeronasal organ (Keverne 1999). As a result, the role of olfaction in communication and

mate choice in birds has received little study. However, recent studies (Leclaire et al. 2012, 2014) have identified a potential source of odour cues through which birds may be able to assess the genetic makeup of potential mates and other conspecific individuals. Avian preen oil is a waxy secretion from the uropygial gland that functions primarily in protecting feathers and maintaining hydrophobicity (Jacob and Ziswiler 1982). The oil is composed of monoesters (Thomas et al. 2010) and diesters (Tuttle et al. 2014) as well as volatile compounds such as linear alcohols (Whittaker et al. 2010). MHC profiles may be associated with preen oil composition either relatively directly (through potentially influencing the microbiome, which modifies preen oil odour components, Strandh et al. 2012) and/or indirectly (if geographic variation in preen oil profiles (Whittaker et al. 2010) corresponds to geographic variation at MHC). Supporting a potential association between preen oil profile and genetic makeup, dark-eyed juncos (*Junco hyemalis*) exposed to conspecific preen oil can discriminate between sexes, and between populations (Whittaker et al. 2010, 2011). Moreover, in seabirds, the diversity of an individual's preen oil compounds reflects its neutral-locus genetic diversity (Leclaire et al. 2012), and pairwise similarity in preen oil composition is associated with pairwise genetic similarity at class II MHC loci (Leclaire et al. 2014), and MHC-disassortative mate preference (Leclaire et al. 2017). Therefore, in birds, the mechanism by which they may be able to discriminate MHC genotypes is likely to be indirectly influenced by preen oil composition rather than direct detection of MHC peptide-ligands found in mammals and fish. Chemical signals of MHC in songbirds remains unexplored, therefore, in Chapter 3 I examine the relationship between the chemical composition of preen oil and MHC class IIβ genotype in song sparrows.

Ornamentation and other condition-related traits may indirectly signal MHC genotype through the effects of MHC on disease resistance and body condition. In great snipe (*Gallinago media*), males with particular MHC genotypes were larger and preferred as mates (Ekblom et al. 2004). Similarly, MHC genotype predicted spur length, a sexually selected trait in ring-necked pheasants (*Phasianus colchicus*, von Schantz et al. 1997). Therefore, even though the majority of MHC genotype signalling appears to involve

chemical communication, other modalities may also contribute to signals of MHC genotype.

In light of exploring indirect signals of MHC, geographic variation in learned vocalizations represents a compelling candidate mechanism by which females may be able to assess a potential mate's population of origin. In oscine songbirds (Passeri), birdsong is learned from older conspecifics, and such learning is often restricted to a sensitive period during early life (Marler and Peters 1987). As a result, in many species, song varies geographically (Podos and Warren 2007), such that the songs within a male's repertoire may reflect his population of origin. Geographic variation in song is behaviourally salient, as females of many species prefer local song types (Baker et al. 1981; Searcy et al. 2002). Such preferences may enable females to obtain locally-good genes at MHC for their offspring. Conversely, in populations where non-additive effects on fitness at MHC are important, females might use song to avoid inbreeding (Grant and Grant 1996), for example by preferentially pairing with males of non-local origin. Locally-good gene effects at MHC may operate concurrently with and in opposition to non-additive genetic effects such as heterozygote advantage (Marr et al. 2002; MacDougall-Shackleton et al. 2005). Whether females use song as a cue to obtain locally-good genes (Baker et al. 1981), or to avoid inbreeding (Grant and Grant 1996), assessing a potential mate's population of origin should be particularly important if geographic variation in parasite assemblages (Pagenkopp et al. 2008) has generated geographic variation at MHC. In Chapter 4, I explore the relationship between variation in birdsong composition and MHC class IIβ genotypes in song sparrows. The possibility that geographic variation in song might advertise geographic variation at MHC suggests an important potential mechanism by which songbirds may assess the genetic quality (including compatibility) of potential mates.

1.5.2 Signals of MHC diversity

Whereas most research into MHC-mediated mate choice has focused on signals of MHC dissimilarity, there is also substantial evidence in several species that individuals prefer mates with a particular level of MHC diversity (reviewed by Ruff et al. 2012). By choosing a mate that is optimally diverse at MHC, the choosier sex may accrue direct (material)

benefits, such as nest quality (Jäger et al. 2007), and offspring care (Knafler et al. 2012) since their mate would be more resistant to parasites, thus allowing them to allocate energy to provisioning offspring, in comparison to less MHC-diverse individuals. Importantly, this does not exclude the possibility that the choosier sex may also accrue indirect genetic benefits when selecting mates that are optimally diverse at MHC. For example, preferences for optimally MHC-diverse mates may increase the likelihood of obtaining locally-adapted alleles for one's offspring, potentially providing these offspring an advantage in defending against common pathogens (Milinski 2016).

Sexually selected signals of MHC diversity have been documented in most vertebrate groups. In mammals, antler size and body size of male white-tailed deer (*Odocoileus virginianus*) were positively associated with MHC diversity (Ditchkoff et al. 2001); MHC-diverse individuals were also more resistant to nematode infection. In fish, intensity of the sexually-selected red coloration of male three-spined sticklebacks was positively associated with diversity at MHC class I, whereas males with the highest quality nests had intermediate, rather than maximal, diversity at MHC class II (Jäger et al. 2007). In birds, train length of male peacocks varied positively with MHC class II diversity (Hale et al. 2009). Similarly, in common yellowthroats the expression of sexually selected plumage traits such as mask size (Wisconsin population, Dunn et al. 2013), and yellow bib brightness (New York population, Whittingham et al. 2015) varies positively with MHC class II diversity. Collectively, these studies suggest that sexually selected visual ornaments may indirectly signal the level of MHC diversity in jawed vertebrates. However, the relationship between a sexually selected acoustic ornament, such as birdsong, has yet to be explored. Song sparrows are an excellent candidate organism to explore the relationship between birdsong repertoire size and MHC diversity. Females prefer males with complex songs (Searcy 1984), and more MHC-diverse males may sing more songs as they would have better immunity during early life when they learned their songs. Therefore, in Chapter 4, I examine the relationship between song repertoire size and MHC class IIβ diversity in song sparrows.

1.6 Study species

The song sparrow is the model organism used in this thesis to study additive and non-additive effects at MHC on fitness and mate choice, and signals of MHC. Song sparrows are oscine songbirds that are widespread across North America (Arcese et al. 2002). Most of the research presented in my thesis was conducted on a seasonally migratory population of song sparrows breeding near Newboro, Ontario (44.6338° N, 76.3308° W). Relatively few females in this population produce extra-pair young (~20% of nests containing any extra-pair young, Potvin and MacDougall-Shackleton 2009), making patterns of social mate choice a good proxy for genetic mate choice. Identifying and pairing with high-quality mates is essential in this species, as direct (material) benefits in the form of paternal care are critical for offspring survival. These benefits include territory quality, mate guarding, nest defence, and provisioning to offspring in the nest and during the fledgling period (Arcese et al. 2002). Male song sparrows learn their songs during the first year of life (Arcese et al. 2002), such that early-life condition is critical to song development (Schmidt et al. 2013). Males vary substantially in the number of songs they can produce: repertoire size ranges from five to 12 songs in this population and is the best supported sexually selected trait for this species, as females prefer to mate with males the largest song repertoires (Searcy 1984).

Song sparrows in the study population interact with a diversity of pathogens, including but not limited to haemosporidian parasites such as *Plasmodium* spp. (Sarquis-Adamson and MacDougall-Shackleton 2016). Sarquis-Adamson and MacDougall-Shackleton (2016) found evidence for a home-field advantage such that song sparrows are more resistant to local *Plasmodium* strains and relatively susceptible to nonlocal strains. This pattern might suggest locally-good genes effects at immune loci, such as MHC. Alternatively, song sparrows may also exhibit compatible-gene effects at MHC, as the immune function decreases with inbreeding level for this species (Reid et al. 2003). Furthermore, in my population, females adjust parental effort based on neutral-locus similarity to their social mate (Potvin and MacDougall-Shackleton 2009). This finding

suggests that song sparrows may be capable of assessing genetic similarity, potentially mediated by MHC cues.

1.7 Dissertation structure

My thesis contains four data chapters, collectively addressing the over-arching hypothesis that MHC affects disease resistance of song sparrows through additive and non-additive genetic effects; that the chemical composition of preen oil and the content of learned song repertoires provide information regarding MHC genotype dissimilarity and diversity; and that song sparrows use this information to pair disassortatively with respect to MHC genotype.

Chapter 2 examines the role of MHC class I and class IIβ in song sparrows' resistance to local and nonlocal strains of *Plasmodium*, previously studied by Sarquis-Adamson and MacDougall-Shackleton (2016). I examined whether locally-good gene effects at MHC class I exon 3 and/or IIβ explained the apparent local adaptation of two song sparrow populations ~ 440 km apart. Additionally, I explored if heterozygote advantage at either MHC class also explained their resistance to *Plasmodium*.

Chapter 3 and 4 explore two candidate signals of MHC in songbirds. In Chapter 3, I explored whether the chemical composition of preen oil covaried with MHC class IIβ genotypes, and which if any specific compounds explained the majority of variation at MHC. In Chapter 4, I explored whether birdsong composition covaried with MHC class IIβ genotypes. Because male song sparrows learn their song at their population of origin, and there may be geographic variation in MHC, I explored whether birdsong reliably signalled MHC genotype. I also explored whether birdsong complexity (repertoire size) might convey information regarding MHC class IIβ diversity.

In Chapter 5, I investigated whether free-living song sparrows show evidence of non-random mating at MHC class IIβ. Furthermore, because codons subject to positive selection might be particularly salient to mate choice, I screened for particular codons at MHC class IIβ that show evidence of past positive selection. I compared patterns of social

mate choice observed in the field to those expected under random mating, to explore whether song sparrows pair non-randomly based on evolutionary or functional distance at the entire exon, and/or at the subset of codons showing evidence of positive selection.

Chapter 6 integrates the findings from the four data chapters and assesses progress towards the goal of testing whether pathogen-mediated selection favours certain MHC alleles or combination of alleles, what mechanisms might allow individuals to assess the MHC of a potential mate, and whether disassortative mating at MHC occurs in this study population, together with identifying promising directions for future research. Collectively, this dissertation tests a leading candidate locus hypothesized to underlie previous findings of enhanced resistance to local parasite strains (chapter 2), identifies two new candidate modalities by which songbirds may assess MHC dissimilarity and diversity of potential mates (chapter 3 and 4), identifies signals of molecular selection, and provides unexpected findings regarding the nature of MHC-mediated mating in a wild population of songbirds.

1.8 References

Andersson, M. 1986. Evolution of condition-dependent sex ornaments and mating preferences: sexual selection based on viability differences. *Evolution* 40:804–816.

Andersson M, 1994. *Sexual selection*. Princeton University Press.

Andersson M, Simmons LW. 2006. Sexual selection and mate choice. *Trends Ecol Evol.* 21:296–302.

Agbali M, Reichard M, Bryjová A, Bryja J, Smith C. 2010. Mate choice for nonadditive genetic benefits correlate with MHC dissimilarity in the rose bitterling (*Rhodeus ocellatus*). *Evolution*. 64:1683–1696.

Apanius V, Penn DJ, Slev PR, Ruff LR, Potts WK. 1997. The nature of selection on the major histocompatibility complex. *Crit Rev Immunol.* 17:179–224.

Arcese P, Sogge MK, Marr AB, Patten AP. 2002. Song Sparrow (*Melospiza melodia*), version 2.0. In The Birds of North America (Rodewald PG, editor). Cornell Lab of Ornithology, Ithaca, New York, USA.

Baker M, Nabors KJS, Bradley DC. 1981. Early experience determines song dialect responsiveness of female sparrows. *Science.* 214:819–821.

Bang BG, Cobb S. 1968. The size of the olfactory bulb in 108 species of birds. *The Au*k. 85:55–61.

Bernatchez L, Landry C. 2003. MHC studies in nonmodel vertebrates: what have we learned about natural selection in 15 years? *J Evol Biol.* 16:363–377.

Bichet C, Moodley Y, Penn DJ, Sorci G, Garnier S. 2015. Genetic structure in insular and mainland populations of house sparrows (*Passer domesticus*) and their hemosporidian parasites. *Ecol Evol*. 5:1639–1652.

Binz T, Largiader C, Müller R, Wedekind C. 2001. Sequence diversity of *Mhc* genes in lake whitefish. *J Fish Biol.* 58:359–373.

Boehm T, Zufall F. 2006. MHC peptides and the sensory evaluation of genotype. *Trends Neurosci*. 29:100–107.

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

Bonneaud C, Chastel O, Federici P, Westerdahl H, Sorci G. 2006. Complex *Mhc*-based mate choice in a wild passerine. *Proc R Soc B*. 273:1111–1116.

Bos DH, Williams RN, Gopurenko D, Bulut Z, Dewoody JA. 2009. Condition-dependent mate choice and a reproductive disadvantage for MHC-divergent male tiger salamanders. *Mol Ecol*. 18:3307–3315.

Bottazzo GF, Todd I, Mirakian R, Belfiore A, Pujol-Borrell R. 1986. Organ-specific autoimmunity: a 1986 overview. *Immunol Rev*. 94:137–169.

Broggi J, Hohtola E, Orell M, Nilsson J-Å, Benkman C. 2005. Local adaptation to winter conditions in a passerine spreading north: a common-garden approach. *Evolution.* 59:1600–1603.

Brown JL. 1997. A theory of mate choice based on heterozygosity. *Behav Ecol.* 8:60–65.

Buchanan KL, Catchpole CK. 2000. Song as an indicator of male parental effort in the sedge warbler. *Proc R Soc Lond B*. 267:321–326.

Clarke B, Kirby DRS. 1966. Maintenance of histocompatibility polymorphisms. *Nature.* 211:999–1000.

Clutton-Brock T. 2007. Sexual selection in males and females. *Science.* 318:1882–1885.

Consuegra S, Garcia de Leaniz C. 2008. MHC-mediated mate choice increases parasite resistance in salmon. *Proc R Soc B.* 275:1397–1403.

Ditchkoff SS, Lochmiller RL, Masters RE, Hoofer SR, Van Den Bussche RA. 2001. Major-histocompatibility-complex-associated variation in secondary sexual traits of white-tailed deer (*Odocoileus virginianus*): evidence for good-genes advertisement. *Evol Int J Org Evol*. 55:616–625.

Dunn PO, Bollmer JL, Freeman-Gallant CL, Whittingham LA. 2013. MHC variation is related to a sexually selected ornament, survival, and parasite resistance in common yellowthroats. *Evolution.* 67:679–687.

Ekblom R, Saether S, Grahn M, Fiske P, Kalas JA, Höglund J. 2004. Major histocompatibility complex variation and mate choice in a lekking bird, the great snipe (*Gallinago media*). *Mol Ecol*. 13:3821–8.

Emlen ST, Oring LW. 1977. Ecology, sexual selection, and the evolution of mating systems. *Science.* 197:215–223.

Evans ML, Neff BD. 2009. Major histocompatibility complex heterozygote advantage and widespread bacterial infections in populations of Chinook salmon (*Oncorhynchus tshawytscha*). *Mol Ecol*. 18:4716–4729.

Forsberg LA, Dannewitz J, Petersson E, Grahn M. 2007. Influence of genetic dissimilarity in the reproductive success and mate choice of brown trout – females fishing for optimal MHC dissimilarity. *J Evol Biol*. 20:1859–1869.

Galaverni M, Caniglia R, Milanesi P, Lapalombella S, Fabbri E, Randi E. 2015. Choosy wolves? Heterozygote advantage but no evidence of MHC-based disassortative mating. *J Hered*. 1–9.

Gillingham MAF, Richardson DS, Løvlie H, Moynihan A, Worley K, Pizzari T. 2009. Cryptic preference for MHC-dissimilar females in male red junglefowl, *Gallus gallus*. *Proc R Soc B.* 276:1083–1092.

Gomez-Mestre I, Tejedo M, Merilä J. 2003. Local adaptation of an anuran amphibian to osmotically stressful environments. *Evolution.* 57:1889–1899.

Grant BR, Grant PR. 1996. Cultural inheritance of song and its role in the evolution of Darwin's finches. *Evolution.* 50:2471.

Hale ML, Verduijn MH, Møller AP, Wolff K, Petrie M. 2009. Is the peacock's train an honest signal of genetic quality at the major histocompatibility complex? *J Evol Biol*. 22:1284–1294.

Hamilton WD, Zuk M. 1982. Heritable true fitness and bright birds: a role for parasites? *Science.* 218:384–387.

Havlicek J, Roberts SC. 2009. MHC-correlated mate choice in humans: a review. *Psychoneuroendocrinology.* 34:497–512.

Hughes AL, Nei M. 1988. Pattern of nucleotide substitution at major histocompatibility complex class I loci reveals overdominant selection. *Nature.* 335:335167a0.

Jacob J, Ziswiler V. 1982. The uropygial gland. In *Avian Biology,* pp. 199-314. New York, NY: Elsevier.

Jäger I, Eizaguirre C, Griffiths SW, Kalbe M, Krobbach CK, Reusch TBH, 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. *J Evol Biol*. 20:2005–2015.

Jeffery KJ, Bangham CR. 2000. Do infectious diseases drive MHC diversity? *Microbes Infect*. 2:1335–1341.

Johnsen A, Andersen V, Sunding C, Lifjeld JT. 2000. Female bluethroats enhance offspring immunocompetence through extra-pair copulations. *Nature*. 406:296-299.

Keverne EB. 1999. The vomeronasal organ. *Science.* 286:716–720.

Keyser AJ, Hill GE. 2000. Structurally based plumage coloration is an honest signal of quality in male blue grosbeaks. *Behav Ecol*. 11:202–209.

Klein J. 1986. *Natural history of the major histocompatibility complex*. John Wiley & Sons: New York, NY, USA.

Knafler GJ, Clark JA, Boersma PD, Bouzat JL. 2012. MHC diversity and mate choice in the Magellanic penguin, *Spheniscus magellanicus*. *J Hered*. 103:759–768.

Kodric-Brown A, Brown JH. 1984. Truth in advertising: the kinds of traits favored by sexual selection. *Am Nat.* 124:309–323.

Kubinak JL, Nelson CA, Ruff JS, Potts WK. 2012. Tradeoffs limiting MHC heterozygosity. In *Ecoimmunology*, pp. 225-258. Oxford University Press, Inc.: New York, NY, USA.

Kulski JK, 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. *Immunol Rev.* 190:95–122.

Landry C, Garant D, Duchesne P, Bernatchez L. 2001. "Good genes as heterozygosity": the major histocompatibility complex and mate choice in Atlantic salmon (*Salmo salar*). *Proc R Soc Lond B*. 268:1279–1285.

Leclaire S, Merkling T, Mulard RH, Bessiere J-M, Lhuillier E, Hatch SA, Danchin E. 2012. Semiochemical compounds of preen secretion reflect genetic make-up in a seabird species. *Proc R Soc B.* 279:1185–1193.

Leclaire S, Strandh M, Mardon J, Westerdahl H, Bonadonna F. 2017. Odour-based discrimination of similarity at the major histocompatibility complex in birds. *Proc R Soc B.* 284:20162466.

Leclaire S, van Dongen WFD, Voccia S, Merkling T, Ducamp C, Hatch SA, Blanchard P, Danchin É, Wagner RH. 2014. Preen secretions encode information on MHC similarity in certain sex-dyads in a monogamous seabird. *Sci Rep*. 4:6920.

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*). *Mol Ecol*. 18:1331–1340.

MacDougall-Shackleton EA, Derryberry EP, Foufopoulos J, Dobson AP, Hahn TP. 2005. Parasite-mediated heterozygote advantage in an outbred songbird population. *Biol Lett*. 1:105–107.

Marler P, Peters S. 1987. A sensitive period for song acquisition in the song sparrow, *Melospiza melodia*: a case of age-limited learning. *Ethology.* 76:89–100.

Marr AB, Keller LF, Arcese P. 2002. Heterosis and outbreeding depression in descendants of natural immigrants to an inbred population of song sparrows (*Melospiza melodia*). *Evolution.* 56:131–142.

Mays HL, Hill GE. 2004. Choosing mates: good genes versus genes that are a good fit. *Trends Ecol. Evol*. 19:554–559.

McClelland EE, Penn DJ, Potts WK. 2003. Major histocompatibility complex heterozygote superiority during coinfection. *Infect Immun*. 71:2079–2086.

Milinski M. 2016. Mate choice optimizes offspring MHC genetics and drives sexual reproduction. *Immunogenet Open Access.* 1:105.

Milinski M. 2006. The major histocompatibility complex, sexual selection, and mate choice. *Annu Rev Ecol Evol Syst*. 37:159–186.

Milinski M, Croy I, Hummel T, Boehm T. 2013. Major histocompatibility complex peptide ligands as olfactory cues in human body odour assessment. *Proc R Soc B.* 280:20122889.

Milinski M, Griffiths S, Wegner KM, Reusch TBH, Haas-Assenbaum A, Boehm T. 2005. Mate choice decisions of stickleback females predictably modified by MHC peptide ligands. *Proc Natl Acad Sci USA*. 102:4414–4418.

Miller HC, Moore JA, Nelson NJ, Daugherty CH. 2009. Influence of major histocompatibility complex genotype on mating success in a free-ranging reptile population. *Proc R Soc B.* 276:1695–1704.

Müller-Schwarze D. 2006. *Chemical ecology of vertebrates*. Cambridge University Press: New York, NY, USA.

Neff BD, Garner SR, Heath JW, Heath DD. 2008. The MHC and non-random mating in a captive population of Chinook salmon. *Heredity.* 101:175–185.

Neff BD, Pitcher TE. 2005. Genetic quality and sexual selection: an integrated framework for good genes and compatible genes. *Mol Ecol*. 14:19–38.

Nevitt GA. 2000. Olfactory foraging by Antarctic Procellariiform seabirds: life at high Reynolds numbers. *Biol Bull*. 198:245.

Niskanen AK, Kennedy LJ, Ruokonen M, Kojola I, Lohi H, Isomursu M, Jansson E, Pyhajarvi T, Aspi J. 2014. Balancing selection and heterozygote advantage in major histocompatibility complex loci of the bottlenecked Finnish wolf population. *Mol Ecol*. 23:875–889.

O'Connor EA, Strandh M, Hasselquist D, Nilsson J-Å, Westerdahl H. 2016. The evolution of highly variable immunity genes across a passerine bird radiation. *Mol Ecol*. 25:977–989.

Oliver MK, Telfer S, Piertney SB. 2009. Major histocompatibility complex (MHC) heterozygote superiority to natural multi-parasite infections in the water vole (*Arvicola terrestris*). *Proc R Soc B.* 276:1119–1128.

Olsson M, Madsen T, Nordby J, Wapstra E, Ujvari B, Wittsell H. 2003. Major histocompatibility complex and mate choice in sand lizards. *Proc R Soc Lond B* (Suppl.). 270:S254–S256.

Pagenkopp KM, Klicka J, Durrant KL, Garvin JC, Fleischer RC. 2008. Geographic variation in malarial parasite lineages in the common yellowthroat (*Geothlypis trichas*). *Conserv Genet*. 9:1577–1588.

Penn DJ, Damjanovich K, Potts WK. 2002. MHC heterozygosity confers a selective advantage against multiple-strain infections. *Proc Natl Acad Sci.* 99:11260–11264.

Piertney SB, Oliver MK. 2006. The evolutionary ecology of the major histocompatibility complex. *Heredity.* 96:7–21.

Podos J, Warren PS. 2007. The evolution of geographic variation in birdsong. *Adv Study Behav*. 37:403–458.

Potts WK, Wakeland EK. 1990. Evolution of diversity at the major histocompatibility complex. *Trends Ecol Evol.* 5:181–187.

Potvin DS, MacDougall-Shackleton EA. 2009. Parental investment amplifies effects of genetic complementarity on growth rates in song sparrows, *Melospiza melodia*. *Anim Behav*. 78:943–948.

Radwan J, Tkacz A, Kloch A. 2008. MHC and preferences for male odour in the bank vole. *Ethology.* 114:827–833.

Rajakaruna RS, Brown JA, Kaukinen KH, Miller KM. 2006. Major histocompatibility complex and kin discrimination in Atlantic salmon and brook trout. *Mol Ecol*. 15:4569–4575.

Reid JM, Arcese P, Keller LF. 2003. Inbreeding depresses immune response in song sparrows (*Melospiza melodia*): direct and inter–generational effects. *Proc R Soc Lond B Biol Sci.* 270:2151–2157.

Reinhold K. 2004. Modeling a version of the good-genes hypothesis: female choice of locally adapted males. *Org Divers Evol*. 4:157–163.

Reusch TBH, Haberli MA, Aeschlimann PB, Milinski M. 2001. Female sticklebacks count alleles in a strategy of sexual selection explaining MHC polymorphism. *Nature.* 414:300–2.

Richardson DS, Komdeur J, Burke T, von Schantz T. 2005. MHC-based patterns of social and extra-pair mate choice in the Seychelles warbler. *Proc R Soc B*. 272:759–767.

Robinson J, Halliwell JA, McWilliam H, Lopez R, Parham P, Marsh SGE. 2013. The IMGT/HLA database. *Nucleic Acids Res*. 41:D1222–D1227.

Ruff JS, Nelson AC, Kubinak JL, Potts WK. 2012. MHC signaling during social communication. In *Self and Nonself*, pp 290–313. Advances in Experimental Medicine and Biology Vol 738. Springer Science & Business Media: Berlin, Germany.

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

Schmidt KL., Moore SD, MacDougall-Shackleton EA, MacDougall-Shackleton SA. 2013. Early-life stress affects song complexity, song learning and volume of the brain nucleus RA in adult male song sparrows. *Anim Behav*. 86:25–35.

Schwensow N, Eberle M, Sommer S. 2008a. Compatibility counts: MHC-associated mate choice in a wild promiscuous primate. *Proc R Soc B*. 275:555–564.

Schwensow N, Fietz J, Dausmann K, Sommer S. 2008b. MHC-associated mating strategies and the importance of overall genetic diversity in an obligate pair-living primate. *Evol Ecol*. 22:617–636.

Searcy WA. 1984. Song repertoire size and female preferences in song sparrows. *Behav Ecol Sociobiol*. 14:281–286.

Searcy WA, Nowicki S, Hughes M, Peters S. 2002. Geographic song discrimination in relation to dispersal distances in song sparrows. *Am Nat*. 159:221–230.

Setchell JM, Charpentier MJE, Abbott KM, Wickings EJ, Knapp LA. 2010. Opposites attract: MHC-associated mate choice in a polygynous primate. *J Evol Biol*. 23:136–148.

Siefferman L, Hill GE. 2003. Structural and melanin coloration indicate parental effort and reproductive success in male eastern bluebirds. *Behav Ecol*. 14:855–861.

Sin YG, Annavi G, Newman C, Buesching C, Burke T, MacDonald DW, Dugdale HL. 2015. MHC class II-assortative mate choice in European badgers (*Meles meles*). *Mol Ecol*. 24:3138–3150.

Slade RW, McCallum HI. 1992. Overdominant vs. frequency-dependent selection at MHC loci. *Genetics*. 132:861–862.

Sommer S. 2005. Major histocompatibility complex and mate choice in a monogamous rodent. *Behav Ecol Sociobiol*. 58:181–189.

Spehr M, Kelliher KR, Li X-H, Boehm T, Leinders-Zufall T, Zufall F. 2006. Essential role of the main olfactory system in social recognition of major histocompatibility complex peptide ligands. *J Neurosci*. 26:1961–1970.

Stewart KA, MacDougall-Shackleton EA. 2008. Local song elements indicate local genotypes and predict physiological condition in song sparrows *Melospiza melodia*. *Biol Lett*. 4:240–242.

Sutton JT, Nakagawa S, Robertson BC, Jamieson IG. 2011. Disentangling the roles of natural selection and genetic drift in shaping variation at MHC immunity genes. *Mol Ecol*. 20:4408–4420.

Thomas RH, Price ER, Seewagen CL, Mackenzie SA, Bernards MA, Guglielmo CG. 2010. Use of TLC-FID and GC-MS/FID to examine the effects of migratory state, diet and

captivity on preen wax composition in White-throated Sparrows *Zonotrichia albicollis*. *Ibis.* 152:782–792.

Thorpe RS, Reardon JT, Malhotra A. 2005. Common garden and natural selection experiments support ecotypic differentiation in the Dominican anole (*Anolis oculatus*). *Am Nat.* 165:495–504.

Tregenza T, Wedell N. 2000. Genetic compatibility, mate choice and patterns of parentage: invited review. *Mol Ecol.* 9:1013–27.

Trowsdale J. 2011. The MHC, disease and selection. *Immunol Lett*. 137:1–8.

Tuttle EM, Sebastian PJ, Posto AL, Soini HA, Novotny MV, Gonser RA. 2014. Variation in preen oil composition pertaining to season, sex, and genotype in the polymorphic white-throated sparrow. *J Chem Ecol.* 40:1025–1038.

von Schantz T, Wittzell H, Göransson G, Grahn M. 1997. Mate choice, male condition-dependent ornamentation and MHC in the pheasant. *Hereditas.* 127:133–140.

Wedekind C, Seebeck T, Bettens F, Paepke AJ. 1995. MHC-dependent mate preferences in humans. *Proc R Soc Lond B.* 260:245–249.

Wegner KM, Kalbe M, Kurtz J, Reusch TBH, Milinski M. 2003. Parasite selection for immunogenetic optimality. *Science.* 301:1343.

Wegner KM, Kalbe M, Schaschl H, Reusch TBH. 2004. Parasites and individual major histocompatibility complex diversity—an optimal choice? *Microbes Infect*. 6:1110–1116.

Whittaker DJ, Richmond KM, Miller AK, Kiley R, Burns CB, Atwell JW, Ketterson ED. 2011. Intraspecific preen oil odor preferences in dark-eyed juncos (*Junco hyemalis*). *Behav Ecol*. 22:1256–1263.

Whittaker DJ, Soini HA, Atwell JW, Hollars C, Novotny MV, Ketterson ED. 2010. Songbird chemosignals: volatile compounds in preen gland secretions vary among individuals, sexes, and populations. *Behav Ecol*. 21:608–614.

Whittingham LA, Freeman-Gallant CR, Taff CC, Dunn PO. 2015. Different ornaments signal male health and MHC variation in two populations of a warbler. *Mol Ecol*. 24:1584–1595.

Yamazaki K, Boyse EA, Miké V, Thaler HT, Mathieson BJ, Abbott J, Boyse J, Zayas ZA, Thomas L. 1976. Control of mating preferences in mice by genes in the major histocompatibility complex. *J Exp Med*. 144:1324–1335.

Yang Z, Swanson WJ. 2002. Codon-substitution models to detect adaptive evolution that account for heterogeneous selective pressures among site classes. *Mol Biol Evol*. 19:49–57.

Yewdell JW, Bennink JR. 1992. Cell biology of antigen processing and presentation to major histocompatibility complex class I molecule-restricted T lymphocytes. *Adv Immunol*. 52:1–123.

Yewdell JW, Bennink JR. 1999. Immunodominance in major histocompatibility complex class I–restricted T lymphocyte responses. *Annu Rev Immunol*. 17:51–88.

Zahavi A. 1975. Mate selection—a selection for a handicap. *J Theor Biol*. 53:205–214.

Zuk M, Thornhill R, Ligon JD, Johnson K, Austad S, Ligon SH, Thornhill NW, Costin C. 1990. The role of male ornaments and courtship behavior in female mate choice of red jungle fowl. *Am Nat*. 136:459–473.

Chapter 2

2 Population differences at MHC do not explain enhanced resistance of song sparrows to local parasites[1]

2.1 Introduction

Infectious disease constitutes an emerging threat to natural populations. Geographic ranges of parasites are expected to shift in response to human activity and changes in climate (Harvell et al. 2002; McCallum et al. 2002), intensifying the risk of invasive new parasites spreading to naïve host populations (Altizer et al. 2011; Garamszegi 2011). For example, malarial parasites (*Plasmodium* spp.) have recently been detected in hatch-year and non-migratory birds as far north as Alaska (Loiseau et al. 2012), demonstrating that malaria is now transmitted at this latitude. The introduction of novel parasites is particularly concerning in systems in which hosts are more vulnerable to infection by unfamiliar (allopatric) parasites than they are to local (sympatric) parasites. The devastating effects of avian malaria (*P. relictum*) and poxvirus on immunologically naïve birds in the Hawaiian and Galápagos Islands represent particularly dramatic examples of susceptibility to novel parasites (Warner 1968; Zylberberg et al. 2012), but similar patterns have been found across a variety of taxa (Greischar and Koskella 2007).

A growing body of evidence now supports the idea that in many systems, hosts are more likely to be infected by or suffer more severe infections from allopatric than sympatric parasites. However, the mechanisms that might contribute to this pattern remain uncertain. Field studies of wild birds have found greater prevalence and/or intensity of infection in dispersing than in philopatric individuals (white-crowned sparrows *Zonotrichia leucophrys*, MacDougall-Shackleton et al. 2002; song sparrows *Melospiza*

---

[1] A version of this chapter has been published and is presented with permission from the *Journal of Heredity*.

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

melodia, Stewart and MacDougall-Shackleton 2008; barn swallows *Hirundo rustica*, Saino et al. 2014). These associations are consistent with hosts being more susceptible to allopatric than sympatric parasites, but cannot rule out the alternative explanation that dispersing versus philopatric individuals may differ in inherent quality. Reciprocal infection experiments, in which host individuals from two or more sites are exposed to sympatric versus allopatric parasites, have the advantage of controlling for individual variation in quality and can more directly address whether hosts are more susceptible to allopatric parasites. For example, Canarian lizards (*Gallotia galloti*) are relatively resistant to sympatric and susceptible to allopatric haemogregarine parasites (Oppliger et al. 1999), with similar patterns reported for interactions between three-spined sticklebacks *Gasterosteus aculeatus* and eye flukes *Diplostomum pseudospathaceum* (Kalbe and Kurtz 2006); great tits *Parus major* and *Plasmodium* spp. (Jenkins et al. 2015); and song sparrows and *Plasmodium* spp. (Sarquis-Adamson and MacDougall-Shackleton 2016).

At least two mechanisms may contribute to hosts being more susceptible to allopatric than sympatric parasites. First, host populations may differ in allele frequencies at immune loci, due to historical coevolution with local parasite strains increasing the frequency of locally-protective alleles within each population (i.e., local adaptation of hosts to parasites; Loiseau et al. 2009). Second, host individuals may have previously encountered sympatric parasite strains, resulting in antigen-specific immune memory (Krzych et al. 2014) such that re-infections are less likely and/or less severe. Local adaptation of host populations and immune memory of host individuals are not mutually exclusive explanations for why hosts could be more susceptible to allopatric than sympatric parasites. However, evaluating their relative importance is valuable because it informs as to whether this pattern arises primarily over an evolutionary or an ecological time-scale.

In jawed vertebrate animals, genes of the major histocompatibility complex (MHC) encode cell-surface proteins that discriminate self from non-self antigens (i.e., pathogens), and present them to T-cell receptors to initiate an immune response (Klein 1986). Two major classes of MHC molecules are involved in antigen presentation: class I molecules interact primarily with intracellular pathogens, and class II molecules primarily with

extracellular pathogens. The MHC thus represents a key component of immune defence (Trowsdale 1995). MHC loci are compelling candidates at which host populations may show genetic differentiation, because geographic variation in parasite communities (Pagenkopp et al. 2008) may generate variation in selection pressures on hosts (Kaltz and Shykoff 1998). Thus, local variation in parasite-mediated selection pressures could induce population differentiation at MHC (Loiseau et al. 2009), which is a requirement for locally protective alleles contributing to enhanced resistance to sympatric parasites. Alternatively, however, balancing selection at MHC loci may reduce population differentiation. Immigrants bearing locally-rare alleles might be favoured by negative frequency-dependent selection (Schierup et al. 2000, Muirhead et al. 2001, Fraser et al. 2010) or the offspring of such immigrants may benefit from heterozygote advantage (Penn et al. 2002). If populations are not genetically differentiated at MHC, local adaptation at MHC cannot explain enhanced resistance to sympatric parasites and susceptibility to allopatric parasites. This would suggest that host individuals' previous experience with local parasites, rather than population-level adaptation to the local parasites, is the main contributor to host resistance to sympatric parasite strains.

In a recent study, Sarquis-Adamson and MacDougall-Shackleton (2016) inoculated wild-caught adult song sparrows, captured from two sites 440 km apart, with either sympatric or allopatric lineages of *Plasmodium*. Infection risk was lower for birds exposed to sympatric as opposed to allopatric parasites, a pattern which might reflect population differentiation and locally protective alleles at immune-related loci such as MHC, or some other explanation such as previous exposure and immune memory. In the current study, I characterize MHC (class I and II) of birds from each source population, in order to assess population genetic differentiation and thus, the potential that locally protective alleles might contribute to enhanced resistance to sympatric parasites. I also investigate individual diversity at MHC of subjects used in the cross-infection experiment, and relate this to infection success in order to assess evidence for balancing selection (heterozygote advantage) at these loci.

2.2 Materials and Methods

2.2.1 Population sampling and infectivity experiment

I characterized MHC genotypes of 36 song sparrows previously used in a cross-infection experiment (Sarquis-Adamson and MacDougall-Shackleton 2016). Subjects in the cross-infection study were captured at their summer breeding grounds in Ontario, Canada during late summer and early fall (July-October). Nineteen birds were captured at a site in eastern Ontario (44°38'38.77"N, 76°20'4.86"W), and 17 at a site 440 km away in western Ontario (43°00'34.00"N, 81°16'52.50"W). Song sparrows are abundant in both these locations, and neither site was physically isolated from species-suitable breeding habitat (old fields and wetlands), thus I assume that neither population is currently isolated from other surrounding populations. Lab members collected a small blood sample from each bird's brachial vein for genetic analysis, and then housed the birds indoors in individual cages with *ad libitum* access to food and water under ambient photoperiod.

Details of the cross-infection experiment are available in Sarquis-Adamson and MacDougall-Shackleton (2016), but in brief, they used nested PCR to amplify and sequence haematozoan mitochondrial DNA (Hellgren et al. 2004) and identified lineages of *Plasmodium* that were apparently confined only to the eastern or only to the western site. Lineage P-SOSP9 (GenBank accession # KT19635; 99% sequence identity to morphospecies *P. relictum*) was found only in birds from the eastern site, and lineage P-SOSP10 (GenBank accession # KT19636; 99% sequence identity to morphospecies *P. homopolare*) was found only in birds from the western site. Additional screening of over 300 song sparrows at and around the eastern site confirmed the absence of P-SOSP10 from this location, although comparable screening at and around the western site to confirm the absence of P-SOSP9 was not conducted (Sarquis-Adamson and MacDougall-Shackleton 2016). Although several other lineages of *Plasmodium* were detected, these occurred at both the eastern and western sites. Sarquis-Adamson and MacDougall-Shackleton (2016) thus used P-SOSP9 and P-SOSP10 as the eastern and western lineages, respectively. DNA sequence divergence between these two experimental lineages was 8%. All of the other

lineages detected within experimental subjects upon capture showed at least 5.2% sequence divergence to both experimental lineages (GenBank accession # KT193627-193634 and KT19637; Sarquis-Adamson and MacDougall-Shackleton 2016).

Lab members inoculated a suspension of whole blood from an eastern bird infected with P-SOSP9 into two previously-uninfected "amplifiers" (i.e., birds inoculated with parasites, allowed to develop an acute infection, then used to inoculate experimental subjects) from the eastern site. Similarly, two previously-uninfected amplifiers from the western site were inoculated with a suspension of whole blood from a western bird infected with P-SOSP10 (details in Sarquis-Adamson and MacDougall-Shackleton 2016). At 18 days post-inoculation, all four amplifiers had infectious stages of *Plasmodium* (asexual meronts) detectable in peripheral blood (eastern average = western average = 2.0 parasites per 10,000 erythrocytes, scored by microscopic examination of thin-film blood smears). Blood from amplifiers was then mixed with buffer and used to inoculate birds in the experimental groups, such that all experimental groups received the same dose of *Plasmodium*. Experimental birds comprised four groups of six birds each (i.e., eastern birds inoculated with their sympatric lineage P-SOSP9, eastern birds inoculated with their allopatric lineage P-SOSP10, western birds inoculated with their allopatric lineage P-SOSP9, western birds inoculated with their sympatric lineage P-SOSP10).

To assess infection success, Sarquis-Adamson and MacDougall-Shackleton (2016) monitored parasitemia in the 24 experimental birds every three days, from six to 30 days post-inoculation. 20 μL of blood was collected via brachial venipuncture and used to prepare a thin-film blood smear. Smears were air-dried, fixed in absolute methanol, and treated with Wright-Giemsa stain, after which researchers examined 10,000 red blood cells per smear and noted the number of cells containing one or more *Plasmodium*. Birds with at least one observation of at least four infected erythrocytes per 10,000 examined (0.04% parasitemia) were considered to have been successfully infected (Sarquis-Adamson and MacDougall-Shackleton 2016). However, because they did not confirm via PCR that 'successful infections' involved the lineages of interest, they could not conclusively exclude the possibility that some may have involved coinfections by multiple lineages or

recurrences of earlier infections. All experimental birds survived to the 30-day endpoint of the experiment, and 15 of 24 (63%) became successfully-infected as defined above.

2.2.2 Characterizing MHC

Passerine birds have undergone extensive gene duplications at MHC (Westerdahl et al. 2000; Hess and Edwards 2002), and this combined with high heterozygosity expected at the peptide-binding region (PBR) precluded direct sequencing of class I and II loci. I characterized most of class I, exon 3 by cloning and sequencing, and characterized class II, exon 2 by next-generation sequencing.

2.2.2.1 Class I

I used polymerase chain reaction (PCR) to amplify most of exon 3 of MHC class I, a region encoding the α2 domain of this molecule's highly variable peptide-binding region (PBR). I amplified a 213 bp fragment of exon 3 using primers that bind within this exon (*GCA21M* and *fA23M*; Loiseau et al. 2009). These primers were designed from cDNA sequences, in order to preferentially amplify transcribed alleles (Loiseau et al. 2009). PCR was performed in a total volume of 25 μL, and included 0.2 U of *Taq* polymerase (Life Technologies), 0.2 μM of each primer, 0.2 mM of dNTPs, 1 mM of $MgCl_2$, 1 x PCR buffer (Sigma-Aldrich: 10 mM Tris-HCl, pH 8.3 at 25°C, 50 mM KCl, 0.001% gelatin), and approximately 25 ng of genomic DNA template. The thermocycler profile consisted of 2 min at 94 °C; 32 cycles of 30 s at 94 °C, 30 s at 57 °C, and 30 s at 72 °C; and a final extension step of 10 min at 72 °C.

I used the pGEM-T Easy Vector System (Promega) to ligate PCR products into a vector which I then introduced into competent cells and screened for colonies containing the desired insert. I selected 20 insert-containing colonies per bird, and used a sterile toothpick to transfer each of these colonies into a new PCR reaction. PCR conditions for these second-round amplifications were the same as described above. After PCR, I treated each reaction with 10 U of Exo I and 10 U of FastAP Thermosensitive Alkaline Phosphatase (Thermo Scientific) for 15 min at 37 °C to degrade any remaining primers,

followed by a 15 min treatment at 85 °C to inactivate the enzymes. Each colony was then sequenced using the forward primer *GCA21M*, on an Applied Biosystems 3730 DNA Analyzer at the London Regional Genomics Centre.

I aligned class I sequences in MEGA 6 (Tamura et al. 2013), queried them against the Basic Local Alignment Search Tool (BLAST; Altschul et al. 1990) implemented in GenBank, and confirmed that the most similar sequences retrieved were other class I, exon 3 sequences from passerine birds. Alignments were trimmed to 192 bp, corresponding to codons 11 through 74 of the putative exon 3. I used the *cluster* command in the program mothur 1.33.3 (Schloss et al. 2009) to identify each individual's alleles based on 100% DNA sequence identity. Low-quality reads (i.e., those containing Ns) were discarded. To reduce potential effects of PCR and sequencing errors, I also discarded singleton sequences, i.e., those found only in a single colony. Thus, only clear sequences that were recovered from two or more colonies (either within the same individual or across multiple individuals) were considered to be true alleles. In all, the number of usable sequences per individual was 16.00 ± 0.36 (mean ± s.e.m.). This number was not related to the number of unique alleles detected for an individual (Pearson's $r_{35} = 0.14$, $p = 0.43$).

2.2.2.2 Class II

Due to high levels of duplication and polymorphism of MHC class II found in other passerine birds (e.g., upwards of 20 loci in common yellowthroats *Geothlypis trichas*; Bollmer et al. 2010), I used next-generation sequencing to characterize class II variation within and between populations. With the help of a collaborator, I designed a degenerate forward primer (*SospMHCint1f*: priming sequence 5'-AGY GGG GAY CCG GGG TGG-3'), to bind to intron 1 of class II MHC. I used this in combination with the reverse primer *Int2r.1* (Edwards et al. 1998) binding to intron 2, to amplify all of exon 2, which encodes the hypervariable PBR (β chain) of class II MHC. In addition to the priming sequences, each primer included an adaptor sequence for the Illumina MiSeq platform, four wobble bases, and an individually-unique barcode sequence of eight bases.

PCR was conducted in a total volume of 30 μL, and included 12.5 μL of GoTaq® Hot Start Colorless Master Mix (Promega), 0.2 μM of each primer, and 25 ng of genomic template. The thermocycling profile consisted of 3 min at 94 °C; 28 cycles of 30 s at 94 °C, 30 s at 62 °C, and 45 s at 72 °C; and a final step of 10 min at 72 °C. I confirmed amplification success for each individual by running part of the PCR product on an agarose gel, and then pooled PCR products from all individuals to form a library. This library was used in next-generation sequencing at the London Regional Genomics Centre, using 300 bp paired-end reads on a single flow cell in the Illumina MiSeq platform.

A collaborator used a pipeline developed by Gloor et al. (2010) to extract sequences, sort by individuals, and collapse into clusters of identical reads. To identify very rare sequences likely to have resulted from PCR errors, I amplified class II, exon 2 of MHC using the primers and PCR conditions described above for one individual, and used cloning (Promega pGEM-T Easy Vector System) to generate multiple colonies each containing a single allele. I included PCR products from four colonies in the flow cell run, along with the library described above. In the absence of PCR or sequencing errors, each colony should generate only one individual sequencing unit (ISU), thus the frequency of rare secondary ISUs provides an estimate of error rates. Based on the observed frequency of secondary ISUs averaged across the four colonies, I established a threshold error rate of 1%. Thus, sequences appearing in <1% of an individual's reads were considered to result from PCR or sequencing errors, and were discarded. Retained sequences were checked for similarity to published exon 2 sequences of passerine class II MHC using BLAST as described above.

2.2.3 Population genetic differentiation at MHC

If locally protective alleles at MHC explain enhanced resistance to sympatric parasites (Sarquis-Adamson and MacDougall-Shackleton 2016), I would expect that MHC allele frequencies differ between sites. I investigated population genetic differentiation at MHC class I exon 3 and class II exon 2. Class I genotypes were available for 35 individuals (18 from the eastern site, 17 from the western site) and class II genotypes for 35 individuals (19 eastern, 16 western).

I observed a large number of MHC alleles (27 unique DNA sequences at class I; 192 unique DNA sequences at class II) with relatively little overlap between individuals, and because of this, classical F_{ST} analyses were not possible. Instead, a collaborator and I performed the analyses on distance matrices constructed from reduced data following Lachance et al. (2013). For each class of MHC, I used MEGA 6.0 (Tamura et al. 2013) to align the putative exon sequences (a 192 bp fragment of class I exon 3, and 222 bp corresponding to the entire putative class II exon 2) from all the observed alleles, and constructed neighbour joining trees. Alleles were then assigned to 'superalleles' based on membership in well-defined clades (Appendix A.1, A.2). As a result, the number of alleles detected in the entire sample was reduced from 27 alleles to four superalleles for class I, and from 192 alleles to 76 superalleles for class II. For each individual, I noted the number of alleles belonging to each superallele, and then my collaborator calculated Euclidean distances between all pairwise combinations of individuals based on their superallele genotypes at class I or class II. F_{ST} was calculated from the formula $(D_T - D_S)/D_T$, where D_S is the mean within-site distance and D_T the between-site distance. To evaluate the significance of each F_{ST} value, the coefficient was recalculated for 500 distance matrices constructed from similarly sized demes of randomized membership. The fractional rank of the observed F_{ST} value was taken to be the probability p of a type I error (Lachance et al. 2013).

To further examine population genetic differentiation at MHC, I used Bayesian cluster analysis implemented in *structure* 2.3.4 (Pritchard et al. 2000). The multilocus nature of MHC sequence data meant that I could determine superallele presence vs. absence in each individual, but not homozygous vs. heterozygous state. Thus, I treated MHC superallele profiles as being analogous to AFLP data (presence or absence of a given band) and used settings for dominant markers (Falush et al. 2007). I used the admixture model with correlated allele frequencies, and to maximize power to detect subtle genetic structuring, I included capture location (east or west) as prior information. For each of class I and class II MHC datasets, I compared support for models with and without genetic differentiation between the two sites ($K = 1$, $K = 2$). Running conditions included an initial burn-in of 100 000 iterations, followed by a run length of 100 000 MCMC iterations, after

which all parameters had stabilized. I used ln likelihood scores to calculate posterior probability for each value of *K* (Pritchard and Wen 2003). To assess consistency between trials, for each of the class I and class II MHC datasets I performed five runs for each value of K.

2.2.4 Individual diversity at MHC and infection risk

Of the birds genotyped at both class I and class II, 22 had been experimentally inoculated with *Plasmodium* as part of a previous study (Sarquis-Adamson and MacDougall-Shackleton 2016). Thus, infection outcome was known for these individuals: 14 became successfully infected (as defined by a criterion of 0.04% parasitemia; Sarquis-Adamson and MacDougall-Shackleton 2016) and eight did not. To assess whether individual genetic diversity at MHC influenced infection likelihood, I used the ExPASy translate tool (Gasteiger et al. 2003) to identify the correct reading frame and to convert nucleic acid sequences of MHC class I and class II alleles into the corresponding amino acid sequences. For each of the 22 birds with known infection outcome, I noted the number of unique amino acid sequences at MHC class I and at MHC class II.

I used *glm* in base R 3.2.3 (R Core Team 2015) to construct generalized linear models with binomial error distribution, each with infection outcome (i.e., whether or not the individual developed parasitemia of 0.04% or higher within 30 days of inoculation) as the dependent variable. I used an information theoretic approach (Burnham & Anderson 2002) to compare support for four alternative models (Table 2.1): univariate models of diversity (i.e., number of different amino acid sequences detected within the individual) at MHC class I and class II, a combination model including MHC class I diversity and MHC class II diversity, and a null model (intercept only).

As a complementary analysis, I also tested a broader candidate set of eleven candidate models (Table 2.1) predicting infection outcome. Predictor variables in this broader model set included MHC class I and class II diversity; eastern vs. western bird origin; eastern vs. western parasite origin; bird origin × parasite origin interaction; and previous infection status (i.e., whether or not the individual was determined to have been

naturally-infected with *Plasmodium* prior to capture). The broader candidate model set also included a null model (intercept only). I calculated model-averaged parameter estimates and unbiased 95% confidence intervals from the full set of AICc-ranked candidate models using the functions *model.avg* and *confint* in the R package MuMIn (Bartoń 2015).

2.3 Results

2.3.1 Individual variation at MHC

I detected 37 variable sites within the 192 bp examined for class I exon 3, and 188 variable sites within the 222 bp examined for class II exon 2. I found between 1 and 4 alleles (i.e., different DNA sequences) per individual at class I (mean ± s.e.m., eastern: 2.78 ± 0.22, n = 18; western: 2.24 ± 0.29, n = 17) and between 10 and 26 alleles per individual at class II (eastern: 17.38 ± 0.96, n = 19; western: 18.89 ± 0.89, n = 16).

2.3.2 Population genetic differentiation at MHC

My collaborator and I found no evidence of genetic differentiation between eastern and western song sparrows, for either class I or class II of MHC (class I: D_T = 3.42, D_S = 2.78, F_{ST} = 0.111, p = 0.12; class II: D_T = 32.7, D_S = 31.1, F_{ST} = 0.050, P = 0.61). Consistent with this lack of genetic differentiation, I found little evidence for alleles that were common at one site but rare at the other site (Figure 2.1a and 2.1b for class I and class II, respectively). Two-thirds (18/27) of the class I alleles were private, i.e., restricted to either the eastern or the western birds. However, each of these were observed in only 1-2 individuals (Figure 2.1a), and I observed no private superalleles at MHC class I.

Bayesian (*structure*) analysis of MHC class I differentiation did not conclusively support either K = 1 (panmixia) or K = 2 (two genetic clusters). Two of the five trials identified an optimal K of 1 (posterior probabilities 0.52, 0.84); two identified an optimal K of 2 (posterior probabilities 0.55, 0.87) and one showed identical support (posterior probability = 0.50) for K = 1 and K = 2. By contrast, all five trials for class II identified K = 2 as the most probable (posterior probabilities = 0.98 – 1.00). However, for both class I

and class II datasets, under $K = 2$ models all individuals were highly admixed regardless of capture location (40-60% genetic membership in each cluster; no individuals strongly assigned to either cluster; Figure 2.2a, b), consistent with a lack of true underlying population genetic structure (Pritchard and Wen 2003). In light of the apparent lack of population differences in superallele frequencies at either class of MHC, and because sample sizes were low relative to the diversity of MHC alleles observed, I did not test further for locally protective effects of specific alleles or superalleles.

2.3.3 Individual diversity at MHC and infection risk

Of the four models predicting infection outcome (MHC class I diversity, MHC class II diversity, combination of both MHC class I and MHC class II diversity, and null), the strongest was the MHC class I diversity model (Table 2.2). This model was approximately three times as likely as the next best model (MHCI + MHCII; Table 2.2) and 10-34 times as likely as either model without MHC class I diversity (intercept only, MHCII; Table 2.2). Based on this top model, the effect of MHC I diversity on infection outcome was significantly different from zero (Estimate ± SE = -1.54 ± 0.72, Z = -2.14, p = 0.033, 95% CI = -3.35 − -0.36). Individuals that remained uninfected following experimental exposure to *Plasmodium* had more amino acid alleles at MHC class I (uninfected: 2.63 ± 0.38 alleles, n = 8; infected: 1.57 ± 0.17 alleles, n = *14*) but not at class II (uninfected: 16.86 ± 1.04 alleles, n = 8; infected: 16.38 ± 1.07 alleles, n = 14).

Comparison and averaging of the broader candidate set of models yielded similar results: the best model was that containing only MHC class I diversity (Table 2.1), and model averaging revealed a significant effect of MHC class I diversity but not of other variables tested (Table 2.3).

Table 2.1 AICc-ranked set of candidate models predicting infection success in 22 song sparrows experimentally inoculated with *Plasmodium*. Predictor variables were MHC class I diversity (i.e., number of distinct amino acid sequences at class I; MHCI), MHC class II diversity (MHCII), bird origin (BO), parasite origin (PO), bird origin × parasite origin interaction (BOPO), and previous infection status (PREV). All models were fit with binomial error distribution.

Candidate models	**K**	**logLik**	**AICc**	**ΔAICc**	**Weight**
MHCI	2	-10.86	26.35	0	0.46
MHCI + PREV	3	-10.19	27.70	1.36	0.23
MHCI + MHCII	3	-10.60	28.53	2.19	0.15
MHCI + MHCII + PREV	4	-10.06	30.47	4.13	0.06
INTERCEPT ONLY (null model)	1	-14.42	31.04	4.70	0.04
PREV	2	-14.06	32.75	6.40	0.02
MHCII	2	-14.37	33.37	7.03	0.01
MHCII + PREV	3	-13.79	34.92	8.58	0.01
BO + PO + BOPO	4	-12.76	35.87	9.52	0
MHCI + MHCII + BO + PO + BOPO	6	-9.78	37.17	10.82	0
MHCI + MHCII + BO + PO + BOPO + PREV	7	-9.50	40.99	14.65	0

Table 2.2 AICc-ranked set of candidate models predicting infection success in 22 song sparrows experimentally inoculated with *Plasmodium*. Predictor variables were MHC class I diversity (i.e., number of distinct amino acid sequences at class I; MHCI), and MHC class II diversity (MHCII). All models were fit with binomial error distribution.

Candidate models	**K**	**logLik**	**AICc**	**ΔAICc**	**Weight**	**Evidence ratio**
MHCI	2	-10.86	26.35	0	0.68	--
MHCI + MHCII	3	-10.60	28.53	2.19	0.23	2.96
NULL (intercept only)	1	-14.42	31.04	4.70	0.07	9.71
MHCII	4	-14.37	33.37	7.03	0.02	34

Table 2.3 Model-averaged predictors of infection success in 22 song sparrows experimentally inoculated with *Plasmodium*. Bold type indicates predictor for which 95% confidence interval does not include zero. Model averaging was conducted on all eleven candidate models from Table 2.1, weighted by Akaike weight.

Predictor	Estimate	SE	95% CI (2.5%, 97.5%)
MHC class I diversity	**-1.49**	**0.91**	**-4.79, -0.29**
MHC class II diversity	-0.026	0.11	-0.65, 0.29
Previous infection status (yes)	-0.43	1.00	-4.26, 1.60
Parasite origin (western)	0.0053	0.17	-4.26, 4.03
Bird origin (western)	-0.0051	0.15	-5.39, 1.71
Bird origin (western) × parasite origin (western)	0.0049	0.20	-3.84, 6.46
Intercept	4.33	3.26	-2.28, 23.50

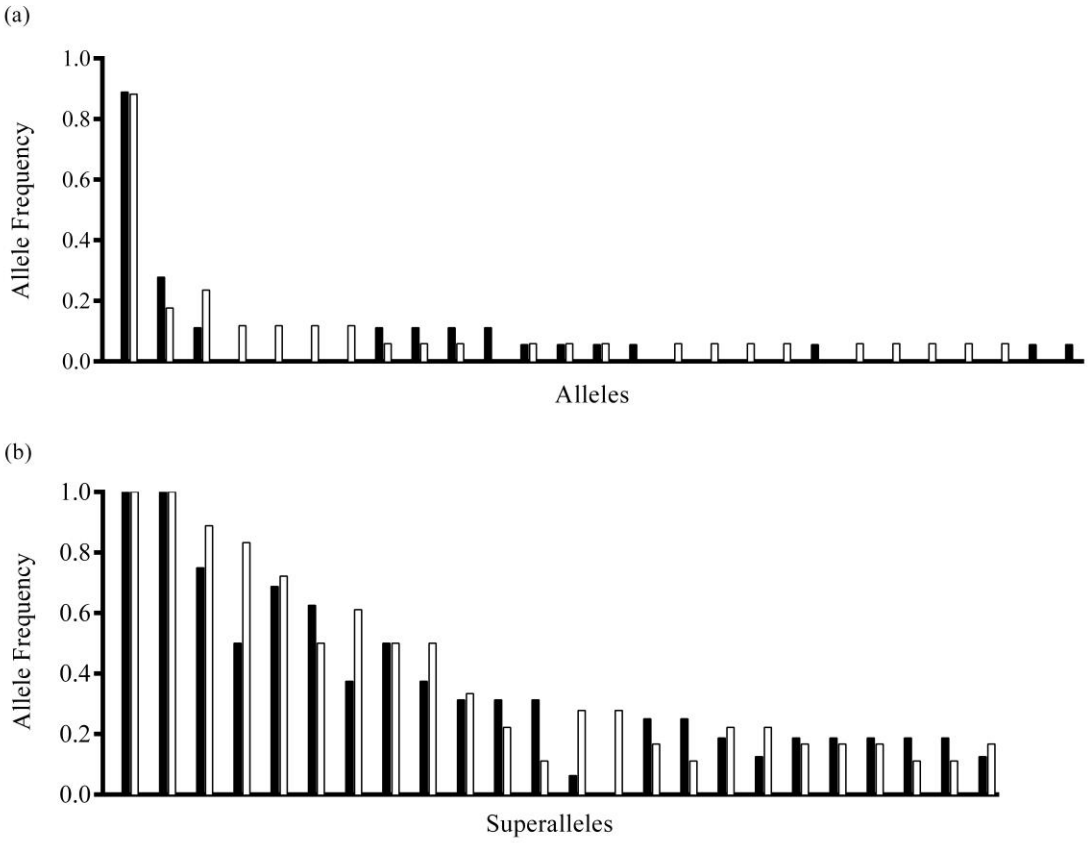


Figure 2.1 Allele frequencies (proportion of individuals in which an allele occurs) for (a) exon 3 of MHC class I, (b) exon 2 of MHC class II. Shaded bars show frequencies at the western site (n = 17 individuals for class 1, 16 for class II), white bars show frequencies at the eastern site (n = 18 individuals for class 1, 19 for class II). Alleles are arranged in decreasing order of frequency. Due to the large number of sequences found at class II, panel (b) presents superalleles rather than alleles, and only those found in five or more individuals.

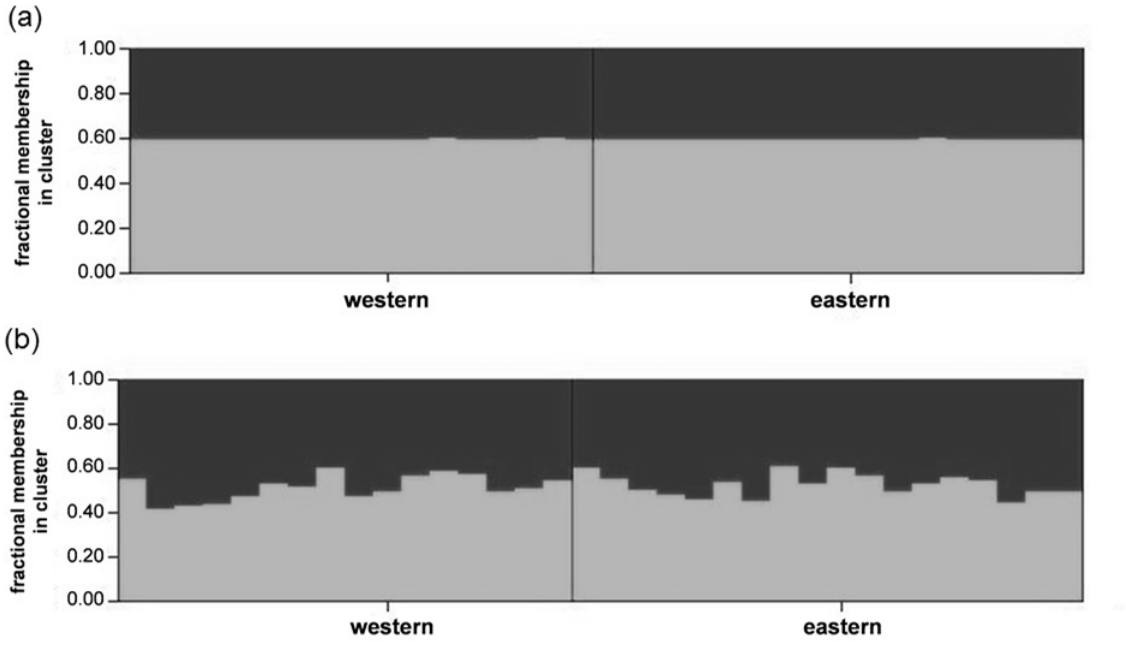


Figure 2.2 *Structure* plots for (a) exon 3 of MHC class I, (b) exon 2 of MHC class II. Each column represents an individual, and each shade denotes a population cluster. Plots were generated for K = 2 populations, using the admixture model with correlated allele frequencies, and included capture location as prior information. No individuals were strongly assigned to either putative cluster.

2.4 Discussion

Enhanced resistance to sympatric parasites and susceptibility to allopatric parasites may result from evolutionary and/or ecological processes. Over evolutionary time, spatial variation in selection pressures may generate geographic variation among host populations at immune-related loci, resulting in host local adaptation. Over ecological time, individual hosts may have previously encountered local parasite strains, resulting in immune memory. I investigated the first of these potential explanations in a study system already determined to show enhanced resistance to sympatric parasites (Sarquis-Adamson and MacDougall-Shackleton 2016). I found no evidence of population genetic differentiation at either class I or II of MHC, suggesting that locally protective alleles at these loci cannot explain greater resistance to sympatric than allopatric parasites. Instead, my findings suggest that individuals' prior immune experience may contribute to home-field advantage in this system. I also found evidence for heterozygote advantage at class I loci, suggesting that variation at MHC is maintained by balancing selection rather than by spatial variation in parasite-mediated selection.

Consistent with many studies of natural populations, I found substantial variation at MHC. Based on the maximum number of different sequences recovered per individual (4 and 26 for class I and class II, respectively), I amplified at least two class I loci and at least 13 class II loci. This finding supports the hypothesis that evolution at MHC in passerine birds has been characterized by multiple gene duplications (e.g., Edwards et al. 1998; Bollmer et al. 2010). Importantly, to reduce effects of PCR or sequencing errors, I retained only sequences that had been recovered from multiple PCR amplifications (class I) or those found in >1% of ISUs (class II). Thus, the diversity reported here may underestimate actual diversity at MHC. As well, because I did not sequence the full PBR at class I, there might be additional, undetected functional variation at exon 2 or at the unsequenced portion of exon 3. I used class I primers designed from cDNA sequences, albeit from another passerine species (house sparrow *Passer domesticus*, Loiseau et al. 2009), thus the class I sequences I recovered are presumably also expressed in song sparrows. By contrast, in order to sequence the entire second exon of class II MHC I used

primers derived from the sequences of the flanking introns, so it is possible that a subset of the class II loci I amplified are not expressed.

Despite considerable variation within and among individuals, my collaborator and I found little evidence for between-site variation (i.e., population genetic structuring) at either class I or class II of MHC. F_{ST} was not significantly different from zero for either class of molecule, we observed no alleles that were common at one site but rare at the other, and my *structure* analysis failed to assign any individuals conclusively to one or the other population based on superallele frequencies at MHC class I or class II. Importantly, sample sizes in my study were low (fewer than 20 individuals genotyped per site) and this restricts they power to detect subtle population structuring. This is a concern particularly at hypervariable loci such as MHC, in which many alleles likely occur at frequencies below 5% and may thus have been missed. More extensive sampling at MHC and at other immune-related loci is required before population genetic differentiation can be conclusively ruled out as an explanation for enhanced resistance to sympatric parasites. Still, based on the apparent lack of differentiation at the MHC regions surveyed, it seems unlikely that locally-protective alleles at MHC can explain Sarquis-Adamson's and MacDougall-Shackleton's (2016) previous finding that song sparrows are less likely to be infected by sympatric than allopatric strains of *Plasmodium*.

I focused on the PBR of class I and class II loci as candidate regions for local adaptation and locally-protective alleles, because protein-sequence variation at these loci determines the repertoire of antigen peptides that can be bound and presented to T-cells. I do not exclude the possibility that locally-protective alleles could be operating at other loci related to immunity (e.g., innate defenses such as β-defensins; Gilroy et al. 2016), or at portions of MHC not sequenced in my study. Locally protective allele effects at MHC have been described mainly in inhabitants of relatively small-scale environments (e.g., river-dwelling Atlantic salmon *Salmo salar*, Dionne et al. 2009; Chinook salmon *Oncorhynchus tshawytscha*, Evans and Neff 2009, Evans et al. 2010; yellow-necked mouse *Apodemus flavicollis*, Meyer-Lucht and Sommer 2005), and could be less pronounced in highly mobile animals such as migratory birds that encounter multiple environments, and multiple

parasite fauna, each year. Alternatively, if one or both populations have been recently bottlenecked, genetic drift at MHC may outweigh selection (e.g., Miller and Lambert 2004) and thus the potential for local adaptation at these loci. However, the high levels of variation I observed at both class I and class II, together with both sites' apparently large population sizes and lack of physical isolation make this latter explanation unlikely.

In light of the apparent lack of population genetic differentiation at MHC, greater resistance to sympatric than allopatric parasites in this system may be driven primarily by ecological rather than evolutionary processes. That is, host individuals may be relatively resistant to the local parasites because they have previously encountered these strains or similar, resulting in immune memory reducing the risk or intensity of re-infections (Møller and Szép 2011; Krzych et al. 2014). Although none of the birds in the cross-infection experiment showed evidence of having been naturally infected by strains with > 95% DNA sequence identity to those used in the experiment (Sarquis-Adamson and MacDougall-Shackleton 2016), molecular screening can fail to detect certain lineages (Valkiūnas et al. 2006). Thus, establishing with certainty the infection history of wild-caught animals is challenging. Conducting experimental infections on captive-raised hosts known to be immunologically naïve (e.g., Jenkins et al. 2015), or using enzyme-linked immunosorbent assays (ELISA) to screen individuals for anti-*Plasmodium* immunoglobulins (e.g., Graczyk et al. 1994), represent promising future directions.

Whereas I found no evidence for population genetic structuring at MHC that would suggest locally-protective alleles might be operating, I did observe an advantage of diversity at MHC class I. Individuals with more amino acid alleles at class I were less likely to become infected when experimentally exposed to *Plasmodium*. This pattern was not observed at class II, quite possibly because *Plasmodium* spp. occur inside the cells of their vertebrate hosts and are thus targeted mainly by class I gene products. Among free-living animals encountering multiple pathogens, non-additive genetic effects at MHC are often observed: individuals with either more alleles at MHC (e.g., striped mouse *Rhabdomys pumilio*, Froeschke and Sommer 2005) or an intermediate number of alleles (three-spined stickleback, Wegner et al. 2003) are less heavily parasitized than individuals with fewer

alleles. In laboratory mice (*Mus musculus domesticus*), similarly, diversity at MHC confers an advantage in resisting and clearing multi-strain bacterial infections (Penn et al. 2002). Although a large body of literature supports the idea that MHC diversity permits recognizing and responding to a broader spectrum of pathogens (i.e., dominant effects), experimental evidence that MHC diversity improves resistance to specific parasites (i.e., overdominant effects) is relatively scant. However, MHC diversity of Gila topminnows (*Poeciliopsis occidentalis*) predicted survival following experimental exposure to a parasitic fluke (Hedrick et al. 2001), suggesting that MHC diversity can influence resistance to specific pathogens.

Overall, the pattern previously observed in this study system, whereby host individuals are relatively resistant to sympatric and relatively susceptible to allopatric strains of *Plasmodium* (Sarquis-Adamson and MacDougall-Shackleton 2016) appears not to be attributable to population differentiation and local adaptation at MHC class I or class II. Although I cannot rule out local adaptation at other immune-related loci, the absence of population differentiation at MHC suggests that immune experience acquired during the course of an individual's lifetime may be a relatively important contributor to home-field advantage in this system. If so, this is cause for some optimism regarding the time-scale over which birds may be able to respond to invasive new parasites. I also observed a protective effect of diversity at MHC class I, emphasizing the importance of adaptive genetic variation in natural populations. Evidence to date suggests that balancing selection, rather than spatial variation in pathogen-mediated selection pressures, contributes to the maintenance of MHC polymorphism in this system.

2.5 References

Altizer S, Bartel R, Han BA. 2011. Animal migration and infectious disease risk. *Science* 331:296–302.

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. *J Mol Biol.* 215:403-410.

Bartoń K. 2015. MuMIn: multi-model inference. *R package version* 1.15.1, https://cran.r-project.org/package=MuMIn.

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

Burnham KP, Anderson DR. 2002. *Model selection and multimodel inference: a practical information-theoretic approach*. Springer-Verlag: New York, NY, USA.

Dionne M, Miller KM, Dodson JJ, Bernatchez L. 2009. MHC standing genetic variation and pathogen resistance in wild Atlantic salmon. *Philos Trans R Soc B*. 364:1555–1565.

Edwards SV, Gasper J, March M. 1998. Genomics and polymorphism of *Agph-DAB1*, an *Mhc* class II B gene in red-winged blackbirds (*Agelaius phoeniceus*). *Mol Biol Evol*. 15:236-250.

Evans ML, Neff BD. 2009. Major histocompatibility complex heterozygote advantage and widespread bacterial infections in populations of Chinook salmon (*Oncorhynchus tshawytscha*). *Mol Ecol*. 18:4716-4729.

Evans ML, Neff BD, Heath DD. 2010. MHC-mediated local adaptation in reciprocally translocated Chinook salmon. *Conserv Genet*. 11:2333-2342.

Falush D, Stephens M, Pritchard JK. 2007. Inference of population structure using multilocus genotype data: dominant markers and null alleles. *Mol Ecol Notes.* 7:574-578.

Fraser BA, Ramnarine IW, Neff BD. 2010. Selection at the MHC class IIB locus across guppy (*Poecilia reticulata*) populations. *Heredity*. 104:155-167.

Froeschke G, Sommer S. 2005. MHC class II DRB variability and parasite load in the striped mouse (*Rhabdomis pumilio*) in the Southern Kalahari. *Mol Biol Evol*. 22:1254-1259.

Garamszegi LZ. 2011. Climate change increases the risk of malaria in birds. *Glob Change Biol*. 17:1751–1759.

Gasteiger E, Gattiker A, Hoogland C, Ivanyi I, Appel RD, Bairoch A. 2003. ExPASy: the proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Res*. 31:3784–3788.

Gilroy D, van Oosterhout C, Komdeur J, Richardson DS. 2016. Avian β-defensin variation in bottlenecked populations: the Seychelles warbler and other congeners. *Conserv Genet*. 17:661-674.

Gloor GB, Hummelen R, Macklaim JM, Dickson RJ, Fernandes AD, MacPhee R, Reid G. 2010. Microbiome profiling by Illumina sequencing of combinatorial sequence-tagged PCR products. *PLoS One*. 5:e15406–e15406.

Graczyk TK, Cranfield MR, Skjoldager ML, Shaw ML. 1994. An ELISA for detecting anti-*Plasmodium* spp. antibodies in African black-footed penguins (*Spheniscus demersus*). *J Parasitol*. 80:60-66.

Greischar MA, Koskella B. 2007. A synthesis of experimental work on parasite local adaptation. *Ecol Lett*. 10:418-434.

Harvell C, Mitchell C, Ward J, Altizer S, Dobson AP, Ostfeld R, Samuel M. 2002. Climate warming and disease risks for terrestrial and marine biota. *Science*. 296:2158-2162.

Hedrick PW, Kim TJ, Parker KM. 2001. Parasite resistance and genetic variation in the endangered Gila topminnow. *Anim Conserv*. 4:103-109.

Hess CM, Edwards SV. 2002. The evolution of the major histocompatibility complex in birds. *BioScience*. 52:423-431.

Jenkins T, Delhaye J, Christe P. 2015. Testing local adaptation in a natural great tit-malaria system: an experimental approach. *PLoS One*. 10:e0141391.

Kalbe M, Kurtz J. 2006. Local differences in immunocompetence reflect resistance of sticklebacks against the eye fluke *Diplostomum pseudospathaceum*. *Parasitology*. 132:105-116.

Kaltz O, Shykoff JA. 1998. Local adaptation in host-parasite systems. *Heredity*. 81:361–370.

Klein J. 1986. *Natural history of the major histocompatibility complex*. John Wiley & Sons: New York, NY, USA.

Krzych U, Zarling S, Pichugin A. 2014. Memory T cells maintain protracted protection against malaria. *Immunol Lett*. 161:189-195.

Lachance M-A, Perri AM, Farahbakhsh AS, Starmer WT. 2013. Genetic structure of *Kurtzmaniella cleridarum*, a cactus flower beetle yeast of the Sonoran and Mojave Deserts: speciation by distance? *FEMS Yeast Res*. 13:674–681.

Loiseau C, Harrigan RJ, Cornel AJ, Guers SL, Dodge M, Marzec T, Carlson JS, Seppi B, Sehgal RNM. 2012. First evidence and predictions of *Plasmodium* transmission in Alaskan bird populations. *PLoS One*. 7:e44729.

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*). *Mol Ecol*. 18:1331–1340.

MacDougall-Shackleton EA, Derryberry EP, Hahn TP. 2002. Nonlocal male mountain white-crowned sparrows have lower paternity and higher parasite loads than males singing local dialect. *Behav Ecol*. 13:682-689.

McCallum H, Dobson AP. 2002. Disease, habitat fragmentation and conservation. *Proc R Soc Lond B.* 269:2041-2049.

Miller HC, Lambert DM. 2004. Genetic drift outweighs balancing selection in shaping post-bottleneck major histocompatibility complex variation in New Zealand robins (Petroicidae). *Mol Ecol*. 13:3709-3721.

Muirhead CA. 2011. Consequences of population structure on genes under balancing selection. *Evolution.* 55:1532-1541.

Møller AP, Szép T. 2011. The role of parasites in ecology and evolution of parasites and migratory connectivity. *J Ornithol*. 152:S141-S150.

Oppliger A, Vernet R, Baez M. 1999. Parasite local maladaptation in the Canarian lizard *Gallotia galloti* (Reptilia: Lacertidae) parasitized by haemogregarian blood parasite. *J Evol Biol*. 12:951-955.

Pagenkopp KM, Klicka J, Durrant KL, Garvin JC, Fleischer RC. 2008. Geographic variation in malarial parasite lineages in the common yellowthroat (*Geothlypis trichas*). *Conserv Genet*. 9:1577–1588.

Penn DJ, Damjanovich K, Potts WK. 2002. MHC heterozygosity confers a selective advantage against multiple-strain infections. *Proc Natl Acad Sci USA*. 99:11260–11264.

Pritchard JK, Wen W. 2003. Documentation for *structure* software: version 2. http://pritchardlab.stanford.edu/structure.html.

Pritchard JK, Stephens M, Donnelly P. 2000. Inference of population genetic structure using multilocus genotype data. *Genetics.* 155:945-959.

R Core Team. 2015. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.r-project.org/

Saino N, Romano M, Scandolara C, Rubolini D, Ambrosini R, Caprioli M, Costanzo A, Romano A. 2014. Brownish, small and lousy barn swallows have greater natal dispersal propensity. *Anim Behav*. 87:137-146.

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

Schierup M, Vekemans X, Charlesworth D. 2000. The effect of subdivision on variation at multi-allelic loci under balancing selection. *Genet Res*. 76:51-62.

Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB, Lesniewski RA, Oakley BB, Parks DH, Robinson CJ, Sahl JW, Stres B, Thallinger GG, Van Horn DJ, Weber CF. 2009. Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Appl Environ Microbiol*. 75:7537-7541.

Stewart KA, MacDougall-Shackleton EA. 2008. Local song elements indicate local genotypes and predict physiological condition in song sparrows *Melospiza melodia*. *Biol Lett*. 4:240-242.

Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. 2013. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Mol Biol Evol*. 30:2725–2729.

Trowsdale J. 1995. “Both man & bird & beast”: comparative organization of MHC genes. *Immunogenetics*. 41:1-17.

Valkiūnas G, Bensch S, Iezhova T, Krizanauskiene A, Hellgren O, Bolshakov C. 2006. Nested cytochrome B polymerase chain reaction diagnostics underestimate mixed infections of avian blood haemosporidian parasites: microscopy is still essential. *J Parasitol*. 92:418-422.

Warner RE.1 968. The role of introduced diseases in the extinction of the endemic Hawaiian avifauna. *Condor*. 70:101-120.

Wegner KM, Reusch TBH, Kalbe M. 2003. Multiple parasites are driving major histocompatibility complex polymorphism in the wild. *J Evol Biol*. 16:224-232.

Westerdahl H, Wittzell H, von Schantz T. 2000. *Mhc* diversity in two passerine birds: no evidence for a minimal essential *Mhc*. *Immunogenetics*. 52:92-100.

Zylberberg M, Lee K, Klasing K, Wikelski M. 2012. Increasing avian pox prevalence varies by species, and with immune function, in Galápagos finches. *Biol Cons*. 153:72-79.

Chapter 3

3 Chemical composition of preen wax reflects major histocompatibility complex similarity in songbirds[2]

3.1 Introduction

In jawed vertebrate animals, the major histocompatibility complex (MHC) is a key component of immune defense. MHC genes encode cell-surface proteins that recognize and bind foreign peptides (antigens), and present them to T cells to initiate an adaptive immune response (Trowsdale 1995). MHC genotype determines the range of antigens to which individuals can respond. Thus, pathogen-mediated selection at MHC often favours locally-adapted alleles (Evans et al. 2010), rare alleles (Takahata and Nei 1990), or certain combinations of alleles (e.g., heterozygote advantage; Penn et al. 2002).

In light of the adaptive importance of MHC, sexual selection should favour mechanisms by which receivers can assess MHC profiles -- and thus the quality or compatibility -- of potential mates. Indeed, MHC influences mating behaviour or preferences in all major vertebrate groups, including fish (Milinski et al. 2005), reptiles (Miller et al. 2009), amphibians (Bos et al. 2009), mammals (Wedekind et al. 1995), and birds (Freeman-Gallant et al. 2003). Although not universal (Westerdahl 2004), these taxonomically widespread patterns imply that MHC profile often varies with some detectable aspect of phenotype.

MHC signalling has been most extensively studied in mammals and fish, taxa in which olfaction is well developed. In mice, for example, MHC peptide ligands (Brennan and Zufall 2006) and volatile nonpeptides (Kwak et al. 2009) occur in urine, and can be

---

[2] A version of this chapter has been published and is presented with permission from *Proceedings of the Royal Society B*.

Citation: Slade JWG, Watson MJ, Kelly TR, Gloor GB, Bernards MA, MacDougall-Shackleton EA. 2016. Chemical composition of preen wax reflects major histocompatibility complex similarity in songbirds. *Proc R Soc B*. 283:20161966.

smelled by conspecific receivers. Similarly, female three-spined sticklebacks (*Gasteroseus aculeatus*) assess MHC profiles of potential mates based on their MHC peptide ligands (Milinski et al. 2005). Chemical communication in birds has received comparatively little study, because birds are generally considered microsmatic relative to other vertebrates (Jones and Roper 1997). However, recent evidence suggests that chemical signalling may be more important in avian communication than previously thought (Balthazart and Taziaux 2009; Caro et al. 2015). Preen wax is secreted from the avian uropygial gland and functions in plumage maintenance (Jacob and Ziswiler 1982), but may also generate semiochemical odour cues. In black-legged kittiwakes (*Rissa tridactyla*), seabirds with well-developed olfaction, pairwise similarity of preen wax chemical profiles correlates with genetic similarity at neutral loci and at MHC (Leclaire et al. 2012; Leclaire et al. 2014). Whether preen wax reflects MHC profiles in other birds − in particular, songbirds (Passeri), in which chemical communication is much less well documented − remains an open question. However, recent work has established that preen wax chemical profiles vary within and between songbird species (Whittaker et al. 2010; Zhang et al. 2013). Further, songbirds can perceive species and sex differences in preen wax composition (Whittaker et al. 2010; Whittaker et al. 2011), making this substance a strong candidate mechanism for chemosignalling.

I examined whether dissimilarity in chemical profiles of preen wax correlates with dissimilarity in MHC genotype in song sparrows (*Melospiza melodia*). Females in the study population adjust levels of parental care based on neutral-locus (microsatellite) similarity to their mates (Potvin and MacDougall-Shackleton 2009), suggesting the presence of some mechanism to assess relatedness. I quantified pairwise chemical distances between individuals' preen wax secretions, measured via gas chromatography, and genetic distances at MHC, measured via next-generation sequencing of the hypervariable peptide-binding region (PBR) of MHC class II, exon 2. I also examined whether chemical diversity of preen wax reflects diversity at MHC. Finally, to further explore the chemical basis by which MHC profiles might be communicated, I identified specific subsets of compounds that best signalled MHC dissimilarity.

3.2 Materials and Methods

3.2.1 Field methods and sample collection

The field research team used seed-baited Potter traps to capture sixty adult song sparrows (19 females, 41 males) at their breeding grounds near Newboro, Ontario, Canada (44°38'38.77"N, 76°20'4.86"W) between April 14 and May 6, 2014. This period corresponds to pair formation and early nesting (first return from wintering grounds April 3rd; first egg May 8th). Upon capturing each bird, a collaborator applied gentle pressure to the uropygial gland at the base of the tail to express ~ 5-10 μL of preen wax. They collected preen wax in an unheparinized glass capillary tube, snapped the capillary tube to fit inside a sterile 1.5 mL microcentrifuge tube, and stored at -20 ºC. They used brachial venipuncture to collect ~ 25 μL of blood, which they blotted onto high wet-strength filter paper saturated with 0.5 M EDTA. They determined sex based on the presence (male) or absence (female) of a cloacal protuberance, supplemented by wing chord measurements. They outfitted birds with unique combinations of coloured leg bands, then released them. In most cases, sex was further confirmed by field observations of sex-specific behaviours including singing, nest-building, or copulation solicitation.

3.2.2 Genetic analysis of MHC

I used polymerase chain reaction (PCR) to amplify the second exon of MHC class II, using a degenerate forward primer (*SospMHCint1f*; 5'-AGY GGG GAY CCG GGG TGG-3') and the reverse primer *Int2r.1* (Edwards et al. 1998) to bind within introns 1 and 2 respectively. In addition to the priming sequences, each primer included an adaptor sequence for the Illumina MiSeq platform, four wobble bases, and a unique 'barcode' of eight bases to assign recovered sequences to individuals. PCR conditions are outlined in section 2.2.2.2.

I confirmed amplification by agarose gel electrophoresis, then pooled products into a library, which I sent for next-generation sequencing using 300 bp paired-end reads on an Illumina MiSeq (London Regional Genomics Centre). A collaborator used a pipeline

(Gloor et al. 2010) to assign sequences to individuals and collapse into clusters of identical reads, and filtered out chimeric sequences (generally fewer than 0.1% of reads) through *de novo* checking in UCHIME (Edgar et al. 2011). I established an error rate of 1% and discarded sequences occurring below this threshold (procedure outlined in section 2.2.2.2).

To confirm that some of the alleles recovered through PCR of genomic DNA (gDNA) are transcribed and expressed, a collaborator and I compared sequences derived from two individuals' gDNA to those derived from the same individuals' cDNA. A collaborator and I collected blood samples for gDNA analysis from two captive-housed song sparrows, then euthanized the birds and dissected out their livers and spleens. We pooled the organs for each bird and homogenized them with 1 mL of TRIzol reagent (ThermoFisher Scientific), isolated total RNA according to the manufacturer's protocol, and quantified RNA using a NanoDrop 2000 spectrophotometer (ThermoFisher Scientific). We reverse-translated RNA to cDNA using the Quanta cDNA synthesis kit (Quanta Biosciences) following the manufacturer's protocol, quantified cDNA by spectrophotometry, and diluted to a working concentration of 20 ng/μL.

For PCR-amplifying class II, exon 2 from cDNA template, I used the forward primer *SongEX1F.2* and the reverse primer *SongEX3R.1* (Aguilar et al. 2006), which bind to exon 1 and 3 respectively. Like the primers used in gDNA PCR, the cDNA primers also included an Illumina MiSeq adaptor sequence, four wobble bases, and a unique barcode of eight bases. PCR and thermocycling conditions were otherwise identical to those used for gDNA. PCR products from each individual's cDNA and gDNA were pooled along with other samples to form a library, which was sent for sequencing at the London Regional Genomics Centre using 300 bp paired-end reads, in a second flow-cell run on an Illumina MiSeq.

Sequences were sorted, collapsed into stacks of identical reads, checked for chimeras and filtered to remove very rare reads, then aligned, trimmed to exon 2, and translated into the corresponding amino acid sequence using MEGA 7.0. One individual had 21 amino acid alleles in its gDNA, seven of which (33.3%) also occurred in its cDNA. The second individual had 26 amino acid alleles in its gDNA, four of which (15.4%) also

occurred in its cDNA. In both cases, the most common alleles in each individual's gDNA were also present in its cDNA, such that shared alleles constituted 49.6% and 30.7% of gDNA sequence reads for the first and second bird, respectively.

I aligned nucleotide sequences in MEGA 7.0 (Kumar et al. 2016), and trimmed out intron sequence based on comparison to other songbird sequences in GenBank. Trimming resulted in alleles of 70-74 amino acids (median = 73), corresponding to the entire putative second exon. I generated a maximum likelihood phylogeny of alleles within each grouping (males only, females only, both sexes) in MEGA 7.0 (Kumar et al. 2016), using a WAG model (Whelan and Goldman 2001) with five discrete gamma categories. I calculated amino acid distances between all pairwise combinations of individuals using the phylogenetic comparison tool UniFrac (Lozupone and Knight 2005). Because genetic data were binary (allele presence vs absence) rather than continuous, I calculated unweighted UniFrac distances, using the R package GUniFrac (Chen 2012).

3.2.3 Chemical analysis of preen wax

A collaborator and I used gas chromatography with flame ionization detection (GC-FID) to separate and quantify chemical compounds in preen wax. Capillary tubes containing preen wax samples were transferred to glass vials and the waxes dissolved in 3 mL of chloroform. A mixture of alkane retention time standards (C19, C30 and C36, 25 ng each in 5 μL) was added to a 100 μL aliquot of each sample. For GC-FID analysis, 1 μL samples were injected into a 5% phenyl methyl siloxane column (DB-5 (Agilent Technologies), 30 m ×0.32 μm ID ×0.25 μm film thickness) on an Agilent 6890N instrument. Samples were injected at 70 ºC (held for one minute), ramped to 130 ºC at 20 ºC per minute, ramped to 320 ºC at 4 ºC per minute, then held at 320 ºC for 10 minutes. Hydrogen was used as a carrier gas at 2.5 mL/min. For preen wax compound identification, a representative sample was analyzed by gas chromatography-mass spectrometry (GC-MS), on a Varian 3800 GC coupled with a Varian MS220 ion trap mass spectrometer. We used the same GC parameters as for the GC-FID analysis, except that He was used as a carrier gas at 1 mL/min. Monoesters were identified on the basis of their $[M]^+$ ion and the fatty acid-

alcohol composition determined by the presence of a protonated fatty acid fragment (Thomas et al. 2010). A mock extraction prepared from an empty glass capillary tube stored inside a microcentrifuge tube, and subsequently extracted as above, yielded no signal in the GC-MS (data not shown).

Because samples varied in the volume of preen wax collected, I quantified relative peak sizes based on peak area relative to that of the full chromatogram. I retained data from peaks comprising at least 0.1% of total chromatogram area (Leclaire et al. 2012) To prevent large peaks from disproportionately influencing distance measures (Leclaire et al. 2014), I normalized data using the *deconstand* function in the R package VEGAN (Dixon 2003). I generated matrices of standardized Bray-Curtis dissimilarity based on chemical distances between all same-sex (male-male, 41 × 41; female-female, 19 × 19) and cross-sex dyads (male-female, 41 × 19) using *bcdist* in the R package ecodist (Goslee and Urban 2007).

3.2.4 Data analysis

I assessed correlations between Bray-Curtis dissimilarity of preen wax and amino acid distance at MHC separately for male-male and female-female dyads, using Mantel tests (*mantel* in VEGAN; Dixon 2003) with 9999 permutations. Using partial Mantel tests to control for capture date did not alter the statistical significance of results, so below I present results derived from simple Mantel tests. Because the pairwise matrix for male-female dyads was not square, a Mantel test was not possible, so I used Spearman's correlation permutation test (*perm.cor.test* in the R package jmuOutlier; Garren 2016), run with 10 000 permutations.

To determine whether individual diversity at MHC (number of alleles) influences chemical diversity of preen wax, I used two measures of chemical diversity for each individual: number of chromatogram peaks, and Shannon's diversity index (calculated using *diversity* in VEGAN; Dixon 2003). I used simple linear regression to model each measure as a function of MHC diversity.

I used BIOENV (Clarke and Ainsworth 1993) implemented in *PRIMER v7* (Clarke and Gorley 2007) to identify the subset of preen wax components that best reflect MHC distance (Stoffel et al. 2015). This approach considers all possible combinations of variables (peaks) at increasing levels of complexity up to a user-specified maximum (here, 8 peaks), and identifies those that maximize the rank correlation between two distance matrices (here, chemical and genetic distance).

Analyses were run in R 3.3.1 (R Core Team 2015). Values are reported as means ± s.e.m., and all statistical tests were two-tailed.

3.3 Results

Across sixty birds, I detected 250 unique DNA sequences (18.47 ± 0.41 alleles/individual) and 69 preen wax peaks (30.05 ± 0.50 peaks/individual). Figure 3.1 shows a sample chromatogram. Males and females did not differ in chemical richness (# peaks; $t_{58} = 1.64$, $p = 0.11$), but pairwise chemical distances were lower for male-male dyads (0.329 ± 0.003) than for female-female (0.384 ± 0.008) or male-female dyads (0.370 ± 0.005; Kruskal-Wallis test, $H_{2,\ 2717} = 63.82$, $p < 0.0001$). Consistent with findings from a closely related species (Thomas et al. 2010), GC-MS indicated that all peaks corresponded to monoesters of varying chain lengths, with total carbon numbers ranging from 23-39 (Appendix B).

Chemical distance of preen wax was positively related to MHC amino acid distance for male-female dyads (correlation permutation test; Spearman's $r = 0.111$, $p = 0.002$, Figure 3.2). This relationship was not significant, however, for same-sex dyads (Mantel test; male-male: $r = -0.058$, $p = 0.80$; female-female: $r = 0.145$, $p = 0.10$). Genetic diversity at MHC (number of alleles) did not predict chemical richness (number of peaks: $R^2_{1,58} = 0.005$, $p = 0.58$) or diversity (Shannon index: $R^2_{1,58} = 0.001$, $p = 0.78$).

BIOENV identified a best combination of four peaks at which chemical and genetic dissimilarity matrices were maximally correlated ($r = 0.222$; Table 3.1). Each of these peaks also occurred repeatedly in the top combinations for other subset sizes (k = 1-8), further suggesting that they are strong candidates for chemosignalling MHC genotype. All

four peaks occurred in the sample used for GC-MS, and could thus be identified (Appendix B). Two were pure substances (peak 32c was a C15 acid esterified to a C17 alcohol, hereafter C15:C17, i.e., heptadecanyl pentadecanoate; peak 33c was C19:C14, i.e., tetradecanyl nonadecanoate). The other two peaks (28b, 34d) were monoester mixtures with the same total carbon number but varying lengths of acid vs alcohol components.

Table 3.1 BIOENV analysis identified subsets of peaks that maximize the rank correlation between chemical (preen wax) and genetic (MHC) distance matrices. The top-ranked subset comprised four peaks (28b, 32c, 33c, 34d), each of which occurred frequently in other subset sizes (denoted by bold type). Chemical composition of peaks is described in Appendix B.

Subset size	**Mantel's *r***	**Peaks**
1	0.125	35c
2	0.190	**28b**, **34d**
3	0.212	**28b**, **32c**, **34d**
4	0.222	**28b**, **32c**, **33c**, **34d**
5	0.221	**28b**, **32c**, **33c**, **34d**, 35b
6	0.219	**28b**, **32c**, 32b, **33c**, **34d**, 36c
7	0.219	**28b**, **32c**, 32b, **33c**, **34d**, 35b, 36c
8	0.215	**28b**, **32c**, 32b, **33c**, **34d**, 35b, 36c, 37b

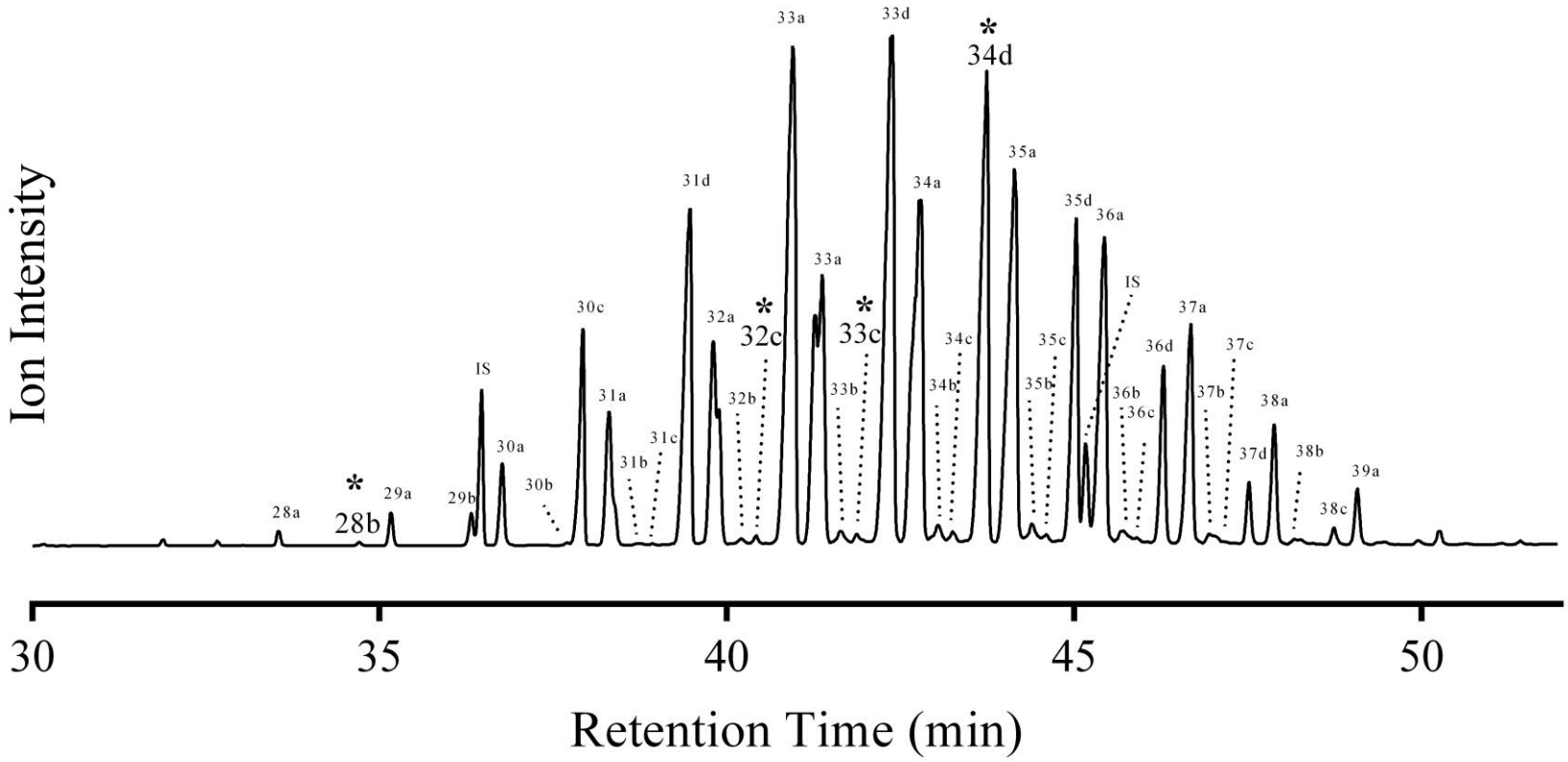


Figure 3.1 Annotated total ion chromatogram from GC-MS analysis of song sparrow preen wax. Peak numbers correspond to the total number of carbons in the compound(s) contributing to the peak; peak letters indicate a subset of monoesters within a given total carbon number category. Asterisks denote peaks for which chemical distance best reflects genetic distance (Table 3.1). IS: internal standard. Chemical compositions of peaks are detailed in Appendix B.

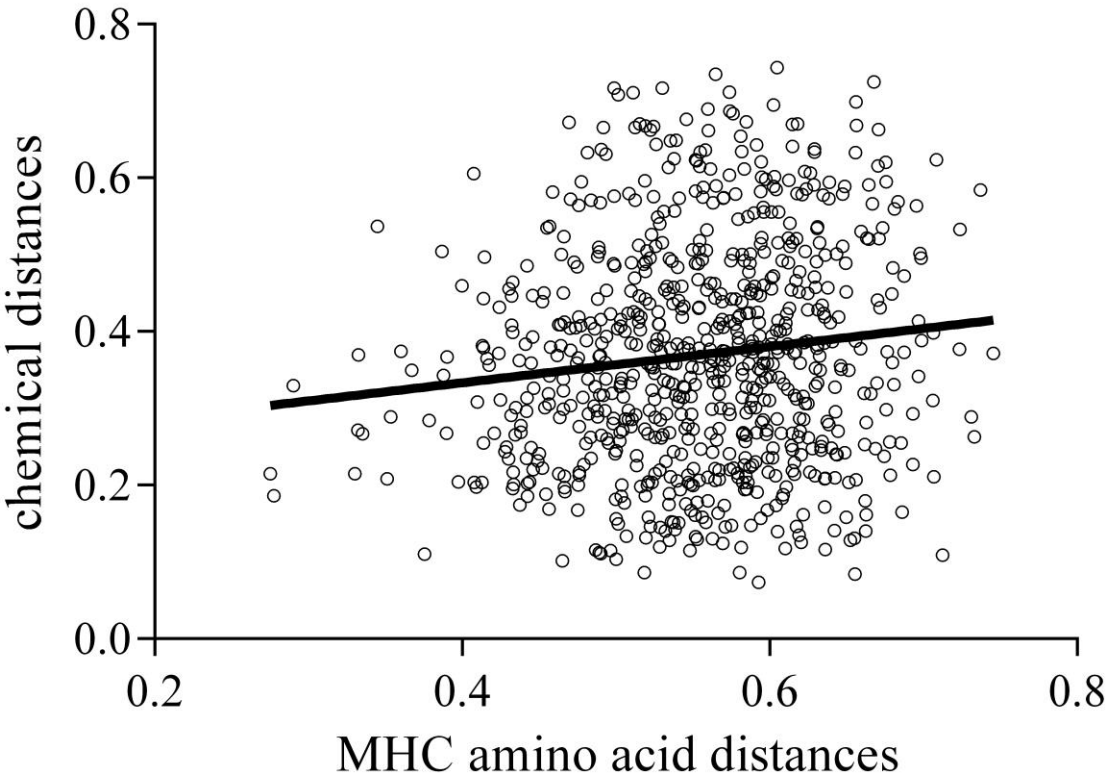


Figure 3.2 Pairwise chemical distances (Bray-Curtis dissimilarity) in preen wax composition of cross-sex song sparrow dyads reflect genetic distances at MHC class II. Solid line shows least-squares regression.

3.4 Discussion

Two conditions are needed for animals to use chemical cues in assessing MHC profiles. First, these cues must covary with MHC; second, animals must be able to perceive such cues. Whereas both conditions are met in fish (Milinski et al. 2005) and mammals (Brennan and Zufall 2006), evidence for chemosignalling in birds has until recently been lacking. Similarity in preen wax composition has recently been shown to reflect MHC similarity in seabirds (Leclaire et al. 2014), and my findings show for the first time that the preen wax of songbirds conveys comparable information. Determining whether song sparrows perceive this variation in chemical signatures, much less use it in mating or other contexts, will require behavioural testing. Still, findings that other songbirds mate nonrandomly at MHC (Freeman-Gallant et al. 2003) and that birds in my study population adjust parental effort based on overall genetic similarity to their mates (Potvin and MacDougall-Shackleton 2009) suggests that some cue exists.

The mechanism by which chemical similarity of preen wax reflects pairwise similarity at MHC remains uncertain (Leclaire et al. 2014). MHC genotype may influence microbial community composition within the uropygial gland, which might generate individual variation in odour profiles either directly or via differences in metabolites (Leclaire et al. 2014; Strandh et al. 2012). Uropygial glands of a closely related species (dark-eyed junco) harbour bacteria capable of synthesizing wax esters (*Acinetobacter* spp.) and others (*Burkholderia* spp., *Pseudomonas* spp.) that may metabolize preen wax into breakdown products of fatty acids and alcohols (Ishige et al. 2002; Whittaker and Theis 2016).

Chemical composition of preen wax is influenced by multiple factors beyond MHC genotype, including diet (Thomas et al. 2010) and seasonal variation in endocrine profiles (Whittaker et al. 2011). For male-male dyads, the effects of seasonal variation in androgens on chemical profiles may have outweighed and obscured effects of MHC. For cross-sex and female-female dyads, I found correlations between MHC and chemical distances (r = 0.11 and 0.15 respectively) comparable to those of free-living seabirds (r = 0.12 and 0.08

respectively; Leclaire et al. 2014). Captive studies, standardizing diet and photoperiod, should generate stronger correlations between MHC and preen wax composition. However, studies on free-living animals (this study; Leclaire et al. 2014) are useful because they permit estimating the degree to which chemosignals reflect MHC under natural conditions. Finally, whereas my study focused on the hypervariable PBR of class II MHC, other genetic factors such as MHC class I and non-MHC loci may also influence chemical signatures.

Of the preen wax compounds detected in this study, likely only a subset can be detected through olfaction. Behavioural experiments represent a key next step in identifying which, if any, compounds might function in chemosignalling. Of particular interest are the four wax esters that best predict MHC genotype (heptadecanyl pentadecanoate, tetradecanyl nonadecanoate, and the mixtures of 28-carbon and 34-carbon waxes), and their fatty acid and alcohol metabolites.

I found a maximum of 26 MHC class II alleles in a single individual, implying at least 13 loci. This is within the range of diversity reported for other New World nine-primary oscines (e.g., 39 alleles in a single individual, implying at least 20 loci, in common yellowthroats *Geothlypis trichas*; Bollmer et al. 2010). Extensive duplication at the avian MHC has generated multiple expressed loci, but also several non-transcribed pseudogenes (Edwards et al. 1998). Indeed, my comparison of gDNA to cDNA profiles suggests that not all gDNA alleles are transcribed: this almost certainly introduces variation into my analyses. However, MHC expression varies with tissue type and infection status (Rohn et al. 1996) thus my estimate of 16-37% of gDNA alleles being transcribed likely underestimates the proportion of loci that are functional.

The salience of MHC genotype to fitness in songbirds (Dunn et al. 2013) suggests that selection should favour the ability to signal and assess MHC profiles. The relationship between MHC and chemical distances for mixed-sex dyads suggests that provided song sparrows can detect chemical cues, this information should be useful in the context of mate choice, regardless of whether a self-referent or a known-kin criterion is used. By contrast, chemical cues do not appear to reflect MHC diversity. My findings implicate preen

secretions as potential semiochemicals in songbirds, a group in which chemical communication has only recently been explored. Further testing is warranted to determine if songbirds can perceive MHC-related variation in chemical profiles. Still, my findings suggest that chemosignalling may be more taxonomically widespread than previously thought, and could help to maintain adaptive genetic diversity in natural populations.

3.5 References

Aguilar A, Edwards SV, Smith TB, Wayne RK. 2006. Patterns of variation in MHC class II β loci of the little greenbul (*Andropadus virens*) with comments on MHC evolution in birds. *J Hered*. 97:133-142.

Balthazart J, Taziaux M. 2009. The underestimated role of olfaction in avian reproduction? *Behav Brain Res*. 200:248-259.

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

Bos DH, Williams RN, Gopurenko D, Bulut Z, Dewoody JA. 2009. Condition-dependent mate choice and a reproductive disadvantage for MHC-divergent male tiger salamanders. *Mol Ecol*. 18:3307-3315.

Brennan PA, Zufall F. 2006. Pheromonal communication in vertebrates. *Nature.* 444:308-315.

Caro SP, Balthazart J, Bonadonna F. 2015. The perfume of reproduction in birds: chemosignaling in avian social life. *Horm Behav*. 68:25-42.

Chen J. 2012 GUniFrac: generalized UniFrac distances. *R Package Version* 1. https://cran.r-project.org/package=GUniFrac

Clarke KR, Ainsworth M. 1993. A method of linking multivariate community structure to environmental variables. *Mar Ecol Prog Ser*. 92:205-219.

Clarke KR, Gorley RN. 2015. *PRIMER v7: User manual/tutorial*. PRIMER-E, Plymouth, UK.

Dunn PO, Bollmer JL, Freeman-Gallant CR, Whittingham LA. 2013 MHC variation is related to a sexually selected ornament, survival, and parasite resistance in common yellowthroats. *Evolution* 67:679-687.

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

Edwards SV, Gasper J, March M. 1998. Genomics and polymorphism of *Agph-DAB1*, an *Mhc* class II B gene in red-winged blackbirds (*Agelaius phoeniceus*). *Mol Biol Evol.* 15:236–250.

Evans ML, Neff BD, Heath DD. 2010. MHC-mediated local adaptation in reciprocally translocated Chinook salmon. *Conserv Genet*. 11:2333-2342.

Freeman-Gallant CR, Meguerdichian M, Wheelwright NT, Sollecito SV. 2003. Social pairing and female mating fidelity predicted by restriction fragment length polymorphism similarity at the major histocompatibility complex in a songbird. *Mol Ecol.* 12:3077-3083.

Garren ST. 2016. jmuOutlier: Permutation tests for nonparametric statistics. *R package version* 1.1. https://cran.r-project.org/package=jmuOutlier.

Gloor GB, Hummelen R, Macklaim JM, Dickson RJ, Fernandes AD, MacPhee R, Reid G. 2010. Microbiome profiling by Illumina sequencing of combinatorial sequence-tagged PCR products. *PLoS One.* 5:e15406-e15406.

Goslee SC, Urban DL. 2007. The ecodist package for dissimilarity-based analysis of ecological data. *J Stat Soft*w. 22:1-19.

Ishige T, Tani A, Takabe K, Sakai Y, Kato N. 2002. Wax ester production from Acinetobacter sp. strain M-1: ultrastructure of cellular inclusions and role of acyl coenzyme A reductase. *Appl Environ Microbiol*. 68:1192-1195.

Jacob J, Ziswiler V. 1982. The uropygial gland. In *Avian Biology*, pp. 199-314. Elsevier: New York, NY, USA.

Jones RB, Roper TJ. 1997. Olfaction in the domestic fowl: a critical review. *Physiol Behav*. 62:1009-1018.

Kwak J, Opiekun MC, Matsumura K, Preti G, Yamakazi K, Beauchamp GK. 2009. Major histocompatibility complex-regulated odortypes: peptide-free urinary volatile signals. *Physiol Behav*. 96:184-188.

Kumar S, Stecher G, Tamura K. 2016. MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. *Mol Biol Evol*. 33:1870-1874.

Leclaire S, Merkling T, Raynaud C, Mulard H, Bessiére JM, Lhuillier E, Hatch SA, Danchin E. 2012. Semiochemical compounds of preen secretion reflect genetic make-up in a seabird species. *Proc R Soc B.* 279:1185–1193.

Leclaire S, van Dongen WFD, Voccia S, Merkling T, Ducamp C, Hatch SA, Blanchard P, Danchin E, Wagner RH. 2014. Preen secretions encode information on MHC similarity in certain sex-dyads in a monogamous seabird. *Sci Rep*. 4:6920.

Lozupone C, Knight R. 2005. UniFrac: a new phylogenetic method for comparing microbial communities. *Appl Environ Microbiol*. 71:8228-8235.

Milinski M, Griffiths S, Wegner KM, Reusch TBH, Haas-Assenbaum A, Boehm T. 2005. Mate choice decisions of stickleback females predictably modified by MHC peptide ligands. *Proc Natl Acad Sci USA.* 102:4414-4418.

Miller HC, Moore JA, Nelson NJ, Daugherty CH. 2009. Influence of major histocompatibility complex genotype on mating success in a free-ranging reptile population. *Proc R Soc B.* 276:1695-1704.

Penn DJ, Damjanovich K, Potts WK. 2002. MHC heterozygosity confers a selective advantage against multiple-strain infections. *Proc Natl Acad Sci USA.* 99:11260-11264.

Potvin DA, MacDougall-Shackleton EA. 2009. Parental investment amplifies effects of genetic complementarity on growth rates in song sparrows, *Melospiza melodia*. *Anim Behav*. 78:943-948.

R Core Team. 2015 R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.r-project.org/

Rohn WM, Lee YJ, Benveniste EN. 1996. Regulation of class II MHC expression. *Crit Rev Immunol*. 16:311-330.

Stoffel MA, Caspers BA, Forcada J, Giannakara A, Baier M, Eberhart-Phillips L, Müller C, Hoffmann JI. 2015. Chemical fingerprints encode mother-offspring similarity, colony membership, genetic quality, and relatedness in fur seals. *Proc Natl Acad Sci USA*. 112:E5005-E5012.

Strandh M, Westerdahl H, Pontarp M, Canbäck B, Dubois MP, 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. *Proc R Soc B.* 279:4457-4463.

Takahata N, Nei M. 1990. Allelic genealogy under overdominant and frequency-dependent selection and polymorphism of major histocompatibility complex loci. *Genetics.* 124:967-978.

Thomas RH, Price ER, Seewagen CL, Mackenzie SL, Bernards MA, Guglielmo CG. 2010. Use of TLC-FID and GC-MS/FID to examine the effects of migratory state, diet and captivity on preen wax composition in white-throated sparrows *Zonotrichia albicollis*. *Ibis.* 152:782-792.

Trowsdale J. 1995. Both man & bird & beast: comparative organization of MHC genes. *Immunogenetics.* 41:1-17.

Wedekind C, Seebeck T, Bettens F, Paepke AJ. 1995. MHC-dependent mate preferences in humans. *Proc R Soc Lond B.* 260:245-249.

Westerdahl H. 2004. No evidence of an MHC-based female mating preference in great reed warblers. *Mol Ecol*. 13:2465-2470.

Whelan S, Goldman N. 2001. A general empirical model of protein evolution derived from multiple protein families using a maximum-likelihood approach. *Mol Biol Evol*. 18:691-699.

Whittaker DJ, Soini HA, Atwell JW, Hollars C, Novotny MV, Ketterson ED. 2010. Songbird chemosignals: volatile compounds in preen gland secretions vary among individuals, sexes, and populations. *Behav Ecol*. 21:608-614.

Whittaker DJ, Richmond KM, Miller AK, Kiley R, Burns CB, Atwell JA, Ketterson ED. 2011. Intraspecific preen oil odor preferences in dark-eyed juncos (*Junco hyemalis*). *Behav Ecol*. 22:1256-1263.

Whittaker DJ, Soini HA, Gerlach NM, Posto AL, Novotny MV, Ketterson ED. 2011. Role of testosterone in stimulating seasonal changes in a potential avian chemosignal. *J Chem Ecol.* 37:1349-1357.

Whittaker, D. J., and K. R. Theis. 2016. Bacterial communities associated with junco preen glands: preliminary ramifications for chemical signaling. In *Chemical Signals in Vertebrates* Vol 13, pp. 105-117. Springer: Cham, Switzerland.

Zhang Y-H, Du Y-F, Zhang J-X. 2013. Uropygial gland volatiles facilitate species recognition between two sympatric sibling bird species. *Behav Ecol*. 24:1271-1278.

Chapter 4

4 Birdsong signals individual diversity at MHC[3]

4.1 Introduction

The major histocompatibility complex (MHC) plays a fundamental role in vertebrate immunity. MHC molecules recognize exogenous peptides (antigens) and present them to T cells, initiating an immune response (Klein 1986). Because MHC genotype determines the suite of antigens that can be recognized, pathogen-mediated selection often favours particular alleles or allelic combinations. Evolutionary arms races with pathogens can impose selection favouring rare alleles, while balancing selection (e.g., heterozygote advantage) often favours individuals with multiple alleles at MHC (Piertney and Oliver 2006). Thus, receivers should benefit by assessing potential mates' MHC profiles (Klein 1986; Piertney and Oliver 2006; Kamiya et al. 2014). Choosing mates with MHC alleles that are dissimilar to one's own or locally-rare should yield offspring that are MHC-diverse or possess rare alleles, conferring genetic benefits via disease resistance (Klein 1986; Potts et al. 1991; Piertney and Oliver 2006; Kamiya et al. 2014). Choosing mates that are themselves MHC-diverse may also enhance offspring MHC diversity (Reusch et al. 2001) and/or inheritance of rare alleles (Hale et al. 2009). Moreover, MHC-diverse mates may provide enhanced parental investment (Knafler et al. 2012) due to superior condition. Signals of locally-rare MHC profiles, and of individual diversity at MHC, are thus likely to be salient to mating decisions (Kamiya et al. 2014).

In mammals, fish, and seabirds, groups with well-developed chemical communication, receivers identify specific (e.g., locally-rare) MHC alleles through olfactory cues from sweat, urine, or preen oil (Penn and Potts 1998; Milinski et al. 2005;

---

[3]A version of this chapter has been published and is presented with permission from *Biology Letters*. Analyses on body condition do not appear in the published version.

Citation: Slade JWG, Watson MJ, MacDougall-Shackleton EA. 2017. Birdsong signals individual diversity at the major histocompatibility complex. *Bio Lett.* 13: 20170430.

Leclaire et al. 2017). Conversely, because individual diversity at MHC affects disease resistance, condition-dependent ornaments may signal MHC diversity (von Schantz et al. 1997; Whittingham et al. 2015). In oscine songbirds, however, learned birdsong could theoretically reflect similarity to the surrounding population as well as individual diversity at MHC. Song learning is generally restricted to early life, meaning that song repertoire content can advertise population of origin (Podos and Warren 2007). Conversely, song complexity can advertise early-life condition (Schmidt et al. 2013) and/or adult immunocompetence (Reid et al. 2005). Song repertoire content and complexity may signal the singer's MHC similarity versus dissimilarity to others in the population (and presumably, compatibility with potential mates), and/or individual diversity at MHC (and presumably, and capacity for parental care). MHC diversity may also influence current body condition, whereby individuals of high MHC diversity are predicted to be in the best condition (for example, have the highest mass corrected for structural size) due to a superior immune system (e.g., Wegner et al. 2003; Lenz et al. 2009).

In song sparrows (*Melospiza melodia*), males learn song during early life. Song varies geographically, and locally-typical repertoires are associated with locally-typical microsatellite genotypes (Stewart and MacDougall-Shackleton 2008), suggesting that repertoire content reflects population of origin. Females prefer local over non-local song (Searcy et al. 2002), and larger over smaller song repertoires (Searcy 1984). I characterized the peptide-binding region of MHC class IIβ to test whether pairwise dissimilarity in song repertoire content reflects MHC dissimilarity to other males in the population (a proxy for locally-rare alleles), and whether song repertoire size and body condition reflects individual diversity at MHC.

4.2 Materials and Methods

4.2.1 Field Sampling

Subjects were 32 male song sparrows at a single breeding site (< 1 km diameter, and not physically isolated from other suitable habitat) near Newboro, Ontario, Canada (44.633°N, 76.330°W). Between 13 April - 3 May, 2015, I captured sparrows in seed-baited traps,

measured the length of tarsus and wing chord to the nearest 0.1 mm with calipers, and measured body mass to the nearest 0.2 g using a Pesola spring scale. I also collected blood for genetic analysis, applied individually-unique colour band combinations for field identification, then released birds.

4.2.2 Song Analysis

A collaborator and I recorded song onto Marantz Professional PMD 671 recorders using Telinga Twin Science Pro parabolic microphones. Recording 200 songs per individual, not necessarily consecutive, is sufficient in most cases to characterize complete repertoires in this population (Potvin et al. 2015). To be conservative, we recorded 300 songs per individual and confirmed by accumulation curves that a plateau had occurred. I digitized recordings in Raven Pro 1.5 (Cornell Lab of Ornithology), inspected spectrograms to identify song types, and noted song repertoire size as the number of different song types each individual produced. I identified a total of 235 syllables (i.e., one or more traces on a spectrogram that always occurred together (Stewart and MacDougall-Shackleton 2008) across all song types. As detailed elsewhere (Stewart and MacDougall-Shackleton 2008), I screened each individual's repertoire for each syllable, constructed a presence-absence syllable matrix, and calculated pairwise Jaccard dissimilarity coefficients adjusted for differences in syllable repertoire size.

4.2.3 Body Condition

I assessed body condition using a scaled mass index (Peig & Green 2009). The formula $\widehat{M}_i = M_i \left[\frac{L_0}{L_i}\right]^{b_{sma}}$ incorporates mass (M) and a linear measurement (L; tarsus or wing cord length) to calculate a scaled mass index ($\widehat{M}_i$). To determine which linear measurement is appropriate to use (tarsus vs. wing cord length), I calculated the correlation of both linear measurements against mass on a log-log scale (Peig and Green 2009). Tarsus length correlated significantly with mass ($r = 0.436$, $p = 0.014$), while wing cord length did not ($r = 0.115$, $p = 0.545$). Therefore, the linear measurement I used was tarsus length. Following the formula, M_i is individual body mass, L_0 is a standardized (mean) tarsus length, and b_{sma}

is the regression coefficient ($\frac{b_{OLS}}{r}$; OLS = ordinary least squares) derived from fitting a line to the formula lnM = ln a + b lnL.

4.2.4 Genetic Analysis

A collaborator and I used primers *SospMHCint1f* (Slade et al. 2017) and *Int2r.1* (Edwards et al. 1998) to amplify exon 2 (β subunit). Details of PCR and sequencing conditions are outlined in section 2.2.2.2 of this thesis. A collaborator sorted sequences into stacks of identical reads using a pipeline (Gloor et al. 2010) and removed chimeras using UCHIME (Edgar et al. 2011). As detailed in section 2.2.2.2, I used a 1% threshold frequency to remove rare reads that could represent PCR or sequencing errors, and compared a subset of reads to cDNA-derived sequences to confirm transcription of some alleles.

I trimmed sequences to remove introns, translated into amino acid sequences of 70-74 codons, and removed apparent pseudogenes based on premature termination codons. Based on a maximum likelihood allele phylogeny with WAG substitution and five discrete gamma categories, I used the unweighted UniFrac algorithm in the R package GUniFrac (Chen 2012) to calculate pairwise genetic distances between individuals. Alleles in the same clade are presumably similar functionally, so to be conservative in estimating genetic diversity, I clustered into "superalleles" based on well-defined clade membership (Slade et al. 2017). I scored each individual's MHC diversity as the number of different superalleles.

4.2.5 Data Analysis

To test whether song dissimilarity signals MHC dissimilarity to other males in the population, I assessed the correlation between Jaccard dissimilarity and UniFrac genetic distance, using a Mantel test with 9999 permutations (mantel in VEGAN (Dixon 2003)).

To test whether song complexity and/or scaled body mass varies with MHC diversity, I compared support for three models predicting song repertoire size: a linear model (number of superalleles), a quadratic model (number of superalleles + squared number of superalleles), and a null model. I ranked models using Akaike's corrected

information criterion (AICc), setting a threshold of 2 AICc units for model averaging. I checked models for highly influential points, and confirmed Cook's distance < 0.5 in all cases.

Analyses were performed in R 3.4.0 (R Core Team 2017); values reported are means ± s.e.m.

4.3 Results

I detected 186 unique alleles at MHC IIβ, which clustered to 91 superalleles (Appendix C.1; 13.6 ± 0.5 superalleles per individual). Repertoire size ranged from 5-12 song types (7.8 ± 0.3 per individual).

Song dissimilarity was not associated with genetic distance at MHC (Mantel's r = -0.01, p = 0.55).

In predicting song repertoire size, the quadratic model received 5-16 times more support than the null or linear models (table 4.1; parameter estimates in table 4.2). The largest song repertoires occurred in males with intermediate MHC diversity (Figure 4.1). Excluding one individual with low repertoire size and high MHC diversity (Figure 4.1, upper leftmost point) did not qualitatively alter significance of results (Appendix C.2, C.3).

There was low support for MHC in predicting scaled body mass as all three models had similar model weights (table 4.3).

Table 4.1 Ranked set of candidate models predicting song sparrow song repertoire size. Predictors were number of MHC superalleles and squared number of MHC superalleles.

Model	**df**	**logLik**	**AICc**	**ΔAICc**	**Model Weight**
Quadratic: #MHC + $(\#MHC)^2$	4	-54.8	119.0	0	0.78
Null: intercept only	2	-58.9	122.2	3.20	0.16
Linear: #MHC	3	-58.8	124.4	5.40	0.05

Table 4.2 Parameter estimates from the best-supported model predicting song sparrow song repertoire size. Repertoire size increased with number of MHC class IIβ superalleles, but decreased with squared number of superalleles. $R^2 = 0.23$, $F_{2,29} = 4.24$.

Parameter	**Estimate ± SE**	**95% CI**
Intercept	-5.27 ± 4.97	-15.4, 4.89
#MHC	2.12 ± 0.76	0.56, 3.67
$(\#MHC)^2$	-0.082 ± 0.029	-0.14, -0.02

Table 4.3 Ranked set of candidate models predicting song sparrow scaled body mass. Predictors were number of MHC superalleles and squared number of MHC superalleles.

Model	**df**	**logLik**	**AICc**	**ΔAICc**	**Model Weight**
Quadratic: #MHC + $(\#MHC)^2$	4	-61.9	133.39	0	0.39
Linear: #MHC	3	-63.2	134.42	0.03	0.38
Null: intercept only	2	-65.0	134.40	1.01	0.23

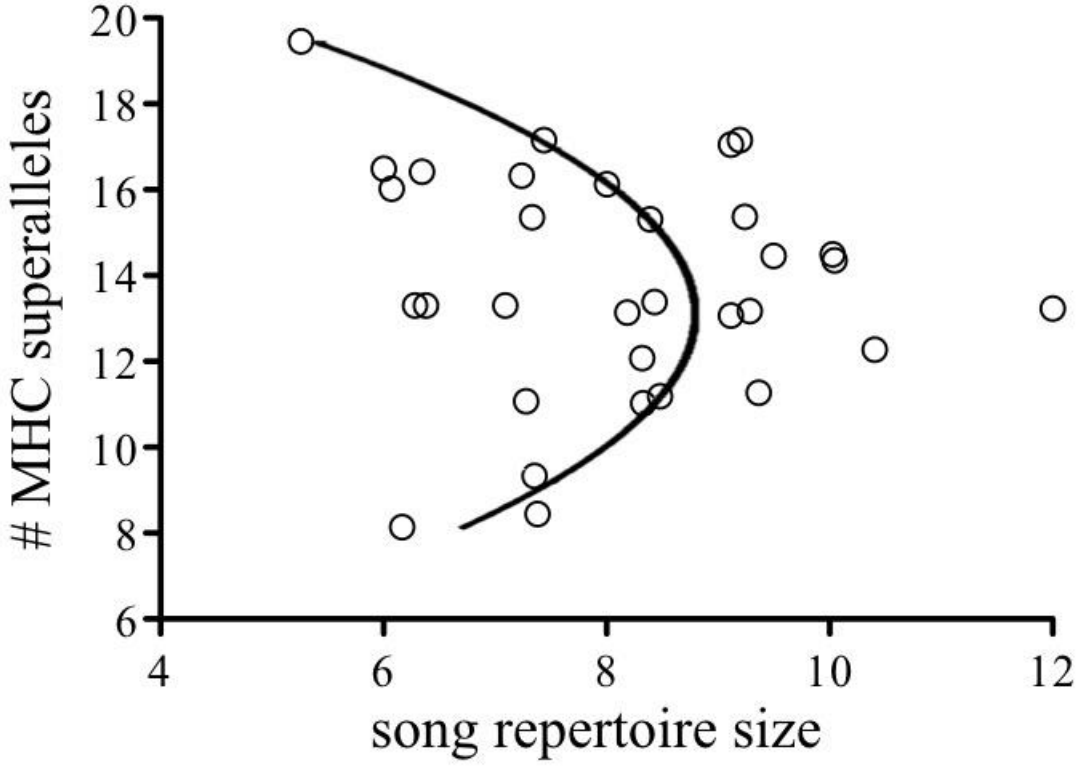


Figure 4.1 Relationship between song repertoire size and MHC class IIβ superallele diversity. Curve depicts best-supported model described in tables 4.1 and 4.2.

4.4 Discussion

MHC-related mating preferences have been observed in all vertebrate classes (Kamiya et al. 2014), raising the question of how animals assess MHC profiles. Cues of compatibility and diversity are generally studied in the contexts of chemosignalling (Penn & Potts 1998; Milinski et al. 2005; Leclaire et al. 2017) and visual ornaments (von Schantz et al. 1997; Hale et al. 2009; Whittingham et al. 2015), respectively. I investigated birdsong, an acoustic ornament as a signal of MHC class IIβ dissimilarity to other individuals in the population (a proxy for locally-rare genotypes). I also investigated birdsong repertoire size, and scaled body mass in respect to individual MHC genetic diversity.

Geographic variation in song has long been proposed to advertise population of origin, suggesting receivers might use song to achieve an optimal balance between inbreeding and outbreeding (Nottebohm 1969; Podos & Warren 2007). Finding no relationship between song and MHC dissimilarity could reflect low genetic differentiation at MHC in this system (Slade et al. 2017): if MHC does not vary with population of origin, song is unlikely to signal MHC dissimilarity. However, I examined only one class of MHC and cannot rule out associations between song repertoire content and class I loci, whose products interact with intracellular pathogens such as viruses (Klein 1986).

Whereas MHC class IIβ dissimilarity to other males at the site was not associated with song dissimilarity, individual diversity at MHC explained 23% of the variation in song repertoire size. This supports previous findings that MHC diversity influences ornamentation (von Schantz et al. 1997; Hale et al. 2009; Whittingham et al. 2015), but an association with an acoustic ornament is novel (although see Reid et al. (2005) for evidence that song repertoire size advertises cell-mediated immunity in this species). Also notable is the nonlinear nature of this relationship: larger song repertoires were associated with intermediate, not maximal, MHC diversity. This might reflect trade-offs between susceptibility to pathogens versus autoimmune disorders (Klein 1986; Piertney & Oliver 2006), or a dilution effect whereby overly-diverse MHC profiles have too few copies of protective alleles. In choosing social mates with complex song, females may obtain

material benefits through increased paternal investment. To the extent that males with optimal MHC diversity produce optimally-diverse offspring, preferences for complex song may also confer genetic benefits.

In light of the significance from a condition-dependent trait learned in early life (birdsong) being associated with MHC class IIβ diversity, I did not find any association with adult body condition. In contrast, long-tailed giant rats (*Leopoldamys sabanus*) with a diversity of highly divergent alleles (thought to be functionally dissimilar) were in better body condition than individuals with less divergent alleles (Lenz et al. 2009), which is thought to be linked to their immunocompetence, as MHC-divergent individuals also had fewer parasites. Thus, my results indicate that a quadratic model of MHC class IIβ diversity is good at predicting birdsong complexity (which is learned during the early stages of a songbird's life) rather than MHC diversity predicting adult male song sparrow's body condition.

4.5 References

Chen J. 2012. GUniFrac: generalized UniFrac distances. *R Package Version* 1. https://cran.r-project.org/package=GUniFrac

Dixon P. 2003. VEGAN, a package of R functions for community ecology. *J Veg Sci*. 14:927–930.

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

Edwards SV, Gasper J, March M. 1998. Genomics and polymorphism of *Agph-DAB1*, an *Mhc* class II B gene in red-winged blackbirds (*Agelaius phoeniceus*). *Mol Biol Evol.* 15:236–250.

Gloor GB, Hummelen R, Macklaim JM, Dickson RJ, Fernandes AD, MacPhee R, Reid G. 2010. Microbiome profiling by Illumina sequencing of combinatorial sequence-tagged PCR products. *PloS One.* 5:e15406–e15406.

Hale ML, Verduijn MH, Møller AP, Wolff K, Petrie M. 2009. Is the peacock's train an honest signal of genetic quality at the major histocompatibility complex? *J Evol Biol*. 22:1284–1294.

Kamiya T, O'Dwyer K, Westerdahl H, Senior A, Nakagawa S. 2014. A quantitative review of MHC-based mating preference: the role of diversity and dissimilarity. *Mol Ecol,* 23:5151–5163.

Klein J. 1986. *Natural history of the major histocompatibility complex*. John Wiley & Sons: New York, NY, USA.

Knafler GJ, Clark JA, Boersma PD, Bouzat JL. 2012. MHC diversity and mate choice in the Magellanic penguin, *Spheniscus magellanicus*. *J Hered.* 103:759–768.

Leclaire S, Strandh M, Mardon J, Westerdahl H, Bonadonna F. 2017. Odour-based discrimination of similarity at the major histocompatibility complex in birds. *Proc R Soc B*. 284:20162466.

Lenz TL, Wells K, Pfeiffer M, Sommer S. 2009. Diverse MHC IIB allele repertoire increases parasite resistance and body condition in the Long-tailed giant rat (*Leopoldamys sabanus*). *BMC Evol Biol.* 9:269.

Milinski M, Griffiths S, Wegner KM, Reusch TBH, Haas-Assenbaum A, Boehm T. 2005. Mate choice decisions of stickleback females predictably modified by MHC peptide ligands. *Proc Natl Acad Sci USA*.102:4414–4418.

Nottebohm F. 1969. The song of the chingolo, *Zonotrichia capensis*, in Argentina: description and evaluation of a system of dialects. *The Condor*. 71:299–315.

Peig J, Green AJ. 2009. New perspectives for estimating body condition from mass/length data: the scaled mass index as an alternative method. *Oikos*. 118:1883–1891.

Penn D, Potts W. 1998. How do major histocompatibility complex genes influence odor and mating preferences. *Adv Immunol*. 69:411–436.

Piertney SB, Oliver MK. 2006. The evolutionary ecology of the major histocompatibility complex. *Heredity*. 96:7–21.

Podos J, Warren PS. 2007. The evolution of geographic variation in birdsong. *Adv Study Behav*. 37:403–458.

Potts WK, Manning CJ, Wakeland EK. 1991. Mating patterns in seminatural populations of mice influenced by MHC genotype. *Nature*. 352:619–21.

Potvin DA, Crawford PW, MacDougall-Shackleton SA, MacDougall-Shackleton EA. 2015. Song repertoire size, not territory location, predicts reproductive success and territory tenure in a migratory songbird. *Can J Zool*. 93:627–633.

R Core Team. 2017. R: a language and environment for statistical computing. R Foundation for Statistical Computing: Vienna, Austria. URL https://www.r-project.org/

Reid J, Arcese P, Cassidy AEV, Marr A, Smith JM, Keller L. 2005. Hamilton and Zuk meet heterozygosity? Song repertoire size indicates inbreeding and immunity in song sparrows (*Melospiza melodia*). *Proc R Soc B.* 72:481–487.

Reusch TBH, Haberli MA, Aeschlimann PB, Milinski M. 2001. Female sticklebacks count alleles in a strategy of sexual selection explaining MHC polymorphism. *Nature*. 414: 300–2.

von Schantz T, Wittzell H, Göransson G, Grahn M. 1997. Mate choice, male condition-dependent ornamentation and MHC in the pheasant. *Hereditas*. 127:133–140.

Schmidt KL, Moore SD, MacDougall-Shackleton EA, MacDougall-Shackleton SA. 2013. Early-life stress affects song complexity, song learning and volume of the brain nucleus RA in adult male song sparrows. *Anim Behav.* 86:25–35.

Searcy WA. 1984. Song repertoire size and female preferences in song sparrows. *Behav Ecol Sociobiol*. 14:281–286.

Searcy WA, Nowicki S, Hughes M, Peters S. 2002. Geographic song discrimination in relation to dispersal distances in song sparrows. *Am Nat.* 159:221–230.

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

Slade JWG, Watson MJ, Kelly TR, Gloor GB, Bernards MA, MacDougall-Shackleton EA. 2016. Chemical composition of preen wax reflects major histocompatibility complex similarity in songbirds. *Proc R Soc B.* 283:20161966.

Stewart KA, MacDougall-Shackleton EA. 2008. Local song elements indicate local genotypes and predict physiological condition in song sparrows *Melospiza melodia*. *Biol Lett.* 4:240–242.

Wegner KM, Kalbe M, Kurtz J, Reusch TBH, Milinski M. 2003. Parasite selection for immunogenetic optimality. *Science.* 301:1343.

Whittingham LA, Freeman-Gallant CR, Taff CC, Dunn PO. 2015. Different ornaments signal male health and MHC variation in two populations of a warbler. *Mol Ecol.* 24: 1584–1595.

Chapter 5

5 Signatures of positive selection at class II MHC accompanied by assortative, not disassortative, mating in song sparrows

5.1 Introduction

The major histocompatibility complex (MHC) is the cornerstone of the immune system in jawed vertebrates (Klein 1986; Trowsdale 1995), as MHC molecules recognize multiple non-self antigens and present them T cells to elicit an immune response (Trowsdale 1995). The MHC is composed of two main classes: class I and class II molecules interact primarily with intracellular and extracellular antigens, respectively. Parasite-host interactions are expected to impose strong selective pressure at MHC genes, particularly at codons within the peptide-binding region (PBR) which encode for amino acids that interact directly with antigens; these amino acids determine the suite of pathogens that can be recognized (e.g., Spurgin and Richardson 2010).

Positive selection, defined as an excess of nonsynonymous substitutions relative to synonymous substitutions, is relatively common at codons within the PBR of both classes of MHC (Hughes and Hughes 1995; Piertney and Oliver 2006). By contrast, positive selection tends to be rare at other protein-coding loci, which generally show evidence of negative (purifying) selection, associated with the removal of deleterious mutations from a population (Yang and Swanson 2002). Indeed, MHC codons that encode amino acids that contact antigens directly due to their position within the PBR show the highest levels of positive selection (Hughes and Nei 1989; Kasahara 2000), presumably due to parasite-mediated selection favouring rare and new variants at these amino acids. Thus, the PBR is generally the major site of genetic variation at both class I and class II MHC (Hughes and Hughes 1995).

At least two, non-mutually exclusive, hypotheses may explain how balancing selection maintains high genetic diversity at MHC and particularly within the PBR. First, non-additive genetic effects, such as heterozygote advantage, often confer a fitness benefit

to individuals that are highly diverse at MHC (e.g., McClelland et al. 2003). For example, heterozygosity at MHC predicted external parasite load in water voles (*Arvicola terrestris*, Oliver et al. 2009). Individuals with more MHC alleles can recognize a wider variety of pathogens than can less MHC-diverse individuals (Kubinak et al. 2012) and are therefore assumed to be more immunocompetent. Second, models of antagonistic co-evolution between hosts and parasites posit that hosts with common alleles may be relatively vulnerable to infection, while individuals with new, rare alleles at MHC have a fitness advantage in the population (Kubinak et al. 2012). Over evolutionary time, negative-frequency dependent selection increases the frequency of the rare allele until it is widespread throughout the population, at which time new rare alleles are favoured (Hughes and Nei 1992; Slade and McCallum 1992). Overall, balancing selection via heterozygote advantage and negative-frequency dependent selection can help maintain MHC diversity.

The fitness advantages associated with being diverse at MHC and/or possessing rare alleles have also been hypothesized to favour MHC-mediated mate choice (Yamazaki et al. 1976; Penn and Potts 1998; Neff and Pitcher 2005; Kubinak et al. 2012; Huchard et al. 2013). Compatible genes models of mate choice predict that genetic quality reflects non-additive interactions between paternal and maternal genomes (Neff and Pitcher 2005). To the extent that MHC heterozygosity improves disease resistance, the most compatible mate for a given individual may be the one with a maximally dissimilar MHC genotype (Mays and Hill 2004; Neff and Pitcher 2005). Supporting this, MHC-disassortative mate preference has been found in multiple vertebrate taxa, including mammals (e.g., lab mice *Mus musculus domesticus* Yamazaki et al. 1979; bank voles *Myodes glareolus*, Radwan et al. 2008; mandrills *Mandrillus sphinx*, Setchell et al. 2010; and humans, Wedekind et al. 1995); fish (e.g., brown trout *Salmo trutta*, Forsberg et al. 2007; rose bitterlings *Rhodeus ocellatus*, Agbali et al. 2010); and birds (Savannah sparrows *Passerculus sandwichensis*, Freeman-Gallant et al. 2003). Indeed, rose bitterling females that paired with MHC-dissimilar males had greater embryo survival when the phylogenetic and functional distance of positively selected codons were the highest (Agbali et al. 2010). However, MHC-disassortative pairings or preferences are not universal (reviewed by Ruff et al. 2012). In some systems, pairing with MHC-similar mates might be beneficial in terms of

avoiding outbreeding depression at this or other loci (Mays and Hill 2004; Bonneaud et al. 2006; Galaverni et al. 2015). In particular, in species with multiple loci expressed at MHC, natural selection may favor intermediate diversity at MHC since being too diverse at these loci may cause autoimmune disorders (Wegner et al. 2003). Therefore, sexual selection might not promote disassortative mating at MHC in order to control for being too outbred at these immune genes.

In this study, I examined MHC-mediated mate choice in a population of free-living song sparrows (*Melospiza melodia*). In my population, class II MHC profiles covary with the chemical composition of preen oil, which provides a potential candidate mechanism for assessing class II MHC profiles (Slade et al. 2016). I hypothesized that these birds select social mates based in part on MHC genotypes, preferentially pairing with MHC-dissimilar individuals. In particular, I expected mating to be disassortative at MHC codons subject to positive selection; thus, I examined patterns of molecular selection within exon 2 (comprising the PBR) of MHC class II to identify specific codons under positive selection in this population. I calculated multiple measures of pairwise genetic distance at MHC, including evolutionary and functional distance measures, at the entire exon 2 and considering only the codons under positive selection. Using these four measures of genetic distance at class II MHC (evolutionary vs functional; entire exon vs positively selected codons only), I calculated pairwise genetic distances at MHC for observed breeding pairs and compared these values to distributions of pairwise genetic distance that would be expected under random mating.

5.2 Materials and Methods

5.2.1 Field methods and sample collection

The field team collected blood samples for genotyping, MHC and identified naturally-occurring social pairs, during the breeding seasons of 2014 and 2015 on a long-term study population of song sparrows breeding near Newboro, Ontario, Canada (44.6338°N, 76.3308°W). Field data collection occurred between April 14 - June 2 in 2014, and April 13 - June 6 in 2015. This time frame corresponds to pair formation, nesting, and offspring

provisioning in this population.

The field team used two-cell, seed-baited Potter traps, checked once per hour three times each day, to capture 69 birds in 2014 (n_{males} = 44, $n_{females}$ = 25) and 87 birds in 2015 (n_{males} = 49, $n_{females}$ = 38), with 28 birds overlapping between the two years. The field team used brachial venipuncture to collect ~ 25 μL of whole blood from each individual, which was blotted onto high-strength filter paper saturated with 0.5M EDTA. Blood blots were dried and stored for subsequent DNA extraction. The field team used wing chord measurements and the presence (male) or absence (female) of a cloacal protuberance to identify sex. All birds received a unique combination of color bands for field identification, and were released at their capture site.

5.2.2 Assigning social pairs

I focused on social mate choice, rather than genetic mate choice as my study population shows a very low occurrence of extra-pair paternity (less than 20% of nests; Potvin and MacDougall-Shackleton 2009). Thus, social mate choice appears to be a good proxy for genetic mate choice in this population. I identified social pairs through a combination of opportunistic behavioral observations on colour-banded individuals (observing a female and male interacting on a territory and noting the occurrence of reproductive behaviours such as nest-building, copulation, provisioning offspring) and trapping records. My criteria for considering a male and female to be socially paired based on trapping records were: (1), the male and female being trapped at the same time in the same trap (occurred in 8.7% of all trappings), or (2) the male and female were trapped in the same trap within 48 hours of each other and were the only two song sparrows to be trapped at that location (occurred in 8.8% of all trappings). Using these criteria, I identified 18 social pairs in 2014 and 22 pairs in 2015 (n_{pairs} = 40). One social pair remained the same over both study years, and was only included once in the analysis. Another three birds (one male, two females) from the 2014 dataset were also included in the 2015 analysis as they paired with different mates each year.

5.2.3 MHC characterization

I extracted DNA from blood blots using an ammonium acetate protocol to salt out proteins (Laitinen et al. 1994). DNA quantity and quality was checked on a spectrophotometer (Thermo Scientific Nanodrop 2000).

I used polymerase chain reaction (PCR) to amplify MHC class II exon 2 (β chain; hypervariable PBR). I used a degenerate forward primer (*SospMHCint1f;* Slade et al. 2017) and used a previously developed reverse primer (*Int2r.1*; Edwards et al. 1998) to bind within introns 1 and 2 respectively. Each primer also consisted of a barcode sequence, four wobble bases, and an adaptor for the Illumina MiSeq platform.

PCR took place in 30 μL reactions using 12.5 μL of GoTaq® Hot Start Master Mix (Promega), 0.2 μM of each primer, and 25-60 ng of gDNA. Thermocycler conditions comprised of 3 min at 94 °C; 28 cycles of 30 s at 94 °C, 30 s at 62 °C, and 45 s at 72 °C; with a final extension step of 10 min at 72 °C.

MHC amplification was checked on a 2% agarose gel. Amplicons were pooled together based on the year (2014 and 2015) to form a library and run on a separate Illumina MiSeq flow cells at the London Regional Genomics Center. Using the barcode sequence, a colleague and I retrieved individual MHC data through a pipeline developed by Gloor et al. (2010). Chimeric sequences were removed using UCHIME (Edgar et al. 2011). To recognize potential sequences that resulted from PCR or sequencing errors, I removed all sequences below 1% of the total reads for each individual following the protocol outlined in section 2.2.2.2 of this thesis. I aligned sequences using MEGA 7.0 (Kumar et al. 2016) and confirmed similarity to other passerine MHC class II exon 2 sequences queried via the Basic Local Alignment Search tool (BLAST; Altschul et al. 1990) implemented in GenBank.

5.2.4 Models of selection at MHC

Using codeml (PAML 4.9; Yang 2007) implemented in the program PAMLx 1.3.1 (Xu and Yang 2013), I tested for positive selection at MHC class II exon 2 using 518 DNA

alleles from my population dataset, each containing up to 225 bases. Codeml does not make a priori assumptions on which codons would be under positive selection, such as those that encode peptide-binding amino acids (Yang and Swanson 2002; Bollmer et al. 2010). I used the M1a (nearly neutral; $\omega_0 < 1$, $\omega_1 = 1$), M2a (positive selection; $\omega_2 > 1$), M7 (β; null model for M8), and M8 (β and ω; $0 < \omega < 1$, $\omega > 1$) models. The M2a (positive selection) and M8 (β and ω) models were of particular interest since they test for which codons are under positive selection. Positively selected codons were determined using the Bayes empirical Bayes approach (Yang et al. 2005). I aligned sequences to MHC class II exon 2 from a New Zealand saddleback (*Philesturnus carunculatus*; Accession No. AGO86446.1; Sutton et al. 2016), and following Brown et al. (1993), I determined putative codons that encode peptide-binding amino acids. Finally, I used Akaike information criterion (AIC; Burnham and Anderson 2002) to evaluate which maximum likelihood codon model best fits the data, using a threshold of ≥ 2 AIC units to identify the best supported model.

5.2.5 Mate choice analysis

To test if song sparrows paired non-randomly at MHC, I calculated genetic distances between social pairs observed in the wild, and compared them to all possible opposite-sex combinations based on genotypes recovered in each year. I used measures of evolutionary (phylogenetic) and functional (chemical) distances using all amino acids in exon 2, and the positively selected amino acids determined using the Bayes empirical Bayes from the positive selection model in codeml. All male-female pairwise distances were calculated using an unweighted UniFrac algorithm (Lozupone and Knight 2005) in the package GUniFrac (Chen 2012) in R 3.4.0 (R Core Team 2017).

5.2.5.1 Evolutionary distance

I aligned MHC amino acid sequences using MEGA 7.0 (Kumar et al. 2016). I calculated pairwise evolutionary distances, based on sequences at the full exon 2 and on the subset of codons found to be under positive selection, with a maximum likelihood phylogeny. This phylogeny was modeled using Whelan and Goldman substitution (Whelan and Goldman 2001) with five discrete gamma categories.

5.2.5.2 Functional distance

The functional properties of amino acids may explain MHC-mediated mate choice (Strandh et al. 2012; Leclaire et al. 2017). To calculate functional distance, I used the chemical properties of each amino acid using five z-score descriptors described by Sandberg et al. (1998): z_1 (hydrophobicity), z_2 (steric bulk), z_3 (polarity), z_4 and z_5 (electronic properties). I aligned MHC amino acid sequences using MEGA 7.0 (Kumar et al. 2016). Because functional amino acid data are continuous, I developed functional trees for both exon 2 and the positively selected amino acids using contml in PHYLIP 3.695 (Felsenstein 2005). Contml requires identical lengths of values for each allele. Therefore, I removed any sequences with indels, which retained 278 alleles for the mate choice analysis for both distance measures.

5.2.5.3 Simulation analysis

Using a Monte Carlo simulation (Manly 1997) in Microsoft Excel, a colleague and I compared the mean distances (evolutionary and functional) of all true pairs to 10 000 randomized pairs sorted by year. The model assumed that every female had the potential to choose a male during their field season. Two-tailed p-values were calculated for each distance measure.

5.3 Results

Song sparrows had an average of 14.1 ± 0.23 MHC class II alleles per individual. M1a (nearly neutral) model indicated that 98.1% of sites within exon 2 of MHC class II were under purifying selection (Table 5.1), while 1.93% of sites had neutral base substitutions. The M2a (positive selection), which is the best representative model for my data (Table 5.1) showed that 94.3% of sites were under purifying selection, while 4.76% had neutral base substitutions, and 0.99% were under positive selection (Table 5.1). Using a β distribution (M7 and M8), 99.1% of sites were under purifying selection, while 0.862% of sites were under positive selection (Table 5.1). Both M2a and M8 estimated the same positively selected codons using the Bayes empirical Bayes approach.

The mean evolutionary distances from true pairs were significantly lower than those of simulated pairs for exon 2 (Figure 5.1; mean evolutionary distance = 0.623, $p = 0.0176$), implying song sparrows may pair assortatively at MHC class II exon 2. The rest of the mean distances from the true pairs were not significantly different than the simulated distances, implying that song sparrows do not pair according to the functional distance of MHC class II exon 2, nor do they pair based on the evolutionary distance and functional distance of the amino acids encoded by positively selected codons within exon 2.

Table 5.1 Codon site models determined using phylogenetic analysis by maximum likelihood. Positively selected codons with asterisks correspond to peptide-binding codons as described by Brown et al. 1993.

Codon ML Model	lnL	AIC	ΔAIC	Parameter Estimates	Positively Selected Sites
M1a: nearly neutral	-145 658	291 324	464	p_0 = 0.981, (p_1 = 0.0193) ω_0 = 0.0492, ω_1 = 1	Not allowed
M2a: positive selection	-145 424	290 860	0 (best model)	p_0 = 0.943, p_1 = 0.0476, (p_2= 0.0099) ω_0 = 0.12, ω_1 = 1, ω_2 = 4.27	1, 2*, 15, 24*, 30, 34, 37*, 38*, 43, 44, 47*, 49*
M7: β	-145 705	291 414	554	p = 0.0149, q = 0.104	Not allowed
M8: β & ω	-145 426	290 862	2	p_0 = 0.991, (p_1 = 0.00862), p = 0.190, q = 1.56, ω = 3.36	1, 2*, 15, 24*, 30, 34, 37*, 38*, 43, 44, 47*, 49*

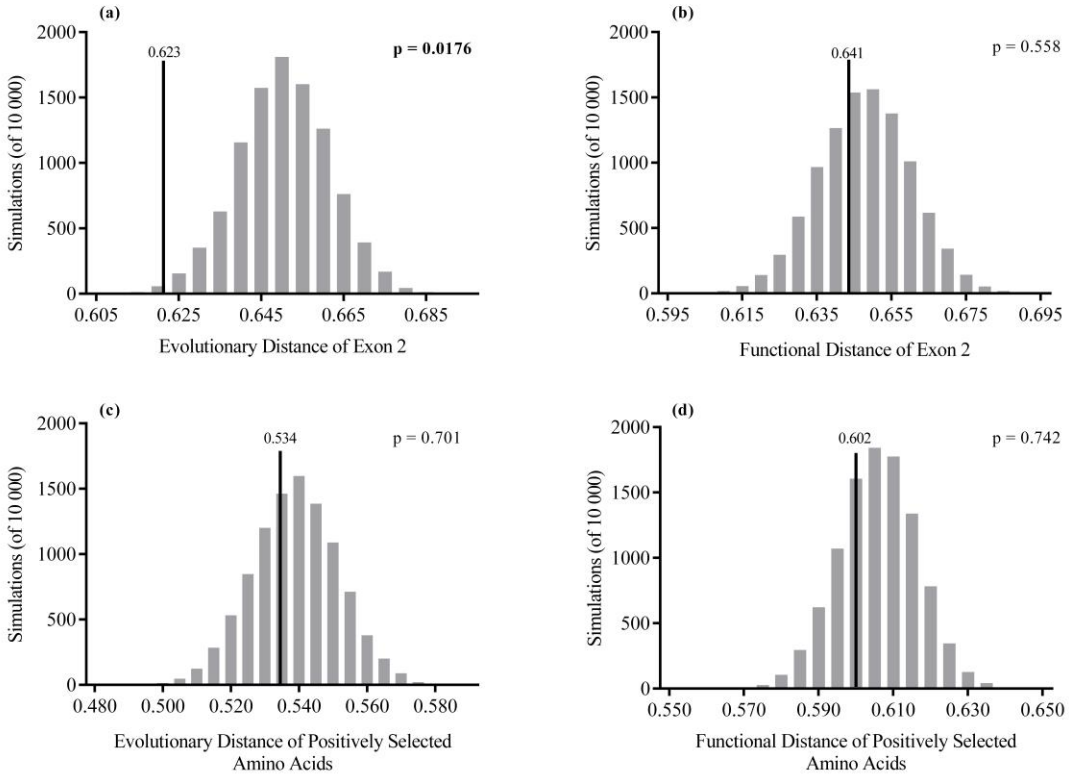


Figure 5.1 Histograms of 10 000 randomized male-female distances of MHC class IIβ. Mean distance of true pairs (i.e., social mates, inferred from behavioural and trapping records) is indicated by a vertical black line. Socially mated pairs were less dissimilar than expected under random mating in terms of evolutionary distance across the entire exon 2, but not in terms of other distance measures.

5.4 Discussion

Evaluating the levels of selection on individual codons within exon 2 (PBR) of MHC class II provides insights for evolutionary processes that may be occurring in the population (e.g., pathogen-mediated selection). These codons encode amino acids that may be involved in direct antigen-recognition and/or are of structural importance to the peptide-binding groove of MHC class II. In my study, I found that the best representative maximum likelihood codon model was the positive selection model (M2a), implying importance for positively selected codons when modeling the evolutionary history of MHC class II exon 2 for my song sparrow population. I found that out of the 12 codons under positive selection, only six encoded for putative peptide-binding amino acids (Brown et al. 1993). This result implies that selection may be acting on codons that encode amino acids which are of structural support for antigen recognition. Alternatively, amino acids that are not putative peptide-binding residues may indeed bind to antigens in song sparrows, as studies model putative peptide-binding amino acids after the human MHC (Brown et al. 1993) and structural changes may have occurred between avian and mammalian lineages.

Offspring from parents with maximally dissimilar MHC genotypes should have maximal MHC diversity, and presumably the ability to recognize more antigens (Klein 1986). Thus, most studies predict disassortative pairing at MHC, as this is generally expected to yield the most disease-resistant offspring (Yamazaki et al. 1976, 1988; Potts et al. 1991; Wedekind and Füri 1997; Landry et al. 2001; Freeman-Gallant et al. 2003; Agbali et al. 2010; Juola and Dearborn 2012; Leclaire et al. 2017). Contrary to this prediction, I did not observe MHC-disassortative pairing in song sparrows, and found that social mates were more, not less, similar at some measures of MHC than predicted by chance. This observation could reflect a mechanism to reduce outbreeding depression and/or reduce the disruption of co-adapted gene complexes (Tregenza and Wedell 2000). Such that, if a female were to pair with a MHC-divergent male, their offspring may have lower fitness since pairing with a divergent mate may disrupt other beneficial loci linked to MHC. Furthermore, pairing with individuals that are highly MHC-dissimilar may disrupt optimal (intermediate) MHC diversity (Bonneaud et al. 2006). Indeed, in my system, song

repertoire size is highest in males with intermediate MHC class II diversity (Slade et al. 2017b); this is consistent with intermediate MHC diversity being optimal because females prefer males with the largest repertoires as mates (Searcy 1984). Additionally, male song sparrow birdsong varies geographically, and female preference for sympatric over allopatric song types (Searcy et al. 2002) may reinforce outbreeding avoidance for this species (Butlin and Tregenza 1997; Tregenza and Wedell 2000).

Similar to my results are those found in tiger salamanders (*Ambystoma tigrinum*, Bos et al., 2009), house sparrows (*Passer domesticus*, Bonneaud et al. 2006), and European badgers (*Meles meles*, Sin et al. 2015). In tiger salamanders, males have a reproductive disadvantage if they are MHC-divergent from females. Female house sparrows tend to pair with males that shared some alleles (Bonneaud et al. 2006). Likewise, European badgers paired assortatively based on the evolutionary and functional distance of MHC class II exon 2 (Sin et al. 2015). These results contradict previous findings of avoiding MHC-similar mates (Yamazaki et al., 1976, 1988; Potts et al. 1991; Freeman-Gallant et al. 2003; Juola and Dearborn 2011, Leclaire et al. 2017), and imply that maximizing MHC diversity may not translate to fitness benefits to offspring for every vertebrate species.

Whereas evolutionary distance of the entire MHC class II exon 2 predicted MHC-mediated mate choice, I did not find that song sparrows paired assortatively or disassortatively based on the positively selected codons or functional distances. In contrast, blue petrels (*Halobaena caerulea*), paired disassortatively based on the functional rather than evolutionary distance at MHC class II exon 2 (Strandh et al. 2012). Furthermore, since I see positive selection on codons directly involved in antigen-binding, but song sparrows do not pair based on these codons, then the observed patterns of molecular evolution at these sites is most likely due to natural selection (e.g., pathogen-mediated selection), not sexual selection (mate choice).

MHC-mediated mate choice has been described in a variety of taxa. However, evidence tends to support disassortative mating in mammals (e.g., Penn and Potts 1998; Jacob et al. 2002), fish (e.g., Landry et al. 2001; Reusch et al. 2001; Eizaguirre et al. 2009), reptiles (e.g., Olsson et al. 2003), and birds (e.g., Freeman-Gallant et al. 2003). I found

positive selection acting on 12 codons within exon 2 of MHC class II. Such variation is likely due to natural selection pressures (pathogen-mediated selection) rather than sexual selection (mate choice), as song sparrows did not pair disassortatively by any measure, and paired assortatively with respect to evolutionary distance of MHC class II exon 2. My results add to a growing body of evidence that MHC-mediated mate choice is not necessarily disassortative (Ruff et al. 2012). Particularly in species in which with highly duplicated MHC, the presence of MHC-assortative mating in my system is likely due to outbreeding avoidance.

5.5 References

Agbali M, Reichard M, Bryjová A, Bryja J, Smith C. 2010. Mate choice for nonadditive genetic benefits correlate with MHC dissimilarity in the rose bitterling (*Rhodeus ocellatus). Evolution*. 64:1683–1696.

Aguilar A, Edwards SV, Smith TB, Wayne RK. 2006. Patterns of variation in MHC class II loci of the little greenbul (*Andropadus virens*) with comments on MHC evolution in birds. *J Hered*. 97:133–142.

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. *J Mol Biol.* 215:403–410.

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

Bonneaud C, Chastel O, Federici P, Westerdahl H, Sorci G. 2006. Complex *Mhc*-based mate choice in a wild passerine. *Proc R Soc B*. 273:1111–1116.

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

Burnham KP, Anderson DR. 2002. *Model selection and multimodel inference: a practical information-theoretic approach*. Springer Science & Business Media: Berlin, Germany.

Butlin RK, Tregenza T. 1997. Is speciation no accident? *Nature.* 387:551–552.

Chen J. 2012. GUniFrac: generalized UniFrac distances. *R Package Version* 1. https://cran.r-project.org/package=GUniFrac

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

Edwards SV, Gasper J, March M. 1998. Genomics and polymorphism of *Agph-DAB1*, an *Mhc* class II B gene in red-winged blackbirds (*Agelaius phoeniceus*). *Mol Biol Evol.* 15:236–250.

Eizaguirre C, Yeates SE, Lenz TL, Kalbe M, Milinski M. 2009. MHC-based mate choice combines good genes and maintenance of MHC polymorphism. *Mol Ecol.* 18:3316–3329.

Felsenstein J. 2005. *PHYLIP (Phylogeny Inference Package).* Department of Genome Sciences, University of Washington, Seattle.

Forsberg L a., Dannewitz J, Petersson E, Grahn M. 2007. Influence of genetic dissimilarity in the reproductive success and mate choice of brown trout – females fishing for optimal MHC dissimilarity. *J Evol Biol*. 20:1859–1869.

Freeman-Gallant CR, Meguerdichian M, Wheelwright NT, Sollecito SV. 2003. Social pairing and female mating fidelity predicted by restriction fragment length polymorphism similarity at the major histocompatibility complex in a songbird. *Mol Ecol.* 12:3077–3083.

Galaverni M, Caniglia R, Milanesi P, Lapalombella S, Fabbri E, Randi E. 2015. Choosy wolves? Heterozygote advantage but no evidence of MHC-based disassortative mating. *J Hered* 107:134–142.

Gloor GB, Hummelen R, Macklaim JM, Dickson RJ, Fernandes AD, MacPhee R, Reid, G. 2010. Microbiome profiling by Illumina sequencing of combinatorial sequence-tagged PCR products. *PloS One.* 5:e15406–e15406.

Huchard E, Baniel A, Schliehe-Diecks S, Kappeler PM. 2013. MHC-disassortative mate choice and inbreeding avoidance in a solitary primate. *Mol Ecol.* 22:4071–4086.

Hughes AL, Hughes MK. 1995. Natural selection on the peptide-binding regions of major histocompatibility complex molecules. *Immunogenetics*. 42:233–243.

Hughes AL, Nei M. 1989. Nucleotide substitution at major histocompatibility complex class II loci: evidence for overdominant selection. *Proc Natl Acad Sci*. 86:958–962.

Hughes AL, Nei M. 1992. Models of host-parasite interaction and *Mhc* polymorphism. *Genetics*. 132:863–864.

Jacob S, McClintock MK, Zelano B, Ober C. 2002. Paternally inherited HLA alleles are associated with women's choice of male odor. *Nat Genet.* 30:175–179.

Juola FA, Dearborn DC. 2012. Sequence-based evidence for major histocompatibility complex-disassortative mating in a colonial seabird. *Proc R Soc B*. 279:153–162.

Kasahara M. 2000. *Major histocompatibility complex: evolution, structure, and function*. Springer: Tokyo, Japan.

Klein J. 1986. *Natural history of the major histocompatibility complex*. John Wiley & Sons: New York, NY, USA.

Kubinak JL, Nelson CA, Ruff JS, Potts WK. 2012. Tradeoffs limiting MHC heterozygosity. In *Ecoimmunology*, pp. 225-258. Oxford University Press, Inc.: New York, NY, USA.

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

Laitinen J, Samarut J, Hölttä E. 1994. A nontoxic and versatile protein salting-out method for isolation of DNA. *Biotechniques*. 17:316–318.

Landry C, Garant D, Duchesne P, Bernatchez L. 2001. 'Good genes as heterozygosity': the major histocompatibility complex and mate choice in Atlantic salmon (*Salmo salar*). *Proc R Soc Lond B.* 268:1279–1285.

Leclaire S, Strandh M, Mardon J, Westerdahl H, Bonadonna F. 2017. Odour-based discrimination of similarity at the major histocompatibility complex in birds. *Proc R Soc B*. 284:20162466.

Lozupone C, Knight R. 2005. UniFrac: a new phylogenetic method for comparing microbial communities. *Appl Environ Microbiol.* 71:8228–8235.

Manly B. 1997. *Randomization, bootstrapping and Monte Carlo methods in biology*. 2nd edn. Chapman & Hall: London, UK.

Mays HL, Hill GE. 2004. Choosing mates: good genes versus genes that are a good fit. *Trends Ecol Evol*. 19:554–559.

McClelland EE, Penn DJ, Potts WK. 2003. Major histocompatibility complex heterozygote superiority during coinfection. *Infect Immun.* 71:2079–2086.

Neff BD, Pitcher TE. 2005. Genetic quality and sexual selection: an integrated framework for good genes and compatible genes. *Mol Ecol.* 14:19–38.

Oliver MK, Telfer S, Piertney SB. 2009. Major histocompatibility complex (MHC) heterozygote superiority to natural multi-parasite infections in the water vole (*Arvicola terrestris*). *Proc R Soc B.* 276:1119–1128.

Olsson M, Madsen T, Nordby J, Wapstra E, Ujvari B, Wittsell H. 2003. Major histocompatibility complex and mate choice in sand lizards. *Proc R Soc Lond B* (Suppl.). 270:S254–S256.

Penn D, Potts W. 1998. How do major histocompatibility complex genes influence odor and mating preferences. *Adv Immunol*. 69:411–436.

Piertney SB, Oliver MK. 2006. The evolutionary ecology of the major histocompatibility complex. *Heredity*. 96:7–21.

Potts WK, Manning CJ, Wakeland EK. 1991. Mating patterns in seminatural populations of mice influenced by MHC genotype. *Nat Lond.* 352:619–21.

Potvin DA, MacDougall-Shackleton EA. 2009. Parental investment amplifies effects of genetic complementarity on growth rates in song sparrows, *Melospiza melodia*. *Anim Behav.* 78:943–948.

R Core Team. 2017. R: a language and environment for statistical computing. R Foundation for Statistical Computing: Vienna, Austria. URL https://www.r-project.org/

Radwan J, Tkacz A, Kloch A. 2008. MHC and preferences for male odour in the bank vole. *Ethology*. 114:827–833.

Reusch TB, HaÈberli MA, Aeschlimann PB, Milinski M. 2001. Female sticklebacks count alleles in a strategy of sexual selection explaining MHC polymorphism. *Nature*. 414:300–302.

Ruff JS, Nelson AC, Kubinak JL, Potts WK. 2012. MHC signaling during social communication. In *Self and Nonself*, pp 290–313. Advances in Experimental Medicine and Biology Vol 738. Springer Science & Business Media: Berlin, Germany.

Sandberg M, Eriksson L, Jonsson J, Sjöström M, Wold S. 1998. New chemical descriptors relevant for the design of biologically active peptides. A multivariate characterization of 87 amino acids. *J Med Chem.* 41:2481–2491.

Searcy WA. 1984. Song repertoire size and female preferences in song sparrows. *Behav Ecol Sociobiol*. 14:281–286.

Searcy WA, Nowicki S, Hughes M, Peters S. 2002. Geographic song discrimination in relation to dispersal distances in song sparrows. *Am Nat*. 159:221–230.

Setchell JM, Charpentier MJE, Abbott KM, Wickings EJ, Knapp LA. 2010. Opposites attract: MHC-associated mate choice in a polygynous primate. *J Evol Biol.* 23:136–148.

Slade RW, McCallum HI. 1992. Overdominant vs. frequency-dependent selection at MHC loci. *Genetics*. 132:861–862.

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

Slade JWG, Watson MJ, Kelly TR, Gloor GB, Bernards MA, MacDougall-Shackleton EA. 2016. Chemical composition of preen wax reflects major histocompatibility complex similarity in songbirds. *Proc R Soc B.* 283:20161966.

Slade JWG, Watson MJ, MacDougall-Shackleton EA. 2017. Birdsong signals individual diversity at the major histocompatibility complex. *Biol Lett*. 13:20170430.

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

Strandh M, Westerdahl H, Pontarp M, Canback 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. *Proc R Soc B.* 279:4457–4463.

Sutton JT, Castro I, Robertson BC, Tompkins DM, Stanton J-AL, Jamieson IG. 2016. MHC genetic diversity and avian malaria prevalence in Mokoia Island saddlebacks. *NZ J Ecol*. 40:0–0.

Tregenza T, Wedell N. 2000. Genetic compatibility, mate choice and patterns of parentage: invited review. *Mol Ecol.* 9:1013–1027.

Trowsdale J. 1995. 'Both man & bird & beast': comparative organization of MHC genes. *Immunogenetics*. 41:1–17.

Wedekind C, Füri S. 1997. Body odour preferences in men and women: do they aim for specific MHC combinations or simply heterozygosity? *Proc R Soc Lond B.* 264:1471–1479.

Wedekind C, Seebeck T, Bettens F, Paepke AJ. 1995. MHC-dependent mate preferences in humans. *Proc R Soc Lond B.* 260:245–249.

Wegner KM, Kalbe M, Kurtz J, Reusch TBH, Milinski M. 2003. Parasite selection for immunogenetic optimality. *Science*. 301:1343.

Whelan S, Goldman N. 2001. A general empirical model of protein evolution derived from multiple protein families using a maximum-likelihood approach. *Mol Biol Evol.* 18:691–699.

Xu B, Yang Z. 2013. PAMLx: a graphical user interface for PAML. *Mol Biol Evol.* 30:2723–2724.

Yamazaki K, Beauchamp GK, Kupniewski D, Bard J, Thomas L, Boyse EA. 1988. Familial imprinting determines H-2 selective mating preferences. *Science.* 240:1331–1333.

Yamazaki K, Boyse EA, Miké V, Thaler HT, Mathieson BJ, Abbott J, Boyse J, Zayas ZA, Thomas L. 1976. Control of mating preferences in mice by genes in the major histocompatibility complex. *J Exp Med.* 144:1324–1335.

Yamazaki K, Yamaguchi M, Baranoski L, Bard J, Boyse EA, Thomas L. 1979. Recognition among mice. Evidence from the use of a Y-maze differentially scented by congenic mice of different major histocompatibility types. *J Exp Med*. 150:755–760.

Yang Z. 2007. PAML 4: phylogenetic analysis by maximum likelihood. *Mol Biol Evol*. 24:1586–1591.

Yang Z, Swanson WJ. 2002. Codon-substitution models to detect adaptive evolution that account for heterogeneous selective pressures among site classes. *Mol Biol Evol*. 19:49–57.

Yang Z, Wong WSW, Nielsen R. 2005. Bayes empirical Bayes inference of amino acid sites under positive selection. *Mol Biol Evol.* 22:1107–1118.

Chapter 6

6 General Discussion

6.1 Summary of findings

In this dissertation, I addressed the over-arching hypothesis that MHC genotype affects disease resistance of song sparrows (*Melospiza melodia*) through additive (locally-good genes) and/or non-additive (heterozygote advantage) genetic effects; that the chemical composition of preen oil and the content of learned song repertoires provide information regarding MHC genotype dissimilarity and/or diversity; and that free-living song sparrows pair disassortatively with respect to MHC genotype. In Chapter 2, I assessed the importance of additive (e.g., locally-protective alleles) and non-additive (e.g., heterozygote advantage) genetic effects at MHC to explain enhanced resistance to sympatric parasites previously described for this species (Sarquis-Adamson and MacDougall-Shackleton 2016). Because natural selection operating at MHC may influence sexual selection, in Chapters 3 and 4 I explored potential mechanisms by which songbirds might signal MHC profiles to potential mates or other conspecifics. Specifically, I tested whether the chemical composition of preen wax varies with MHC profile, which would suggest a potential role for olfaction in MHC-mediated mate choice. Because learned birdsong is an important sexually selected trait in this species, I also addressed whether song type sharing, and song complexity, might signal MHC similarity and diversity, respectively. Finally, in Chapter 5 I investigated whether free-living song sparrows pair disassortatively based on MHC genotypes. This analysis quantified MHC similarity at two scales: over the entire second exon of MHC class II and at codons within exon 2 which I identified as showing the signature of positive molecular selection.

6.1.1 Additive and non-additive genetic effects at MHC on *Plasmodium* resistance

Additive genetic effects at MHC have been experimentally documented, and have been shown to be associated with single-strain infection (e.g., Lohm et al. 2002), however, a less explored area of additive MHC fitness benefits are locally-good genes at MHC. Genetic

differentiation between populations is likely to occur if there is a selective advantage for choosing locally-adapted mates (Reinhold 2004). Therefore, given that geographic variation in parasite assemblages occur (Pagenkopp et al. 2008), then there should be selection for geographically-specific MHC alleles across populations (Loiseau et al. 2009). Also, non-additive genetic effects, such as heterozygote advantage at MHC, may confer fitness benefits against a single-strain infection (e.g., Westerdahl et al. 2005). Therefore, both mechanisms can increase MHC diversity across populations via diversifying selection (e.g., Loiseau et al. 2009) and balancing selection (e.g., Kubinak et al. 2012).

Sarquis-Adamson and MacDougall-Shackleton (2016) experimentally demonstrated that song sparrows separated by 440 km are more resistant to a sympatric *Plasmodium* lineage than to an allopatric *Plasmodium* lineage. In chapter 2, I tested whether genetic differentiation at either MHC class I exon 3 or MHC class IIβ might explain the apparent local adaptation to sympatric strains of *Plasmodium*. I characterized both classes of MHC and recovered up to four class I alleles and 26 class II alleles per individual. However, I did not find any population genetic differentiation at either class of MHC. This suggests that ecological processes (i.e., prior experience with local parasite strains) rather than evolutionary processes (i.e., locally-protective alleles at MHC) may best explain song sparrows' enhanced resistance to local *Plasmodium* reported by Sarquis-Adamson and MacDougall-Shackleton (2016). Prior exposure to the local or similar strain is expected to result in immune memory, potentially reducing the infection intensity or overall infection risk (Møller and Szép 2011; Krzych et al. 2014; Sarquis-Adamson and MacDougall-Shackleton 2016). Furthermore, other immune genes such as β-defensins and/or toll-like receptors may vary geographically and better explain the observed resistance to local *Plasmodium* strains.

Whereas geographic variation at MHC failed to explain Sarquis-Adamson and MacDougall-Shackleton's (2016) results, overall allelic diversity at MHC class I exon 3 did appear to influence song sparrows' infectivity risk. Individuals that were more diverse at MHC class I exon 3 were less likely to become infected with *Plasmodium*, indicating a parasite-mediated heterozygote advantage at class I. Therefore, although I found no

evidence of population genetic structuring at MHC that would indicate spatial variation in selection pressures, balancing selection in the form of heterozygote advantage is likely to maintain MHC class I diversity within each population.

6.1.2 Signals of MHC genotype

In chapter 3, I explored whether pairwise chemical distances of preen wax reflected pairwise genetic distances at MHC class IIβ. I found that for male-female dyads, the comparison most salient to mate choice, pairwise distance in preen wax composition varied positively with amino acid distance of MHC class IIβ. This relationship suggests that preen wax is a potential cue of MHC similarity that may be used during cross-sex interactions. Although the explanatory power of this relationship is low (r = 0.111), yet statistically significant (p = 0.002), it is difficult to estimate whether the potential signal from preen wax is biologically important. However, given that diet can influence preen wax composition (Thomas et al. 2010), and that not all gDNA MHC alleles may be translated, then it would be beneficial to reduce noise by studying this relationship in a controlled environment (e.g., standardized diet), whereby cDNA MHC alleles are considered. Overall, this finding demonstrates that avian chemical secretions do covary with MHC genotype: such a finding has the potential to considerably expand our understanding of sexual signalling in birds, a taxon previously assumed to communicate exclusively through visual and auditory signals.

MHC has been found to vary geographically in a variety of taxa (e.g, Rodríguez et al. 2011). This raises the possibility that geographically variable cues such as birdsong may also signal MHC genotype, in so far as both birdsong and MHC vary geographically. In song sparrows, song varies geographically (Peters et al. 2000), and more locally-typical song types are correlated with locally-typical neutral locus (i.e., microsatellite) genotypes (Stewart and MacDougall-Shackleton 2008). Therefore, I hypothesized that geographic variation in birdsong might signal variation in MHC. In Chapter 4, I examined whether syllable sharing across male song sparrows is associated with MHC class IIβ similarity. Contrary to my expectation, I found that syllable sharing was not associated with similarity at MHC. Therefore, even though birdsong reliably indicates population of origin, my

findings potentially support results in chapter 2, whereby I found a lack of population genetic structuring at MHC class IIβ.

6.1.3 Signals of MHC diversity

Although most research on MHC-mediated mate choice has addressed preferences for MHC-dissimilar mates, in many systems (particularly those in which males provide extensive material benefits to females and offspring, as is the case in song sparrows) assessing the overall MHC diversity of a potential mate may also be favoured because mates with optimal MHC diversity are presumably healthier and may provide superior material benefits. In Chapter 4, I investigated whether male song complexity (i.e., repertoire size) in song sparrows is associated with MHC diversity. Unexpectedly, I found that MHC diversity varied non-linearly with song complexity, such that males with intermediate numbers of alleles at MHC class IIβ had the largest song repertoires. This is the first study to demonstrate a learned courtship signal being associated with MHC diversity. It also demonstrates that for song sparrows, as in other taxa such as three-spined sticklebacks (Milinski 2003; Wegner et al. 2003, 2004), optimal diversity at class II MHC may involve intermediate rather than maximal numbers of alleles. This finding is in contrast to patterns seen in rodents (McClelland et al. 2003; Oliver et al. 2009), salmonids (Landry et al. 2001; Evans and Neff 2009) and another songbird species (Dunn et al. 2013; Whittingham et al. 2015), where pathogen-mediated selection appears to favour maximal rather than intermediate diversity at MHC. Because male song sparrows learn their songs within their first year of life, my results imply that early-life condition in general, and brain development in particular, may reflect a balance between pathogen resistance and the disadvantages of being too diverse at MHC.

6.1.4 Selection and mate choice at MHC

Theory suggests that jawed vertebrates should prefer mates with MHC alleles dissimilar to their own in order to maximize or optimize (Kubinak et al. 2012; Ruff et al. 2012) the MHC diversity of their offspring. In Chapter 5, I tested for positive selection at the peptide-binding region (PBR) of MHC class II (β chain) in song sparrows, and whether these birds socially paired disassortatively based on their MHC class IIβ genotypes. I found 12

positively selected codons within the PBR, half of them belonging to putative peptide-binding codons (Brown et al. 1993). This reflects that the PBR of song sparrows in my populations are most likely under balancing selection, which diversifies the range of pathogens they can detect (Piertney and Oliver 2006). However, observed patterns of social mate choice, a proxy for genetic mate choice (i.e., low level of extra-pair young in my population; Potvin and MacDougall-Shackleton 2009), revealed that song sparrows paired with individuals that are more MHC-similar than expected by chance (i.e., assortative rather than disassortative mating).

6.2 New Insights

6.2.1 Immunity and MHC

Additive and non-additive genetic effects at MHC on immunity have been documented in various vertebrate taxa (e.g., Langefors et al. 2001; Madsen and Ujvari 2006; Pitcher and Neff 2006; Oliver et al. 2009). Theory predicts that non-additive effects (heterozygote advantage or selection favouring maximal diversity) at MHC should be most pronounced when encountering multi-strain infections (Kubinak et al. 2012). For example, in water voles (*Arvicola terrestris*), individuals that were more diverse at MHC class IIβ had fewer ectoparasites (Oliver et al. 2009). Conversely, resistance to infection by a single pathogen strain should be mediated primarily by additive rather than non-additive effects, such that we might expect to find one 'best' allele at MHC conferring maximal protection against a specific pathogen. Supporting this, in an experiment on Atlantic salmon (*Salmo salar*), heterozygous individuals did not have a fitness advantage against a bacterial infection of *Aeromonas salmonicida*, but individuals carrying a particular allele had a fitness advantage of up to 49%, implying an additive genetic effect (Lohm et al. 2002). Indeed, there are a variety of studies that showcase additive and non-additive effects at MHC against infection (table 6.1). In chapter 2, I found there was a heterozygote advantage at MHC class I exon 3 against single-strain *Plasmodium* infections in song sparrows. In birds, both good genes and heterozygote advantage may influence resistance to avian malaria. In great tits (*Parus major*), the presence of a specific MHC supertype conferred protection against a specific *Plasmodium* strain (Sepil et al. 2013). However, supporting my findings, in great reed

warblers (*Acrocephalus arundinaceus*), being more diverse at MHC class I exon 3 provided protection against the GRW2 *Plasmodium* strain (Westerdahl et al. 2005). Additionally, even a single strain of *Plasmodium* has the capability of producing multiple antigens thereby helping it to evade the host's immune system (Singh et al. 2014). Therefore, it may be possible that being diverse at MHC confers protection against multiple antigens from a single *Plasmodium* strain.

6.2.2 Diversity at MHC

MHC diversity varies greatly among vertebrate taxa (table 6.1). Many early studies used restriction fragment length polymorphism or single strand conformation polymorphism to identify MHC alleles. Although these tools were sufficient to generate an estimate of allelic richness at MHC loci, it was difficult to look at fine-scale variation at the nucleotide level, or to estimate the number of putative loci at either MHC class. These older approaches may have greatly underestimated the number of loci at MHC for many species, and may explain the vast differences between similar taxa (e.g., mice vs. voles) (table 6.1) studied using older techniques instead of DNA sequencing.

In song sparrows, I observed a maximum of four MHC class I alleles per individual implying at least two class I loci (chapter 2). However, class IIβ was much more diverse (chapter 2, 3, 4, and 5), with up to 26 alleles recovered from a single individual: this indicates that song sparrows have at least 13 MHC class II loci. In birds, diversity at MHC class I and class II varies substantially within and between taxa (table 6.1). In general, Galliformes tend to have low MHC diversity (Hess and Edwards 2002). For example, golden pheasants (*Chrysolophus pictus*) only have a maximum of two alleles (one locus) per individual at MHC class I (Zeng et al. 2016). Furthering this, greater prairie chickens (*Tympanuchus cupido*) have minimal MHC class I (one locus), and relatively low class II diversity (two loci) (Bateson et al. 2015). Supporting my findings, other songbird species tend to have high MHC diversity (Hess and Edwards 2002, table 6.1). For example, an early study using gene cloning showed that great reed warblers have a minimum of four loci at both MHC class I (Westerdahl et al. 1999) and class II (Westerdahl et al. 2000). With the advent of technological advances such as next-generation sequencing, MHC class

II has been further characterized in songbirds, revealing extensive diversity. For instance, using high-throughput sequencing revealed high MHC class II diversity in common yellowthroats (20 putative loci, Bollmer et al. 2010), New Zealand saddlebacks (*Philesturnus* spp., 10 putative loci, Sutton et al. 2013), and little greenbuls (*Andropadus viren*, seven loci from cDNA, Aguilar et al. 2006). Thus, there is evidence for high MHC diversity in songbirds, and it is speculated that this diversity occurred through recent duplication events (Hess and Edwards 2002), and may be maintained through heterozygote advantage, antagonistic co-evolution, and MHC-mediated mate choice (Kubinak et al. 2012).

6.2.3 Signals of MHC genotype

Research on mice (Yamazaki et al. 1976) and fish (Milinski et al. 2005) revealed that signals of MHC genotypes could be sent through urine and detected by conspecifics. The information contained in the chemical signal was found to mediate mate choice decisions (Penn and Potts 1998). Chemical signals of MHC genotype were found mainly in mammals (including humans), sand lizards (*Lacerta agilis*), and three-spine sticklebacks (table 6.1). However, while chemical signals of MHC were studied in mammals and fish, research on signals of MHC genotype in birds remained unexplored. This is most likely due to Bang and Cobb's (1968) review on olfactory bulb sizes across avian taxa. The study revealed that seabirds had the largest olfactory bulb to brain size (25%), while songbirds had the smallest (5%). Since then, scientists assumed that seabirds had high olfactory senses and songbirds were either microsmatic or anosmic. Indeed, seabirds are macrosmatic (Nevitt 2000), and research on blue petrels (*Halobaena caerulea*) reveal that they pair disassortatively based on MHC class II (Strandh et al. 2012), and that these signals are chemical (Leclaire et al. 2017). This research provides support for preen wax to be a semiochemical that signals MHC genotype to conspecifics. Chapter 3 provides the first indicator that songbirds may chemically signal MHC genotype, as I found that the chemical composition of preen wax covaries with MHC amino acid distances. Along with my results, and the relationship between preen wax and MHC in blue petrels and black-legged kittiwakes (*Rissa tridactyla*, Leclaire et al. 2014), the field of chemical communication has the potential to greatly expand.

6.2.4 Signals of MHC diversity

Signals of MHC diversity are found predominately in sexually selected signals, such as plumage (table 6.1). For example, peacocks (*Pavo cristatus*) with longer tails had greater MHC diversity (Hale et al. 2009). Likewise, MHC maximal diversity was associated plumage quality in common yellowthroats (*Geothlypis trichas*, Dunn et al. 2013; Whittingham et al. 2015). By expanding the field of honest signals of genetic quality by using candidate loci, researchers can obtain a better understanding of how ornamentation signals genetic quality (Mays and Hill 2004). Chapter 4 is the first study to show how a learned courtship signal (birdsong) is related to MHC diversity, indicating how MHC diversity could affect song learning during early life. Furthermore, my study adds to the small evidence that intermediate (four studies, table 6.1) rather than maximal MHC diversity may increase fitness.

6.2.5 Positive selection at the peptide-binding region (PBR) of MHC

The MHC comprises the most polymorphic genes in jawed vertebrates (Klein 1986). The main cause of this polymorphism is thought to be the intense selective pressure that pathogens place on their hosts (Borghans et al. 2004). Pathogen-mediated balancing selection and/or diversifying selection leads to the evolution of new MHC alleles at both class I (Prugnolle et al. 2005; Loiseau et al. 2009), and class II (Alcaide et al. 2008; Hawley and Fleischer 2012), which, combined with gene duplication at these loci (Bollmer et al. 2010; O'Connor et al. 2016), diversifies the MHC. The peptide-binding region (PBR) of each class of MHC is the contact area between antigens and MHC, and the MHC-bound antigen gets presented to T cells to elicit an immune response (Trowsdale 2011). Compared to other areas of MHC genes, the codons in the PBR are under the strongest positive selective pressure to combat emerging pathogens and enhance the adaptive immunity of their host (Hughes and Hughes 1995). When comparing codons between alleles within this region (exon 2 and 3 for MHC class I, and exon 2 for class II) within and between species, there are high levels of nonsynonymous base substitutions over synonymous ones (table 6.1). In most vertebrates, whereby researchers have explored MHC-mediated mate choice and/or MHC fitness relationships, studies also show positive selection at the PBR (table

6.1). Many of these studies compare the ratio of nonsynonymous (dN) to synonymous (dS) substitutions, and infer positive selection when dN/dS >1. However, theory suggests that codons which encode amino acids that bind directly to antigens should have higher levels of dN than codons that encode amino acids that do not make contact with antigens. Indeed, in chapter 5, I found that six of the codons that encode amino acids which make direct contact with antigens in the PBR of song sparrow MHC class II showed signs of positive selection. However, these codons were inferred by Brown et al (1993), who mapped the amino acids that make direct contact with antigens for MHC class II in humans. This is one of the best proxies to infer peptide-binding amino acids, but it is possible that the other six codons do encode amino acids that do directly contact antigens in song sparrows, or have structural importance to support other peptide-binding amino acids. Ideally, crystallization of MHC proteins in other taxa would shine light on the variation in peptide-binding sites between vertebrates.

6.2.6 MHC-mediated mate choice

MHC mating preferences vary across taxa (table 6.1). MHC-disassortative pairing appears to be the most dominant pattern observed (50% of species, able 6.1), with MHC-assortative mating being the second most prevalent pattern observed (13.8% of species, table 6.1). Other MHC mate preferences incorporate choosing a mate based on maximum MHC diversity (11.5% of species, table 6.1), and optimal diversity, whereby the animal chooses a mate that has intermediate MHC diversity (3.8% of species, table 6.1). Finally, some animals appear to pair randomly with respect to MHC (19.2% of species, table 6.1).

MHC-disassortative pairing corroborates evidence of heterozygote advantage at MHC. Individuals that pair with MHC-dissimilar mates will produce offspring that are more diverse at MHC, thus giving them the ability to fight off a wide range of pathogens (Neff and Pitcher 2005); this mating pattern seems to occur for most vertebrates where MHC mate preference was studied. Mating systems and MHC diversity may have an influence on MHC-disassortative pairing. For example, in strictly monogamous animals (such as many seabirds), the female cannot choose an extra-pair mate if they are paired with a male of poor genetic quality. Therefore, monogamous animals that have few MHC

loci should particularly benefit by pairing with MHC-dissimilar mates. For example, blue petrels (*Halobaena caerulea*) have only two MHC class II loci (Strandh et al. 2012). Thus, it would be advantageous for these monogamous seabirds to choose a mate that is compatible at MHC to increase pathogen resistance for their offspring.

MHC-assortative pairing can pose a paradox, since the resulting offspring should be less diverse than average at MHC, and thus would not benefit from heterozygote advantage. Possible explanations for this phenomenon include avoiding outbreeding depression, especially if the animals have local adaptations to the environment (Tregenza and Wedell 2000). Avoiding divergent genotypes may also be due to retaining co-adapted gene complexes. It is possible that MHC is linked to other beneficial loci within the genome, and by choosing divergent mates, these co-adapted gene complexes would be disrupted, and thus produce offspring with lower fitness. In chapter 5, I found that song sparrows pair assortatively based on the evolutionary distances of the PBR at MHC class II. The evolutionary history at this class of MHC in my population may be important to other adaptive loci within the song sparrow.

6.2.7 MHC diversity, signalling, mate choice and parasite-mediated selection in other species: the need for an integrated approach

In the five classes surveyed (mammals, birds, reptiles, amphibians, and fishes), there was little consistency between the level of MHC diversity in closely related taxa, except for salmonids (table 6.1). Also, there does not seem to be a relationship between MHC diversity, direction of MHC mate preference, or fitness relationship (table 6.1). However, not all studies (including this thesis) have determined what proportion of alleles in the genome are expressed. This could overestimate the number of loci present in each animal. Furthermore, many alleles are functionally similar (Sepil et al. 2013) and to treating each allele as if its function is unique to other alleles in the species may overestimate MHC functional diversity in a population.

The majority of studies investigating signals of MHC revealed that MHC genotype is signalled via olfactory (chemical) cues (table 6.1). These studies were found mainly in

mammals, seabirds, and in one reptile, which are believed to have a good sense of smell (Müller-Schwarze 2006). Interestingly, the majority of these studies also found MHC-disassortative mating preferences, implying heterozygote advantage. In contrast, my study showed the opposite trend, whereby I found a candidate chemical cue of MHC genotype (chapter 3), but found these birds pair assortatively based on MHC class IIβ (chapter 5). Supporting MHC-assortative pairing are studies on two mammals, one songbird, and a salamander (table 6.1). These studies propose MHC-assortative pairing as a mechanism to reduce outbreeding depression and/or prevent the disruption of co-adapted gene complexes.

MHC is of key selective importance since there is high non-synonymous base substitutions over synonymous ones at the PBR. I found 100% of the studies that investigated selection at MHC found positive selection at the PBR (table 6.1), implying pathogen-mediated selection. Out of the 29 species surveyed in table 6.1, 15 studies showed a benefit for being diverse at MHC. The majority of these studies analyzed heterozygote advantage against multi-strain infections (except Westerdahl et al. 2005), however, not all studies measured defense against pathogens. For example, female Megallanic penguins (*Spheniscus magellanicuas*) that had maximal MHC diversity hatched more eggs than homozygotes (Knafler et al 2012). This was most likely a proxy for general health, and those females diverse at MHC were less likely to become infected, and more able to provide direct benefits. In the 15 species that showed heterozygote advantage at MHC, only six paired disassortatively at MHC (table 6.1). One would expect that if heterozygote advantage is widespread at MHC, then the majority of jawed vertebrates would pair disassortatively to maximize MHC diversity in their offspring. Even though some studies do not show heterozygote advantage and MHC-disassortative mate preference for a species, it could be possible that tests against multi-strain infection would reveal a heterozygote advantage. Also, the two songbird species that showed a heterozygote advantage against infection but lacked MHC-mediated mate choice, may pair randomly at MHC because they are already highly diverse at these loci (table 6.1), which may still provide offspring with sufficient MHC diversity.

Three studies showed a fitness advantage for having intermediate MHC diversity (table 6.1), and were not specific to any taxonomic class. Bank voles (*Clethrionomys glareolus*, Kloch et al. 2010) and three-spine sticklebacks (*Gasterosteus aculeatus*, Wegner et al. 2003) with intermediate MHC diversity had fewer parasites than those with low or high MHC diversity. Also, house sparrow (*Passer domesticus*) females with intermediate MHC diversity had larger clutch sizes (Bonneaud et al. 2004). Supporting fitness benefits for intermediate MHC diversity, male song sparrows in my system with intermediate MHC class IIβ diversity had the largest song repertoires (chapter 4). By having too few MHC alleles, individuals may be prone to infection as they do not have the MHC molecules to fight off emerging pathogens. Multiple hypotheses exist for being too diverse at MHC. For example, when an individual is too diverse at MHC, only a certain amount of T cells survive during development, and the missing T lymphocytes would not be able to respond to the many antigens presented by multiple MHC variants, but those with an intermediate level of MHC diversity may produce the appropriate matching T lymphocytes (De Boer and Perelson 1993; Milinski 2006). Another explanation is that autoimmune disorders can develop (e.g., Wegner et al. 2003, Trowsdale 2011). The majority of research on MHC-related autoimmune disorders have been studied in humans (Trowsdale 2011), therefore, it would be beneficial to study MHC-related autoimmune disorders in other organisms to understand how being too diverse at MHC could negatively impact fitness.

A taxonomic bias is evident in MHC-associated studies. Mammals, birds, and fish dominate studies exploring MHC-mediated mate choice and MHC fitness relationships (table 6.1). Based on my literature review (table 6.1), it is evident that reptiles and amphibians are highly underrepresented on studies investigating the role of MHC on mate choice and fitness. To fully understand the role of adaptive immune genes in jawed vertebrates, future studies need to focus on these underrepresented taxa.

Finally, because there is a lack of consistent patterns between allele number, mate choice, and fitness, then the evolutionary and ecological processes at MHC are more complex than theorized. Furthermore, there are many animals whereby information on MHC signals, MHC-mediated mate choice, and MHC fitness components are missing, but

the genes have been characterized. Thus, to fully understand the role that MHC plays on the evolutionary fitness in a species, more research must be done in these areas of evolutionary ecology.

Table 6.1 Subset of MHC studies highlighting the putative number of MHC loci, how MHC is signaled for mate choice, MHC mating preferences, MHC fitness relationship, and whether the peptide binding region (PBR) of the MHC class is undergoing positive selection.

Species	MHC class	Putative no. of MHC loci	MHC signal	MHC mating preference	MHC fitness relationship	Positive selection at the PBR?
Mammals						
House mouse (*Mus musculus domesticus)*	I	>30 (Kumánovics and Lindahl 2004)	Chemical (genotype) (Leinders-Zufall 2004)	Disassortative (Yamazaki et al. 1976; Yamazaki et al. 1988)	Maximal diversity against multi-strain infection (Penn et al. 2002)	Yes (Hughes et al. 1990)
Bank vole (*Clethrionomys glareolus*)	II	4 (Kloch et al. 2010)	Chemical (genotype) (Radwan et al. 2008)	Disassortative (Radwan et al. 2008)	Intermediate diversity against parasite infection (Kloch et al. 2010)	Yes (Axtner & Sommer 2007)
Water vole (*Arvicola terrestris*)	II	1 (Oliver and Piertney 2006)	Unknown	Unknown	Maximal diversity against multi-strain ectoparasite infection (Oliver et al. 2009)	Yes (Oliver and Piertney 2006)
Malagasy giant jumping rat (*Hypogeomys antimena*)	II	1 (Sommer et al. 2002)	Unknown	Assortative (Sommer 2005)	Unknown	Yes (Sommer 2003)
European badgers (*Meles meles*)	II	1 (Sin et al. 2012)	Unknown	Assortative (Sin et al. 2015)	Maximal diversity against parasite infection (Sin et al. 2014)	Yes (Sin et al. 2012)
Fat-tailed dwarf lemur (*Cheirogaleus medius*)	II	2 (Schwensow et al. 2008)	Chemical (genotype) (Schwensow et al. 2008)	Disassortative (Schwensow et al. 2008)	Rare alleles against parasite infection (Schwensow et al. 2007)	Yes (Schwensow et al. 2007)
Mandrill (*Mandrillus sphinx*)	II	6 (Abbott et al. 2006)	Chemical (genotype) (Setchell et al. 2011)	Disassortative (Setchell et al. 2010)	Good alleles cause increased male red facial coloration (Setchell et al. 2009)	Yes (Abbott et al. 2006)

Human (*Homo sapiens*)	I, II	6 (I) 11 (II) (Shiina et al. 2004)	Chemical (genotype) (Wedekind et al. 1995; Jacob et al. 2002; Milinski et al. 2013)	Disassortative (Wedekind et al. 1995; Jacob et al. 2002)	Complicated. Heterozygote advantage and disadvantage. Good alleles, and bad alleles (autoimmune diseases) (Trowsdale 2011)	Yes (I) (Hughes et al. 1990) Yes (II) (Salamon et al. 1999)
Birds						
Red jungle fowl (*Gallus gallus*)	I, II	2 (Worley et al. 2008)	Unknown	Disassortative (cryptic) (Gillingham et al. 2009)	Maximal diversity against coccidosis; (Worley et al. 2010)	Yes (Worley et al. 2008)
Ring-necked pheasant (*Phasianus colchicus*)	II	2 (Baratti et al. 2012)	Unknown	Intermediate MHC similarity (Baratti et al. 2012)	Good alleles associated with male spur length (von Schantz et al. 1997)	Yes (Baratti et al. 2012)
Peafowl (*Pavo cristatus*)	II	3 (Hale et al. 2009)	Visual (male tail length; diversity) (Hale et al. 2009)	Maximal diversity (Hale et al. 2009)	Maximal diversity increases immune response, egg weight, and clutch size (Hale et al. 2009)	Unknown
Great snipe (*Gallinago media*)	II	2 (Ekblom et al. 2004)	Unknown	None (Ekblom et al. 2004)	Males with good alleles likely to reproduce (Ekblom et al. 2004)	Yes (Ekblom et al. 2010)
Blue petrel (*Halobaena caerulea*)	II	2 (Strandh et al. 2011)	Chemical (genotype) (Leclaire et al. 2017)	Disassortative (Strandh et al. 2012; Leclaire et al. 2017)	Unknown	Yes (Strandh et al. 2011)
Black-legged kittiwake (*Rissa tridactyla*)	II	2 (Leclaire et al. 2014)	Chemical (genotype) (Leclaire et al. 2012; Leclaire et al. 2014)	Disassortative (implied) (Leclaire et al. 2014)	Unknown	Yes (Leclaire et al. 2014)
Magellanic penguin (*Spheniscus magellanicus*)	II	1 (Knafler et al. 2012)	Unknown	None (Knafler et al. 2012)	Maximal diversity associated with more hatched eggs (Knafler et al. 2012)	Yes (Knafler et al. 2012)

House sparrow (*Passer domesticus*)	I	5 (Loiseau et al. 2011)	Unknown	Assortative and maximal (Bonneaud et al. 2006)	Good alleles against avian malaria infection (Loiseau et al. 2011). Intermediate diversity associated with larger clutch sizes (Bonneaud et al. 2004)	Unknown
Seychelle warbler (*Acrocephalus sechellensis*)	I	4 (Richardson et al. 2005)	Unknown	Maximal diversity for extra-pair copulations (Richardson et al. 2005)	Unknown	Yes (Richardson & Westerdahl 2003)
Great reed warbler (*Acrocephalus arundinaceus*)	I	6 (Westerdahl et al. 2005)	Unknown	None (Westerdahl 2004)	Good allele and heterozygote advantage against avian malaria (Westerdahl et al. 2005)	Yes (Richardson & Westerdahl 2003)
Common yellowthroats (*Geothlypis trichas*)	I, II	8 (I) - (Dunn et al. 2013) 20 (II) - (Bollmer et al. 2010)	Male plumage (II) (bib brightness and mask size; diversity) (Dunn et al. 2013; Whittingham et al. 2015)	None (Bollmer et al. 2012)	Maximal diversity (I), and good allele (II) against avian malaria (Whittingham et al. 2015)	Yes (II) (Bollmer et al. 2010)
Great tits (*Parus major*)	I	5 (Sepil et al. 2012)	Unknown	None (Sepil et al. 2015)	Specific MHC supertype conferred protection against avian malaria (Sepil et al 2013)	Yes (Sepil et al. 2012)
Reptiles						
Sand lizard (*Lacerta agilis*)	I	Unknown	Chemical (genotype) (Olsson et al. 2003)	Disassortative (Olsson et al. 2003)	Good alleles against ectoparasites; (Olsson et al. 2005)	Unknown
Tuatara (*Sphenodon* spp.)	I	3 (Miller et al. 2010)	Unknown	Disassortative (Miller et al. 2009)	Unknown	Yes (Miller et al. 2010)
Amphibians						
African clawed frog (*Xenopus laevis*)	II	2 (Mable et al. 2015)	Unknown	Unknown	Good allele against a single strain bacterial infection (Barribeau et al. 2008)	Yes (Mable et al. 2015)

Tiger salamanders (*Ambystoma tigrinum*)	II	1 (Bos & DeWoody 2005)	Unknown	Assortative (Bos et al. 2009)	Unknown	Yes (Bos & DeWoody 2005)
Fishes						
Rose bitterling (*Rhodeus ocellatus*)	II	1 (Agbali et al. 2010)	Unknown	Disassortative; (Agbali et al. 2010)	Maximal diversity associated with increased embryo survival (Agbali et al. 2010)	Yes (Agbali et al. 2010)
Three-spine stickleback (*Gasterosteus aculeatus*)	II	5 (Milinski 2006)	Chemical (genotype) (Milinski et al. 2005)	Optimal MHC diversity (Milinski 2006)	Intermediate diversity associated with increased parasite resistance (Wegner et al. 2003)	Yes (Reusch & Langefors 2005)
Atlantic salmon (*Salmo salar*)	II	1 (Landry & Bernatchez 2001)	Unknown	Disassortative (Consuegra & Garcia de Leaniz 2008)	Good allele against bacterial infection (Lohm et al. 2002)	Yes (Langefors et al. 2001)
Chinook salmon (*Oncorhynchus tshawytscha*)	II	1 (Neff et al. 2008)	Unknown	Disassortative (Neff et al. 2008)	Maximal diversity against bacterial infection (Evans & Neff 2009)	Yes (Miller et al. 1997)
Arctic charr (*Salvelinus alpinus*)	II	1 (Kekäläinen et al. 2009)	Unknown	Unknown	Maximal diversity increases fertilization success (Skarstein et al. 2005) and increases parasite resistance (Kekäläinen et al. 2009)	Unknown

6.3 Future Directions

6.3.1 MHC susceptibility to avian malaria

My thesis has increased the knowledge of additive and non-additive effects of MHC on resistance to avian *Plasmodium*. I found no evidence for population differences in allele frequencies at MHC, suggesting that locally-good genes at the MHC regions surveyed do not explain previous findings that song sparrows are more resistant to local than nonlocal strains of *Plasmodium* (Sarquis-Adamson and MacDougall-Shackleton 2016). However, the high variability of these loci makes it particularly difficult to detect population structuring, particularly when sample sizes are relatively low as in my study, so future studies with larger sample sizes should increase power to detect locally-protective alleles at MHC if they exist. Geographic surveys of other candidate immune genes, such as toll-like receptors (Netea et al. 2012) or β-defensins (Hollox and Armour 2008), represent another logical next step to identifying the genetic basis (if any) for enhanced resistance to sympatric parasites.

6.3.2 Preen wax as a signal of MHC genotype

By investigating a novel chemical cue for MHC genotype, and demonstrating that the chemical composition of preen wax co-varies with amino acid distances of MHC class IIβ, I have raised the question of how, mechanistically, preen wax might signal MHC class IIβ genotypes. My working hypothesis is that the microbiome within and around the preen gland may help mediate the chemical composition of the preen wax of song sparrows. Odour-producing symbiotic bacteria were identified in the preen glands of a closely related species, the dark-eyed junco (Whittaker and Theis 2016). Two candidate bacterial genera discovered were *Burkholderia* and *Pseudomonas*, both known to produce the volatile compound 2-tridecanone. The proportion of this compound is associated with sex differences in preen wax chemistry for this species (Whittaker et al. 2010) and predicts reproductive success in males (Whittaker et al. 2013).

To further understand the role that MHC plays in the preen gland microbiome (and thus preen wax composition), one could inject antibiotics in and around the preen gland to

destroy the bacteria in this region. Next, taking bacterial samples throughout multiple time points of microbial recolonization in and around the preen gland would allow verifying the genera and abundance of bacterial strains. Assuming the microbes in and around the preen gland thrive in different conditions, then the next step would be to isolate and cultivate preen gland microbial colonies using various environmental conditions that the bacteria may thrive in (e.g., oxygenated, hypoxic, and anoxic). One could then identify the bacterial metabolites (e.g., wax esters; Ishige et al. 2003). Indeed, I would predict the bacterial metabolites to be associated with the four compounds I found covaried with MHC distance. Therefore, as my results from Chapter 3 suggest, if MHC plays a role in the chemical composition of preen wax (hypothesized to be mediated by the preen gland microbiome), then characterizing the bird's MHC, their bacteria, and the microbial metabolites should help reveal the mechanism in which songbird MHC indirectly influences the chemical composition of preen wax.

In addition to determining how MHC may influence chemical composition of preen wax, it also remains to be determined whether songbirds can detect MHC-related variation in preen wax chemical composition. Behavioural trials are a clear next step in assessing the salience of this variation to mating behaviour. To corroborate that preen wax acts a cue of MHC genotype, male and female songbirds might be given a choice experiment (e.g., Y-maze design; Whittaker et al. 2011) between the preen wax of potential mates of varying MHC similarity. Based on my results in Chapter 5, I would predict that if preen wax is a signal of MHC genotype, then song sparrows may prefer mates with similar, rather than dissimilar, preen oil chemical profiles. This is because I found that song sparrows mate assortatively based the evolutionary distance of MHC amino acid alleles (the same measure used to assess similarity of preen wax to MHC in Chapter 3).

6.3.3 Birdsong as a signal of MHC diversity

The lack of a relationship between pairwise birdsong syllable dissimilarity and MHC amino acid distance suggests that birdsong is not a reliable indicator of pairwise similarity at MHC. Even though song sparrow song varies geographically (Peters et al. 2000), the absence of a relationship between pairwise dissimilarity in repertoire content and MHC

dissimilarity between males could be due to the low genetic differentiation of MHC as indicated in Chapter 2.

My research has added to a growing body of evidence that optimal MHC diversity may in some cases be intermediate rather than maximal. Male song sparrows with the largest repertoires (a sexually selected trait associated with higher reproductive success; Reid et al. 2004) were not the most diverse, but rather had intermediate diversity, at MHC class IIβ. This finding is consistent with other evidence of the benefits of having intermediate levels of MHC alleles. For example, in three-spined sticklebacks, parasite load is lowest in individuals with intermediate MHC diversity (Milinski 2003), and males with intermediate MHC diversity had the best nest quality (Jäger et al. 2007). I suggest that being optimally (intermediately) diverse at MHC class IIβ helps male song sparrows ward off pathogens during song learning in early life (Marler and Peters 1987). Pathogen stress during this developmental window can reduce song complexity (Spencer et al. 2005; Brumm et al. 2009). This hypothesis could be tested by captive experiments, subjecting young male songbirds to a standard set of infectious agents concurrently with song learning: based on the results of Chapter 4, I predict that males with intermediate MHC IIβ diversity will show the best song learning. Additionally, field experiments aimed at testing whether males of intermediate MHC diversity provide more care to their offspring than males with low or high MHC diversity will also help illuminate what direct benefits females might procure by choosing males with the largest repertoires.

6.3.4 Selection at MHC class IIβ and MHC-mediated mate choice

Signals of positive selection on codons at the PBR, and specifically, at putative peptide-binding codons (Brown et al. 1993) for MHC class IIβ in song sparrows is consistent with many other taxa (e.g., steelhead *Onocorhynchus mykiss*, Aguilar and Garza 2006; great snipe *Gallinago media*, Ekblom et al. 2010; lowland leopard frog *Lithobates yavapaiensis*, Savage and Zamudio 2011). However, MHC-mediated mate choice is unlikely to be driving positive selection on the entire PBR since I found assortative pairing based on the evolutionary distance of MHC class IIβ rather than on the subset of codons showing positive selection. The most likely hypothesis is that positive selection at these codons is

pathogen-mediated. For example, an increase in parasitic nematode species richness measured in 116 primate species explained a species-specific increase in non-synonymous codon changes at the PBR for class IIβ (Garamszegi and Nunn 2011). To further untangle pathogen-mediated positive selection on codons at the MHC PBR, a similar study could be performed on a subset of the songbird subspecies that are widespread (and thus likely have geographically variable selection pressure) like song sparrows (Arcese et al. 2002) dark-eyed juncos (Nolan et al. 2002), and yellow-rumped warblers (*Setophaga coronata*, Hunt and Flaspohler 1998). These studies could measure the richness of each subspecies' haematozoan parasites, and compare that to the number of non-synonymous changes on codons at the MHC PBR for each subspecies. I would expect sub-species subjected to multiple haematozoa species (identified via molecular methods; Waldenström et al. 2004; Bell et al. 2015) would have more non-synomymous codon changes at the PBR of MHC class I exon 3 (since class I molecules detect intracellular pathogens, such as haematozoans) than synonymous changes. This would shine light on the variation in spatial selective pressure that parasites inflict on their host, and how parasites can shape the evolution of MHC within a widespread species.

My thesis has added to the scant evidence that assortative MHC-mediated mate choice can occur in the wild, rather than disassortative. I found assortative, rather than disassortative MHC-mediated mate choice between social parents in my population. Although my population has low incidence of extra-pair young (~20% of nests; Potvin and MacDougall-Shackleton 2009), by investigating the evolutionary MHC distances between the social mother and extra-pair father, I could reveal if MHC has an influence on choosing a mate purely for genetic benefits. For example, in the Seychelles warbler (*Acrocephalus sechellensis*), social mates did not pair at MHC class I exon 3, however, if females were paired with a male that had low MHC class I diversity, they were more likely to obtain extra-pair paternity in their nest, implying an indirect benefit if class I diversity was critical to offspring fitness (Richardson et al. 2005).

Given that MHC class IIβ is highly diverse in the study population of song sparrows (Chapter 2), and maximal MHC diversity is not associated with larger song repertoires

(Chapter 4), then choosing a mate with a relatively similar MHC genotype may enhance offspring fitness by reducing outbreeding depression. Overly diverse individuals at MHC class IIβ may have a fitness disadvantage due to autoimmune disorders (Klein 1986; Piertney and Oliver 2006; Trowsdale 2011) or a dilution of beneficial alleles (de Boer and Perelson 1993). Additionally, investigating whether socially mated pairs are more similar in preen wax chemical composition than expected by chance, may further corroborate that assortative mating at MHC is achieved through assessing chemical cues found in preen wax.

6.4 Conclusion

Overall, my thesis has identified ways in which MHC genotype affects disease resistance of song sparrows through non-additive genetic effects; established that the chemical composition of preen wax and the content of learned song repertoires provide information regarding MHC dissimilarity and diversity, respectively; and demonstrated that free-living song sparrows pair assortatively, rather than disassortatively, with respect to MHC genotype. Importantly, this work has identified novel cues of MHC genotype in songbirds, and will form the basis for further research on MHC signalling through chemical and acoustic cues. Furthermore, I found evidence of positive selection occurring at MHC class IIβ in song sparrows, combined with assortative rather than disassortative social pairing in the wild. Thus, this dissertation has helped expand our understanding of the ecology and evolution of MHC in songbirds through analyzing the role MHC plays in immunity, signalling, pathogen-mediated selection, and mate choice.

6.5 References

Abbott KM, Wickings EJ, Knapp LA. 2006. High levels of diversity characterize mandrill (*Mandrillus sphinx*) Mhc-DRB sequences. *Immunogenetics*. 58:628–640.

Agbali M, Reichard M, Bryjová A, Bryja J, Smith C. 2010. Mate choice for nonadditive genetic benefits correlate with MHC dissimilarity in the rose bitterling (*Rhodeus ocellatus*). *Evolution*. 64:1683–1696.

Aguilar A, Edwards SV, Smith TB, Wayne RK. 2006. Patterns of variation in MHC class II loci of the little greenbul (*Andropadus virens*) with comments on MHC evolution in birds. *J Hered.* 97:133–142

Aguilar A, Garza JC. 2006. A comparison of variability and population structure for major histocompatibility complex and microsatellite loci in California coastal steelhead (*Oncorhynchus mykiss* Walbaum). *Mol Ecol*. 15:923–937.

Alcaide M, Edwards SV, Negro JJ, Serrano D, Tella JL. 2008. Extensive polymorphism and geographical variation at a positively selected MHC class II B gene of the lesser kestrel (*Falco naumanni*). *Mol Ecol*. 17:2652–2665.

Arcese P, Sogge MK, Marr AB, Patten AP. 2002. Song Sparrow (*Melospiza melodia*), version 2.0. In The Birds of North America (Rodewald PG, editor). Cornell Lab of Ornithology, Ithaca, New York, USA.

Axtner J, Sommer S. 2007. Gene duplication, allelic diversity, selection processes and adaptive value of MHC class II *DRB* genes of the bank vole, *Clethrionomys glareolus*. *Immunogenetics*. 59:417–426.

Bang BG, Cobb. 1968. The size of the olfactory bulb in 108 species of birds. *The Auk*. 85:55–61.

Baratti M, Dessì-Fulgheri F, Ambrosini R, Bonisoli-Alquati A, Caprioli M, Goti E, Matteo A, Monnanni R, Ragionieri L, Ristori E, Romano M, Rubolini D, Scialpi A, Saino N.

2012. MHC genotype predicts mate choice in the ring-necked pheasant *Phasianus colchicus*. *J Evol Biol*. 25:1531–1542.

Barribeau SM, Villinger J, Waldman B. 2008. Major histocompatibility complex based resistance to a common bacterial pathogen of amphibians. *PLoS ONE*. 3:e2692.

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

Bell JA, Weckstein JD, Fecchio A, Tkach VV. 2015. A new real-time PCR protocol for detection of avian haemosporidians. *Parasit Vectors*. 8:383.

Bernatchez L, Landry C. 2003. MHC studies in nonmodel vertebrates: what have we learned about natural selection in 15 years? *J Evol Biol*. 16:363–377.

Boehm T, Zufall F. 2006. MHC peptides and the sensory evaluation of genotype. *Trends Neurosci*. 29:100–107.

de Boer RJ, Perelson AS. 1993. How diverse should the immune system be? *Proc R Soc Lond B*. 252:171–175.

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

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

Bollmer JL, Ruder EA, Johnson JA, Eimes JA, Dunn PO. 2011. Drift and selection influence geographic variation at immune loci of prairie-chickens. *Mol Ecol*. 20:4695–4706.

Bonneaud C, Chastel O, Federici P, Westerdahl H, Sorci G. 2006. Complex *Mhc*-based mate choice in a wild passerine. *Proc R Soc B*. 273:1111–1116.

Bonneaud C, Mazuc J, Chastel O, Westerdahl H, Sorci G, Poulin R. 2004. Terminal investment induced by immune challenge and fitness traits associated with major histocompatibility complex in the house sparrow. *Evolution*. 58:2823–2830.

Borghans JAM, Beltman JB, Boer RJD. 2004. MHC polymorphism under host-pathogen coevolution. *Immunogenetics*. 55:732–739.

Bowen L, Aldridge BM, DeLong R, Melin S, Godinez C, Zavala A, Gulland F, Lowenstine L, Stott JL, Johnson ML. 2006. MHC gene configuration variation in geographically disparate populations of California sea lions (*Zalophus californianus*). *Mol Ecol*. 15:529–533.

Bos DH, DeWoody JA. 2005. Molecular characterization of major histocompatibility complex class II alleles in wild tiger salamanders (*Ambystoma tigrinum*). *Immunogenetics*. 57:775–781.

Bos DH, Williams RN, Gopurenko D, Bulut Z, DeWoody JA. 2009. Condition-dependent mate choice and a reproductive disadvantage for MHC-divergent male tiger salamanders. *Mol Ecol*. 18:3307–3315.

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

Brumm H, Zollinger SA, Slater PJB. 2009. Developmental stress affects song learning but not song complexity and vocal amplitude in zebra finches. *Behav Ecol Sociobiol*. 63:1387–1395.

Cammen K, Hoffman JI, Knapp LA, Harwood J, Amos W. 2011. Geographic variation of the major histocompatibility complex in Eastern Atlantic grey seals (*Halichoerus grypus*). *Mol Ecol*. 20:740–752.

Coleman SW. 2009. Taxonomic and sensory biases in the mate-choice literature: there are far too few studies of chemical and multimodal communication. *Acta Ethologica*. 12:45–48.

Consuegra S, Garcia de Leaniz C. 2008. MHC-mediated mate choice increases parasite resistance in salmon. *Proc R Soc B.* 275:1397–1403.

Dunn PO, Bollmer JL, Freeman-Gallant CR, Whittingham LA. 2013. MHC variation is related to a sexually selected ornament, survival, and parasite resistance in common yellowthroats. *Evolution*. 67:679–687.

Edwards SV, Hedrick PW. 1998. Evolution and ecology of MHC molecules: from genomics to sexual selection. *Trends Ecol Evol*. 13:305–311.

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

Eizaguirre C, Lenz TL, Kalbe M, Milinski M. 2012. Rapid and adaptive evolution of MHC genes under parasite selection in experimental vertebrate populations. *Nat Commun.* 3:621.

Ekblom R, Saether S, Grahn M, Fiske P, Kalas JA, Höglund J. 2004. Major histocompatibility complex variation and mate choice in a lekking bird, the great snipe (*Gallinago media*). *Mol Ecol*. 13:3821–3828.

Ekblom R, Sæther SA, Fiske P, Kålås JA, Höglund J. 2010. Balancing selection, sexual selection and geographic structure in MHC genes of great snipe. *Genetica*. 138:453–461.

Evans ML, Neff BD. 2009. Major histocompatibility complex heterozygote advantage and widespread bacterial infections in populations of Chinook salmon (*Oncorhynchus tshawytscha*). *Mol Ecol*. 18:4716–4729.

Garamszegi LZ, Nunn CL. 2011. Parasite-mediated evolution of the functional part of the MHC in primates. *J Evol Biol*. 24:184–195.

Gillingham MAF, Richardson DS, Løvlie H, Moynihan A, Worley K, Pizzari T. 2009. Cryptic preference for MHC-dissimilar females in male red junglefowl, *Gallus gallus*. *Proc R Soc B*. 276:1083–1092.

Hale ML, Verduijn MH, Møller AP, Wolff K, Petrie M. 2009. Is the peacock's train an honest signal of genetic quality at the major histocompatibility complex? *J Evol Biol*. 22:1284–1294.

Hasselquist D, Bensch S, von Schantz T. 1996. Correlation between male song repertoire, extra-pair paternity and offspring survival in the great reed warbler. *Nature*. 381:229–232.

Hawley DM, Fleischer RC. 2012. Contrasting epidemic histories reveal pathogen-mediated balancing selection on class II MHC diversity in a wild songbird. *PloS One*. 7:e30222.

Hedrick PW, Kim TJ, Parker KM. 2001. Parasite resistance and genetic variation in the endangered Gila topminnow. *Anim Conserv*. 4:103–109.

Hendrickson SL, Mayer GC, Wallen EP, Quigley K. 2000. Genetic variability and geographic structure of three subspecies of tigers (*Panthera tigris*) based on MHC class I variation. *Anim Conserv.* 3:135–143.

Hess CM, Edwards SV. 2002. The evolution of the major histocompatibility complex in birds. *BioScience* 52:423–431.

Hollox EJ, Armour JA. 2008. Directional and balancing selection in human beta-defensins. *BMC Evol Biol*. 8:113.

Hughes AL, Hughes MK. 1995. Natural selection on the peptide-binding regions of major histocompatibility complex molecules. *Immunogenetics*. 42:233–243.

Hughes AL, Ota T, Nei M. 1990. Positive Darwinian selection promotes charge profile diversity in the antigen-binding cleft of class I major-histocompatibility-complex molecules. *Mol Biol Evol*. 7:515–524.

Hunt PD, Flaspohler DJ. 1998. Yellow-rumped Warbler (*Setophaga coronata*), version 2.0. In The Birds of North America (Rodewald PG, editor). Cornell Lab of Ornithology, Ithaca, New York, USA.

Ishige T, Tani A, Sakai Y, Kato N. 2003. Wax ester production by bacteria. *Curr Opin Microbiol*. 6:244–250.

Jacob S, McClintock MK, Zelano B, Ober C. 2002. Paternally inherited HLA alleles are associated with women's choice of male odor. Nat Genet. 30:175–179.

Jäger I, Eizaguirre C, Griffiths SW, Kalbe M, Krobbach CK, Reusch TBH, 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. *J Evol Biol*. 20:2005–2015.

Kekäläinen J, Vallunen JA, Primmer CR, Rättyä J, Taskinen J. 2009. Signals of major histocompatibility complex overdominance in a wild salmonid population. *Proc R Soc B*. rspb20090727.

Klein J. 1986. *Natural history of the major histocompatibility complex*. John Wiley & Sons: New York, NY, USA.

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

Knafler GJ, Clark JA, Boersma PD, Bouzat JL. 2012. MHC diversity and mate choice in the Magellanic penguin, *Spheniscus magellanicus*. *J Hered*. 103:759–768.

Krzych U, Zarling S, Pichugin A. 2014. Memory T cells maintain protracted protection against malaria. *Immunol Lett*. 161:189–195.

Kubinak JL, Nelson CA, Ruff JS, Potts WK. 2012. Tradeoffs limiting MHC heterozygosity. In *Ecoimmunology*, pp. 225-258. Oxford University Press, Inc.: New York, NY, USA.

Kumánovics A, Lindahl KF. 2004. Good copy, bad copy: choosing animal models for HLA-linked diseases. *Curr Opin Genet Dev*. 14:258–263.

Lampe HM, Saetre G-P. 1995. Female pied flycatchers prefer males with larger song repertoires. *Proc R Soc Lond B*. 262:163–167.

Landry C, Bernatchez L. 2001. Comparative analysis of population structure across environments and geographical scales at major histocompatibility complex and microsatellite loci in Atlantic salmon (*Salmo salar*). *Mol Ecol*. 10:2525–2539.

Landry C, Garant D, Duchesne P, Bernatchez L. 2001. "Good genes as heterozygosity": the major histocompatibility complex and mate choice in Atlantic salmon (*Salmo salar*). *Proc R Soc Lond B*. 268:1279–1285.

Langefors Å, Lohm J, Grahn M, Andersen Ø, Schantz T von. 2001. Association between major histocompatibility complex class IIB alleles and resistance to *Aeromonas salmonicida* in Atlantic salmon. *Proc R Soc Lond B*. 268:479–485.

Leclaire S, Merkling T, Raynaud C, Mulard H, Bessiere J-M, Lhuillier E, Hatch SA, Danchin E. 2012. Semiochemical compounds of preen secretion reflect genetic make-up in a seabird species. *Proc R Soc B*. 279:1185–1193.

Leclaire S, Strandh M, Mardon J, Westerdahl H, Bonadonna F. 2017. Odour-based discrimination of similarity at the major histocompatibility complex in birds. *Proc R Soc B*. 284:20162466.

Leclaire S, van Dongen WFD, Voccia S, Merkling T, Ducamp C, Hatch SA, Blanchard P, Danchin É, Wagner RH. 2014. Preen secretions encode information on MHC similarity in certain sex-dyads in a monogamous seabird. *Sci Rep*. 4:6920.

Leinders-Zufall T. 2004. MHC class I peptides as chemosensory signals in the vomeronasal organ. *Science*. 306:1033–1037.

Lerner H, Fleischer R. 2010. Prospects for the use of next-generation sequencing methods in ornithology. *The Auk.* 127:4–15.

Lohm J, Grahn M, Langefors Å, Andersen Ø, Storset A, von Schantz T. 2002. Experimental evidence for major histocompatibility complex–allele–specific resistance to a bacterial infection. *Proc R Soc Lond B*. 269:2029–2033.

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*). *Mol Ecol*. 18:1331–1340.

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

Mable BK, Kilbride E, Viney ME, Tinsley RC. 2015. Copy number variation and genetic diversity of MHC Class IIb alleles in an alien population of *Xenopus laevis*. *Immunogenetics*. 67:591–603.

Madsen T, Ujvari B. 2006. MHC class I variation associates with parasite resistance and longevity in tropical pythons. *J Evol Biol*. 19:1973–1978.

Marler P, Peters S. 1987. A sensitive period for song acquisition in the song sparrow, *Melospiza melodia*: a case of age-limited learning. *Ethology*. 76:89–100.

Mays HL, Hill GE. 2004. Choosing mates: good genes versus genes that are a good fit. *TREE*. 19:554–559.

McClelland EE, Penn DJ, Potts WK. 2003. Major histocompatibility complex heterozygote superiority during coinfection. *Infect Immun*. 71:2079–2086.

McGregor PK, Krebs JR, Perrins CM. 1981. Song repertoires and lifetime reproductive success in the great tit (*Parus major*). *Am Nat*. 118:149–159.

Milinski M. 2003. The function of mate choice in sticklebacks: optimizing *Mhc* genetics. *J Fish Biol*. 63:1–16.

Milinski M. 2006. The major histocompatibility complex, sexual selection, and mate choice. *Annu Rev Ecol Evol Syst*. 37:159–186.

Milinski M, Croy I, Hummel T, Boehm T. 2013. Major histocompatibility complex peptide ligands as olfactory cues in human body odour assessment. *Proc R Soc B*. 280:20122889.

Milinski M, Griffiths S, Wegner KM, Reusch TBH, Haas-Assenbaum A, Boehm T. 2005. Mate choice decisions of stickleback females predictably modified by MHC peptide ligands. *Proc Natl Acad Sci USA*. 102:4414–4418.

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

Miller HC, Moore JA, Nelson NJ, Daugherty CH. 2009. Influence of major histocompatibility complex genotype on mating success in a free-ranging reptile population. *Proc R Soc B*. rspb.2008.1840.

Miller KM, Withler RE, Beacham TD. 1997. Molecular evolution at *Mhc* genes in two populations of chinook salmon *Oncorhynchus tshawytscha*. *Mol Ecol*. 6:937–954.

Møller AP, Szép T. 2011. The role of parasites in ecology and evolution of migration and migratory connectivity. *J Ornithol*. 152:141–150.

Müller-Schwarze D. 2006. *Chemical ecology of vertebrates*. Cambridge University Press: New York, USA.

Neff BD, Garner SR, Heath JW, Heath DD. 2008. The MHC and non-random mating in a captive population of Chinook salmon. *Heredity*. 101:175–185.

Neff BD, Pitcher TE. 2005. Genetic quality and sexual selection: an integrated framework for good genes and compatible genes. *Mol Ecol*. 14:19–38.

Nevitt GA. 2000. Olfactory foraging by antarctic Procellariiform seabirds: life at high reynolds numbers. *Biol Bull*. 198:245-253.

Netea MG, Wijmenga C, O'Neill LAJ. 2012. Genetic variation in Toll-like receptors and disease susceptibility. *Nat Immunol*. 13:535–542.

Nolan Jr. V, Ketterson ED, Cristol DA, Rogers CM, Clotfelter ED, Titus RC, Schoech SJ, Snajdr E. 2002. Dark-eyed Junco (*Junco hyemalis*), version 2.0. In The Birds of North America (Rodewald PG, editor). Cornell Lab of Ornithology, Ithaca, New York, USA.

O'Connor EA, Strandh M, Hasselquist D, Nilsson J-Å, Westerdahl H. 2016. The evolution of highly variable immunity genes across a passerine bird radiation. *Mol Ecol*. 25:977–989.

Oliver MK, Telfer S, Piertney SB. 2009. Major histocompatibility complex (MHC) heterozygote superiority to natural multi-parasite infections in the water vole (*Arvicola terrestris*). *Proc R Soc B*. 276:1119–1128.

Olsson M, Madsen T, Nordby J, Wapstra E, Ujvari B, Wittsell H. 2003. Major histocompatibility complex and mate choice in sand lizards. *Proc R Soc Lond B (Suppl.)*. 270:S254–S256.

Olsson M, Madsen T, Wapstra E, Silverin B, Ujvari B, Wittzell H. 2005. MHC, health, color, and reproductive success in sand lizards. *Behav Ecol Sociobiol*. 58:289–294.

Pagenkopp KM, Klicka J, Durrant KL, Garvin JC, Fleischer RC. 2008. Geographic variation in malarial parasite lineages in the common yellowthroat (*Geothlypis trichas*). *Conserv Genet*. 9:1577–1588.

Penn DJ. 2002. The scent of genetic compatibility: sexual selection and the major histocompatibility complex. *Ethology*. 108:1–21.

Penn DJ, Damjanovich K, Potts WK. 2002. MHC heterozygosity confers a selective advantage against multiple-strain infections. *Proc Natl Acad Sci*. 99:11260–11264.

Penn D, Potts W. 1998. MHC–disassortative mating preferences reversed by cross–fostering. *Proc R Soc Lond B.* 265:1299–1306.

Peters S, Searcy WA, Beecher MD, Nowicki S. 2000. Geographic variation in the organization of song sparrow repertoires. *The Auk.* 117:936–942.

Piertney SB, Oliver MK. 2006. The evolutionary ecology of the major histocompatibility complex. *Heredity*. 96:7–21.

Pitcher TE, Neff BD. 2006. MHC class IIB alleles contribute to both additive and nonadditive genetic effects on survival in Chinook salmon. *Mol Ecol*. 15:2357–2365.

Potvin DA, MacDougall-Shackleton EA. 2009. Parental investment amplifies effects of genetic complementarity on growth rates in song sparrows, *Melospiza melodia*. *Anim Behav*. 78:943–948.

Prugnolle F, Manica A, Charpentier M, Guégan JF, Guernier V, Balloux F. 2005. Pathogen-driven selection and worldwide HLA class I diversity. *Curr Biol*. 15:1022–1027.

Radwan J, Tkacz A, Kloch A. 2008. MHC and preferences for male odour in the bank vole. *Ethology*. 114:827–833.

Reid JM, Arcese P, Cassidy ALEV, Hiebert SM, Smith JNM, Stoddard PK, Marr AB, Keller LF. 2004. Song repertoire size predicts initial mating success in male song sparrows, *Melospiza melodia*. *Anim Behav*. 68:1055–1063.

Reinhold K. 2004. Modeling a version of the good-genes hypothesis: female choice of locally adapted males. *Org Divers Evol.* 4:157–163.

Reusch TBH, HaÈberli MA, Aeschlimann PB, Milinski M. 2001. Female sticklebacks count alleles in a strategy of sexual selection explaining MHC polymorphism. *Nature.* 414:300–302.

Reusch TBH, Langefors Å. 2005. Inter- and intralocus recombination drive MHC class IIB gene diversification in a teleost, the three-spined stickleback *Gasterosteus aculeatus*. *J Mol Evol*. 61:531–541.

Richardson DS, Komdeur J, Burke T, von Schantz T. 2005. MHC-based patterns of social and extra-pair mate choice in the Seychelles warbler. *Proc R Soc B Biol Sci.* 272:759–767.

Richardson DS, Westerdahl H. 2003; MHC diversity in two *Acrocephalus* species: the outbred great reed warbler and the inbred Seychelles warbler. *Mol Ecol*. 12:3523–3529.

Rodríguez A, Alcaide M, Negro JJ, Pilard P. 2011. Using major histocompatibility complex markers to assign the geographic origin of migratory birds: examples from the threatened lesser kestrel. *Anim Conserv*. 14:306–313.

Ruff JS, Nelson AC, Kubinak JL, Potts WK. 2012. MHC signaling during social communication. In *Self and Nonself*, pp 290–313. Advances in Experimental Medicine and Biology Vol 738. Springer Science & Business Media: Berlin, Germany.

Salamon H, Klitz W, Easteal S, Gao X, Erlich HA, Fernandez-Viña M, Trachtenber EA, McWeeney SK, Nelson MP, Thomson G. 1999. Evolution of HLA class II molecules: allelic and amino acid site variability across populations. *Genetics*. 152:393–400.

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

Savage AE, Zamudio KR. 2011. MHC genotypes associate with resistance to a frog-killing fungus. *Proc Natl Acad Sci USA*. 108:16705–16710.

Schwensow N, Fietz J, Dausmann K, Sommer S. 2008. MHC-associated mating strategies and the importance of overall genetic diversity in an obligate pair-living primate. *Evol Ecol*. 22:617–636.

Schwensow N, Fietz J, Dausmann KH, Sommer S. 2007. Neutral versus adaptive genetic variation in parasite resistance: importance of major histocompatibility complex supertypes in a free-ranging primate. *Heredity*. 99:265.

Searcy WA. 1984. Song repertoire size and female preferences in song sparrows. *Behav Ecol Sociobiol*. 14:281–286.

Sepil I, Lachish S, Hinks AE, Sheldon BC. 2013. *Mhc* supertypes confer both qualitative and quantitative resistance to avian malaria infections in a wild bird population. *Proc R Soc B*. 280:20130134–20130134.

Sepil I, Moghadam HK, Huchard E, Sheldon BC. 2012. Characterization and 454 pyrosequencing of Major Histocompatibility Complex class I genes in the great tit reveal complexity in a passerine system. *BMC Evol Biol*. 12:68.

Sepil I, Radersma R, Santure AW, De Cauwer I, Slate J, Sheldon BC. 2015. No evidence for MHC class I-based disassortative mating in a wild population of great tits. *J Evol Biol*. 28:642–654.

Setchell JM, Charpentier MJE, Abbott KM, Wickings EJ, Knapp LA. 2010. Opposites attract: MHC-associated mate choice in a polygynous primate. *J Evol Biol*. 23:136–148.

Setchell JM, Charpentier MJE, Abbott KM, Wickings EJ, Knapp LA. 2009. Is brightest best? Testing the Hamilton-Zuk hypothesis in mandrills. *Int J Primatol*. 30:825–844.

Setchell JM, Vaglio S, Abbott KM, Moggi-Cecchi J, Boscaro F, Pieraccini G, Knapp LA. 2011. Odour signals major histocompatibility complex genotype in an Old World monkey. *Proc R Soc B*. 278:274–280.

Shiina T, Inoko H, Kulski JK. 2004. An update of the HLA genomic region, locus information and disease associations: 2004. *Tissue Antigens* 64:631–649.

Sin YW, Annavi G, Dugdale HL, Newman C, Burke T, MacDonald DW. 2014. Pathogen burden, co-infection and major histocompatibility complex variability in the European badger (*Meles meles*). *Mol Ecol.* 23:5072–5088.

Sin YW, Annavi G, Newman C, Buesching C, Burke T, MacDonald DW, Dugdale HL. 2015. MHC class II-assortative mate choice in European badgers (*Meles meles*). *Mol Ecol*. 24:3138–3150.

Sin YW, Dugdale HL, Newman C, MacDonald DW, Burke T. 2012. MHC class II genes in the European badger (*Meles meles*): characterization, patterns of variation, and transcription analysis. *Immunogenetics*. 64:313–327.

Singh V, Gupta P, Pande V. 2014. Revisiting the multigene families: *Plasmodium* var and vir genes. *J Vector Borne Dis.* 51:75–81.

Skarstein F, Folstad I, Liljedal S, Grahn M. 2005. MHC and fertilization success in the Arctic charr *Salvelinus alpinus*. *Behav Ecol Sociobiol*. 57:374–380.

Sommer S. 2003. Effects of habitat fragmentation and changes of dispersal behaviour after a recent population decline on the genetic variability of noncoding and coding DNA of a monogamous Malagasy rodent. *Mol Ecol*. 12:2845–2851.

Sommer S. 2005. The importance of immune gene variability (MHC) in evolutionary ecology and conservation. *Front Zool*. 2:16.

Sommer S, Schwab D, Ganzhorn JU. 2002. MHC diversity of endemic Malagasy rodents in relation to geographic range and social system. *Behav Ecol Sociobiol*. 51:214–221.

Sommer S, Volahy AT, Seal US. 2002. A population and habitat viability assessment for the highly endangered giant jumping rat *Hypogeomys antimena*, the largest extant endemic rodent of Madagascar. *Anim Conserv* Forum. 5:263–273.

Spencer KA, Buchanan KL, Leitner S, Goldsmith AR, Catchpole CK. 2005. Parasites affect song complexity and neural development in a songbird. *Proc R Soc B*. 272:2037–2043.

Stewart KA, MacDougall-Shackleton EA. 2008. Local song elements indicate local genotypes and predict physiological condition in song sparrows *Melospiza melodia*. *Biol Lett.* 4:240–242.

Strandh M, Lannefors M, Bonadonna F, Westerdahl H. 2011. Characterization of MHC class I and II genes in a subantarctic seabird, the blue petrel, *Halobaena caerulea* (Procellariiformes). *Immunogenetics*. 63:653–666.

Strandh M, Westerdahl H, Pontarp M, Canback 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. *Proc R Soc B*. 279:4457–4463.

Sutton JT, Castro I, Robertson BC, Tompkins DM, Stanton, J-A, Jamieson I. 2016. MHC genetic diversity and avian malaria prevalence in Mokoia Island saddlebacks. *New Zealand J of Ecol*. 40:0–0.

Thomas RH, Price ER, Seewagen CL, Mackenzie SA, Bernards MA, Guglielmo CG. 2010. Use of TLC-FID and GC-MS/FID to examine the effects of migratory state, diet and captivity on preen wax composition in White-throated Sparrows *Zonotrichia albicollis*. *Ibis.* 152:782–792.

Tregenza T, Wedell N. 2000. Genetic compatibility, mate choice and patterns of parentage: invited review. *Mol Ecol.* 9:1013–1027.

Trowsdale J. 2011. The MHC, disease and selection. *Immunol Lett*. 137:1–8.

von Schantz T, Wittzell H, Göransson G, Grahn M. 1997. Mate choice, male condition-dependent ornamentation and MHC in the pheasant. *Hereditas*. 127:133–140.

Waldenström J, Bensch S, Hasselquist D, Östman Ö. 2004. A new nested polymerase chain reaction method very efficient in detecting *Plasmodium* and *Haemoproteus* infections from avian blood. *J Parasitol*. 90:191–194.

Wedekind C, Seebeck T, Bettens F, Paepke AJ. 1995. MHC-dependent mate preferences in humans. *Proc R Soc Lond B*. 260:245–249.

Wegner KM, Kalbe M, Kurtz J, Reusch TBH, Milinski M. 2003. Parasite selection for immunogenetic optimality. *Science*. 301:1343.

Wegner KM, Kalbe M, Schaschl H, Reusch TBH. 2004. Parasites and individual major histocompatibility complex diversity—an optimal choice? *Microbes Infect*. 6:1110–1116.

Westerdahl H. 2004. No evidence of an MHC-based female mating preference in great reed warblers. *Mol Ecol*. 13:2465–2470.

Westerdahl H, Waldenstrom J, Hansson B, Hasselquist D, von Schantz T, Bensch S. 2005. Associations between malaria and MHC genes in a migratory songbird. *Proc R Soc B*. 272:1511–1518.

Westerdahl H, Wittzell H, von Schantz T. 1999. Polymorphism and transcription of *Mhc* class I genes in a passerine bird, the great reed warbler. *Immunogenetics*. 49:158–170.

Westerdahl H, Wittzell H, von Schantz T. 2000. *Mhc* diversity in two passerine birds: no evidence for a minimal essential *Mhc*. *Immunogenetics*. 52:92–100.

Whittaker DJ, Gerlach NM, Soini HA, Novotny MV, Ketterson ED. 2013. Bird odour predicts reproductive success. *Anim Behav*. 86:697–703.

Whittaker DJ, Richmond KM, Miller AK, Kiley R, Bergeon Burns C, Atwell JW, Ketterson ED. 2011. Intraspecific preen oil odor preferences in dark-eyed juncos (*Junco hyemalis*). *Behav Ecol*. 22:1256–1263.

Whittaker DJ, Soini HA, Atwell JW, Hollars C, Novotny MV, Ketterson ED. 2010. Songbird chemosignals: volatile compounds in preen gland secretions vary among individuals, sexes, and populations. *Behav Ecol*. 21:608–614.

Whittaker, D. J., and K. R. Theis. 2016. Bacterial communities associated with junco preen glands: preliminary ramifications for chemical signaling. In *Chemical Signals in Vertebrates* Vol 13, pp. 105-117. Springer: Cham, Switzerland.

Whittingham LA, Freeman-Gallant CR, Taff CC, Dunn PO. 2015. Different ornaments signal male health and MHC variation in two populations of a warbler. *Mol Ecol*. 24:1584–1595.

Worley K, Collet J, Spurgin LG, Cornwallis C, Pizzari T, Richardson DS. 2010. MHC heterozygosity and survival in red junglefowl. *Mol Ecol*. 19:3064–3075.

Worley K, Gillingham M, Jensen P, Kennedy LJ, Pizzari T, Kaufman J, Richardson DS. 2008. Single locus typing of MHC class I and class II B loci in a population of red jungle fowl. *Immunogenetics*. 60:233–247.

Xu S, Ren W, Zhou X, Zhou K, Yang G. 2010. Sequence polymorphism and geographical variation at a positively selected MHC-DRB gene in the finless porpoise (*Neophocaena phocaenoides*): implication for recent differentiation of the Yangtze finless porpoise? *J Mol Evol*. 71:6–22.

Yamazaki K, Boyse EA, Miké V, Thaler HT, Mathieson BJ, Abbott J, Boyse J, Zayas ZA, Thomas L. 1976. Control of mating preferences in mice by genes in the major histocompatibility complex. *J Exp Med*. 144:1324–1335.

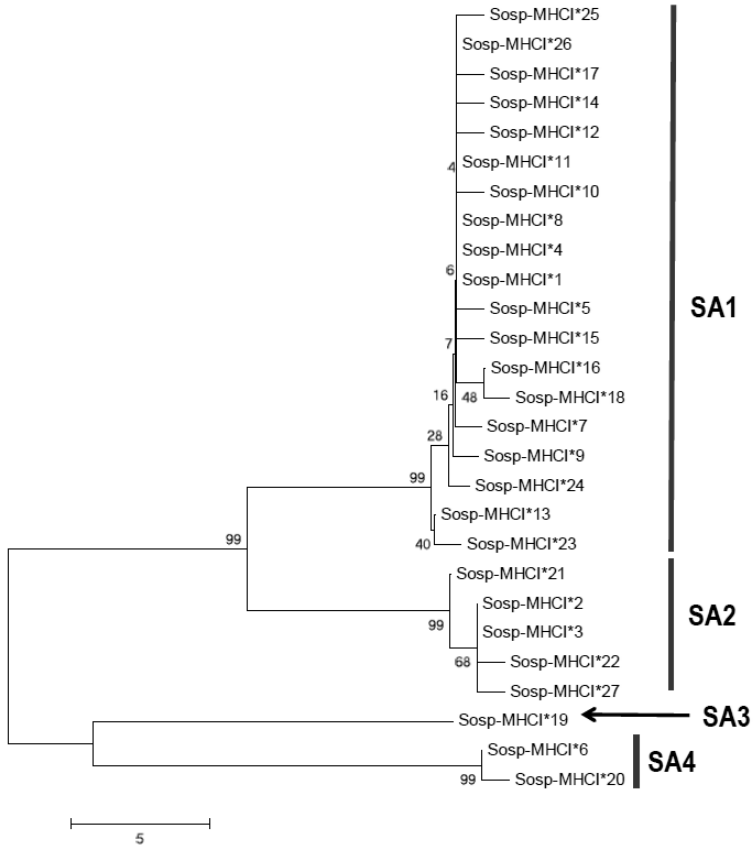
Yamazaki K, Yamaguchi M, Baranoski L, Bard J, Boyse EA, Thomas L. 1979. Recognition among mice. Evidence from the use of a Y-maze differentially scented by congenic mice of different major histocompatibility types. *J Exp Med*. 150:755–760.

Yamazaki K, Beauchamp GK, Kupniewski D, Bard J, Thomas L, Boyse EA. 1988. Familial imprinting determines H-2 selective mating preferences. Science. 240:1331–1333.

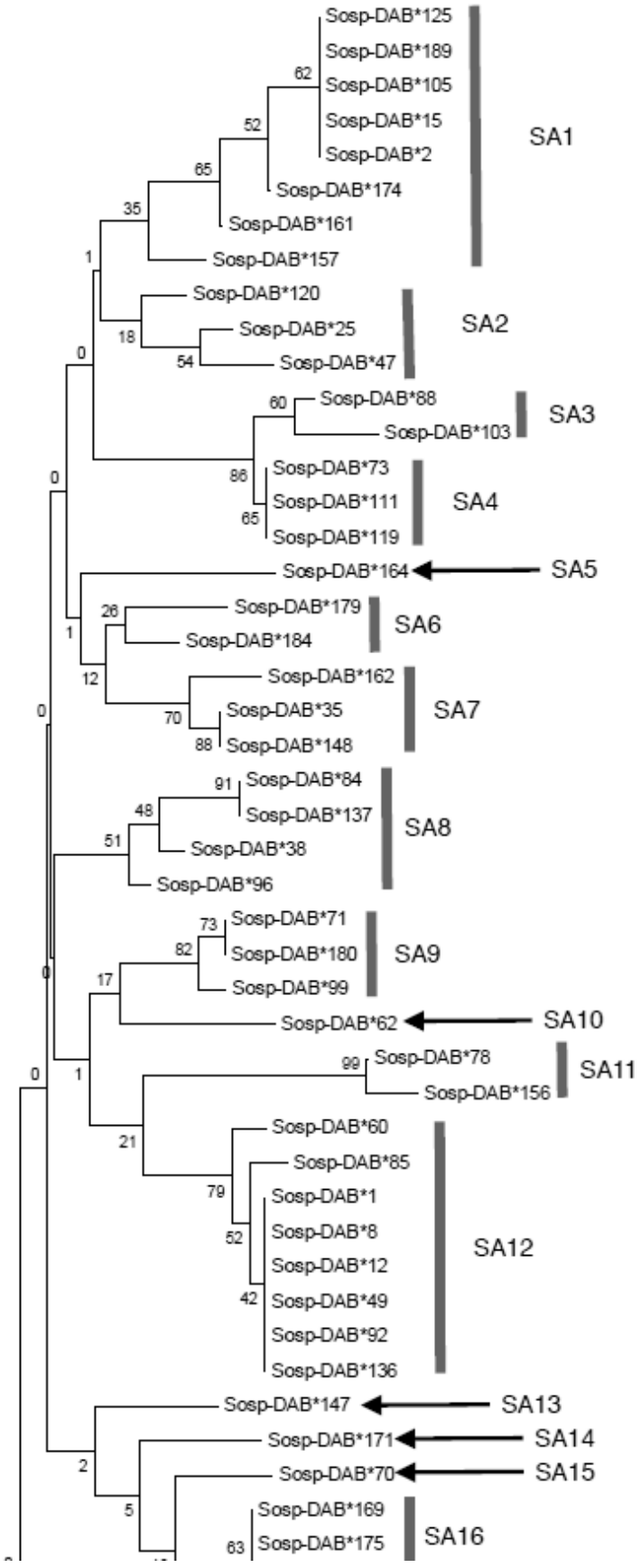
Zeng Q-Q, He K, Sun D-D, Ma M-Y, Ge Y-F, Fang S-G, Wan Q-H. 2016. Balancing selection and recombination as evolutionary forces caused population genetic variations in golden pheasant MHC class I genes. *BMC Evol Biol.* 16:42.

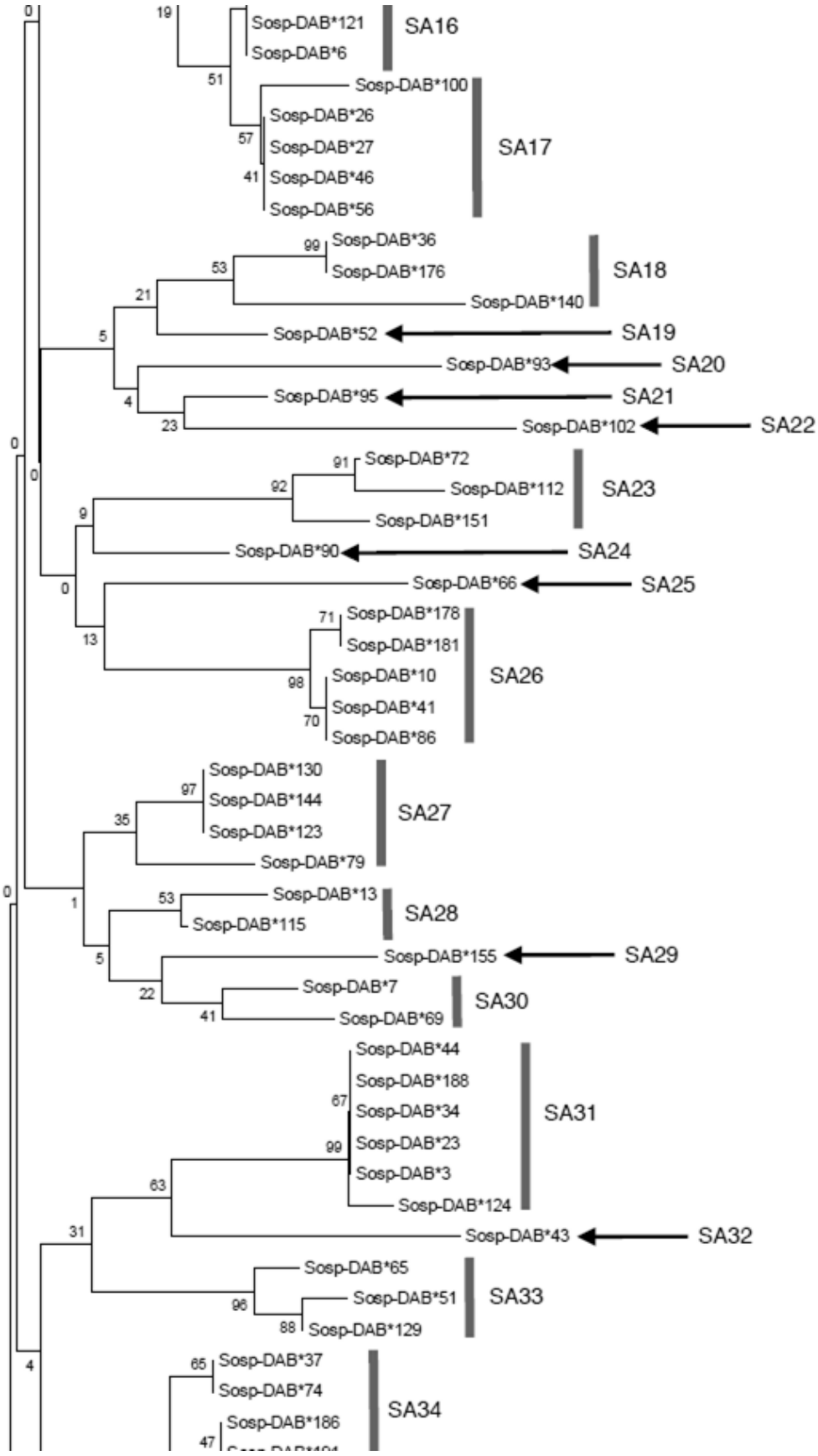
Appendices

Appendix A

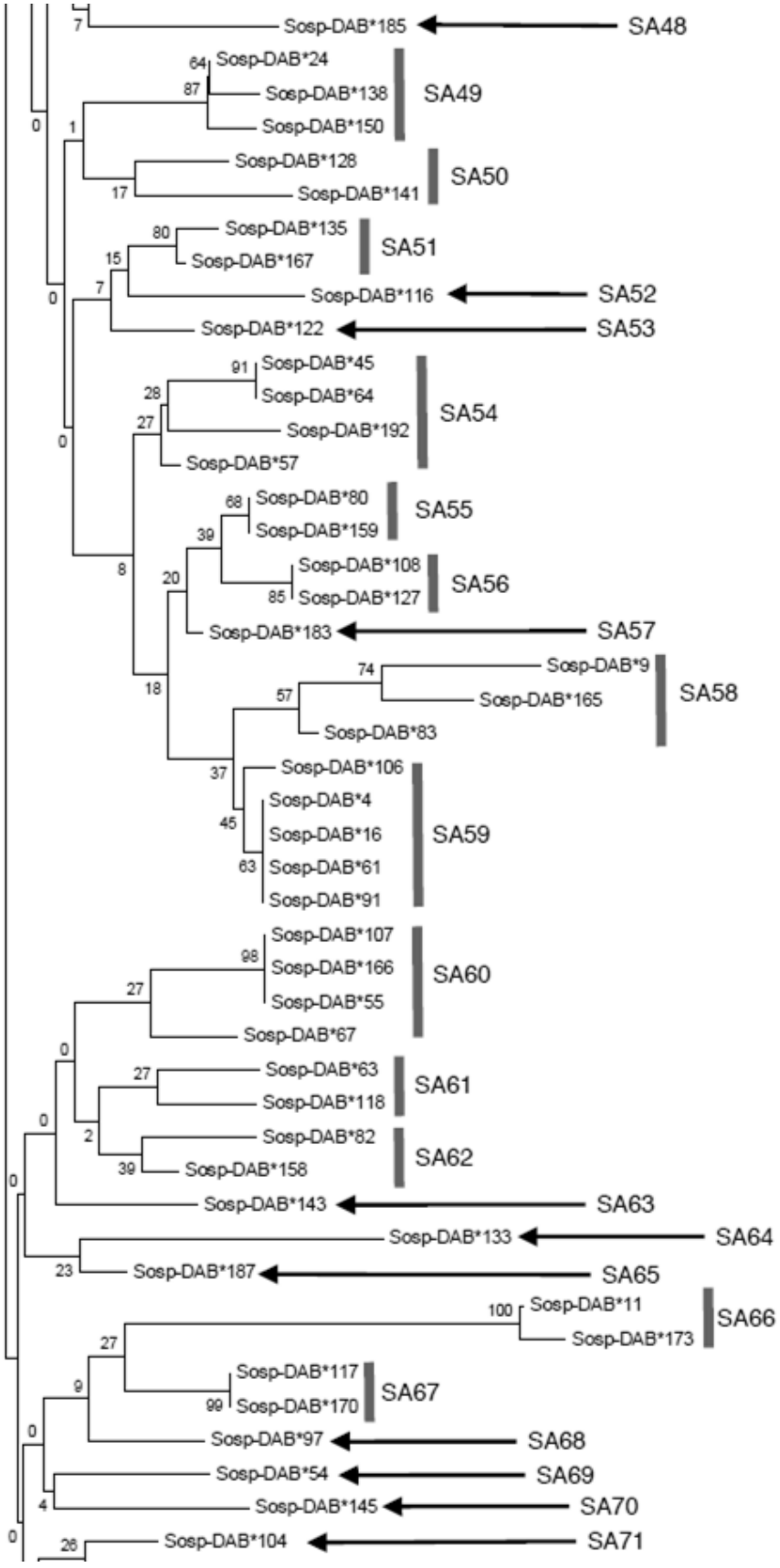


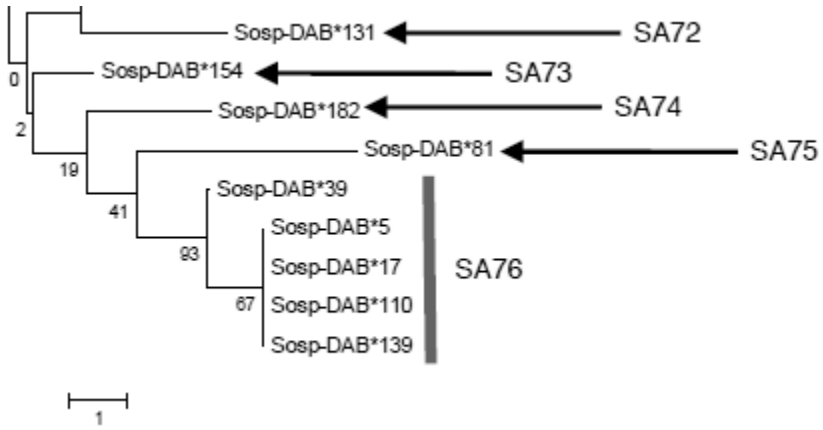
Appendix A.1 Evolutionary relationships of song sparrow MHC class I exon 3. The evolutionary history was inferred using the Neighbor-Joining method. The optimal tree with the sum of branch length = 70.22 is shown. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (500 replicates) are shown next to the branches. The tree is drawn to scale, and the evolutionary distances were computed using the number of differences. The analysis involved 27 nucleotide sequences (Sosp-MHCI*1-Sosp-MHC*27). Each thick vertical line indicates a superallele (SA) that contains more than one sequence, and the black arrow indicates a single sequence that belonged to its own SA.


Sosp-DAB*125
Sosp-DAB*189
Sosp-DAB*105
Sosp-DAB*15
Sosp-DAB*2
Sosp-DAB*174
Sosp-DAB*161
Sosp-DAB*157
SA1
Sosp-DAB*120
Sosp-DAB*25
Sosp-DAB*47
SA2
Sosp-DAB*88
Sosp-DAB*103
SA3
Sosp-DAB*73
Sosp-DAB*111
Sosp-DAB*119
SA4
Sosp-DAB*164
SA5
Sosp-DAB*179
Sosp-DAB*184
SA6
Sosp-DAB*162
Sosp-DAB*35
Sosp-DAB*148
SA7
Sosp-DAB*84
Sosp-DAB*137
Sosp-DAB*38
Sosp-DAB*96
SA8
Sosp-DAB*71
Sosp-DAB*180
Sosp-DAB*99
SA9
Sosp-DAB*62
SA10
Sosp-DAB*78
Sosp-DAB*156
SA11
Sosp-DAB*60
Sosp-DAB*85
Sosp-DAB*1
Sosp-DAB*8
Sosp-DAB*12
Sosp-DAB*49
Sosp-DAB*92
Sosp-DAB*136
SA12
Sosp-DAB*147
SA13
Sosp-DAB*171
SA14
Sosp-DAB*70
SA15
Sosp-DAB*169
Sosp-DAB*175
SA16


Sosp-DAB*121
Sosp-DAB*6
SA16
Sosp-DAB*100
Sosp-DAB*26
Sosp-DAB*27
Sosp-DAB*46
Sosp-DAB*56
SA17
Sosp-DAB*36
Sosp-DAB*176
Sosp-DAB*140
SA18
Sosp-DAB*52
SA19
Sosp-DAB*93
SA20
Sosp-DAB*95
SA21
Sosp-DAB*102
SA22
Sosp-DAB*72
Sosp-DAB*112
Sosp-DAB*151
SA23
Sosp-DAB*90
SA24
Sosp-DAB*66
SA25
Sosp-DAB*178
Sosp-DAB*181
Sosp-DAB*10
Sosp-DAB*41
Sosp-DAB*86
SA26
Sosp-DAB*130
Sosp-DAB*144
Sosp-DAB*123
Sosp-DAB*79
SA27
Sosp-DAB*13
Sosp-DAB*115
SA28
Sosp-DAB*155
SA29
Sosp-DAB*7
Sosp-DAB*69
SA30
Sosp-DAB*44
Sosp-DAB*188
Sosp-DAB*34
Sosp-DAB*23
Sosp-DAB*3
Sosp-DAB*124
SA31
Sosp-DAB*43
SA32
Sosp-DAB*65
Sosp-DAB*51
Sosp-DAB*129
SA33
Sosp-DAB*37
Sosp-DAB*74
Sosp-DAB*186
SA34


Sosp-DAB*191
Sosp-DAB*28
Sosp-DAB*75
Sosp-DAB*33
Sosp-DAB*14
Sosp-DAB*30
Sosp-DAB*87
Sosp-DAB*101
Sosp-DAB*113
SA34
Sosp-DAB*126
Sosp-DAB*149
Sosp-DAB*21
Sosp-DAB*19
Sosp-DAB*146
Sosp-DAB*18
Sosp-DAB*168
SA35
Sosp-DAB*177
Sosp-DAB*50
Sosp-DAB*134
Sosp-DAB*89
Sosp-DAB*29
Sosp-DAB*153
SA36
Sosp-DAB*114
Sosp-DAB*22
Sosp-DAB*58
SA37
Sosp-DAB*32
Sosp-DAB*77
Sosp-DAB*163
SA38
Sosp-DAB*40
Sosp-DAB*98
SA39
Sosp-DAB*53
Sosp-DAB*152
SA40
Sosp-DAB*190
SA41
Sosp-DAB*42
Sosp-DAB*31
Sosp-DAB*142
SA42
Sosp-DAB*59
Sosp-DAB*94
Sosp-DAB*20
SA43
Sosp-DAB*68
Sosp-DAB*160
SA44
Sosp-DAB*132
SA45
Sosp-DAB*48
Sosp-DAB*76
Sosp-DAB*172
SA46
Sosp-DAB*109
SA47


Sosp-DAB*185
SA48
Sosp-DAB*24
Sosp-DAB*138
Sosp-DAB*150
SA49
Sosp-DAB*128
Sosp-DAB*141
SA50
Sosp-DAB*135
Sosp-DAB*167
SA51
Sosp-DAB*116
SA52
Sosp-DAB*122
SA53
Sosp-DAB*45
Sosp-DAB*64
Sosp-DAB*192
Sosp-DAB*57
SA54
Sosp-DAB*80
Sosp-DAB*159
SA55
Sosp-DAB*108
Sosp-DAB*127
SA56
Sosp-DAB*183
SA57
Sosp-DAB*9
Sosp-DAB*165
Sosp-DAB*83
SA58
Sosp-DAB*106
Sosp-DAB*4
Sosp-DAB*16
Sosp-DAB*61
Sosp-DAB*91
SA59
Sosp-DAB*107
Sosp-DAB*166
Sosp-DAB*55
Sosp-DAB*67
SA60
Sosp-DAB*63
Sosp-DAB*118
SA61
Sosp-DAB*82
Sosp-DAB*158
SA62
Sosp-DAB*143
SA63
Sosp-DAB*133
SA64
Sosp-DAB*187
SA65
Sosp-DAB*11
Sosp-DAB*173
SA66
Sosp-DAB*117
Sosp-DAB*170
SA67
Sosp-DAB*97
SA68
Sosp-DAB*54
SA69
Sosp-DAB*145
SA70
Sosp-DAB*104
SA71



Appendix A.2 Evolutionary relationships of MHC class II exon 2 in song sparrows. The evolutionary history was inferred using the Neighbor-Joining method. The optimal tree with the sum of branch length = 374.46 is shown. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (500 replicates) are shown next to the branches. The tree is drawn to scale, and the evolutionary distances were computed using the number of differences method. The analysis involved 192 nucleotide sequences (Sosp-DAB*1 to Sosp-DAB*192). Each thick vertical line indicates a superallele (SA) that contains more than one sequence, and each black arrow indicates a single sequence that belonged to its own SA.

Appendix B

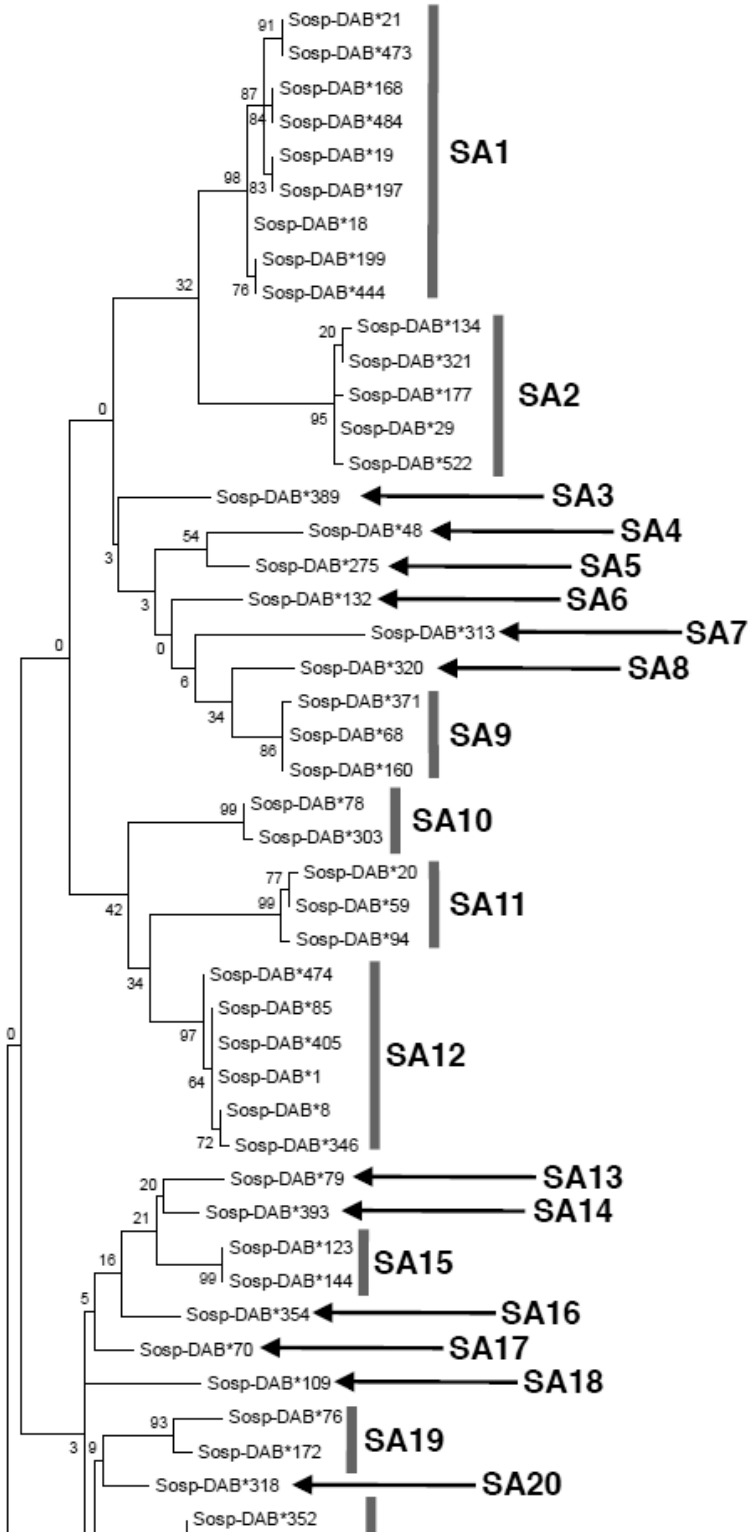
Compounds identified by GC-MS of preen wax from one song sparrow. Peaks clustered into groups of 2-4 with the same total number of carbons but varying chain lengths of the acid versus alcohol portions (e.g., 28a, 28b). Peak names correspond to those used in figure 3.1 and table 3.1. Some peaks consisted of a single monoester (e.g., 29a, 100% C13:C16) and others were monoester mixtures. Nomenclature for each monoester describes the number of carbons in the acid and alcohol portion, respectively, such that C13:C15 indicates a 13-carbon fatty acid esterified to a 15-carbon 1-alcohol. Asterisks denote peaks identified by BIOENV as being in the maximally informative subset.

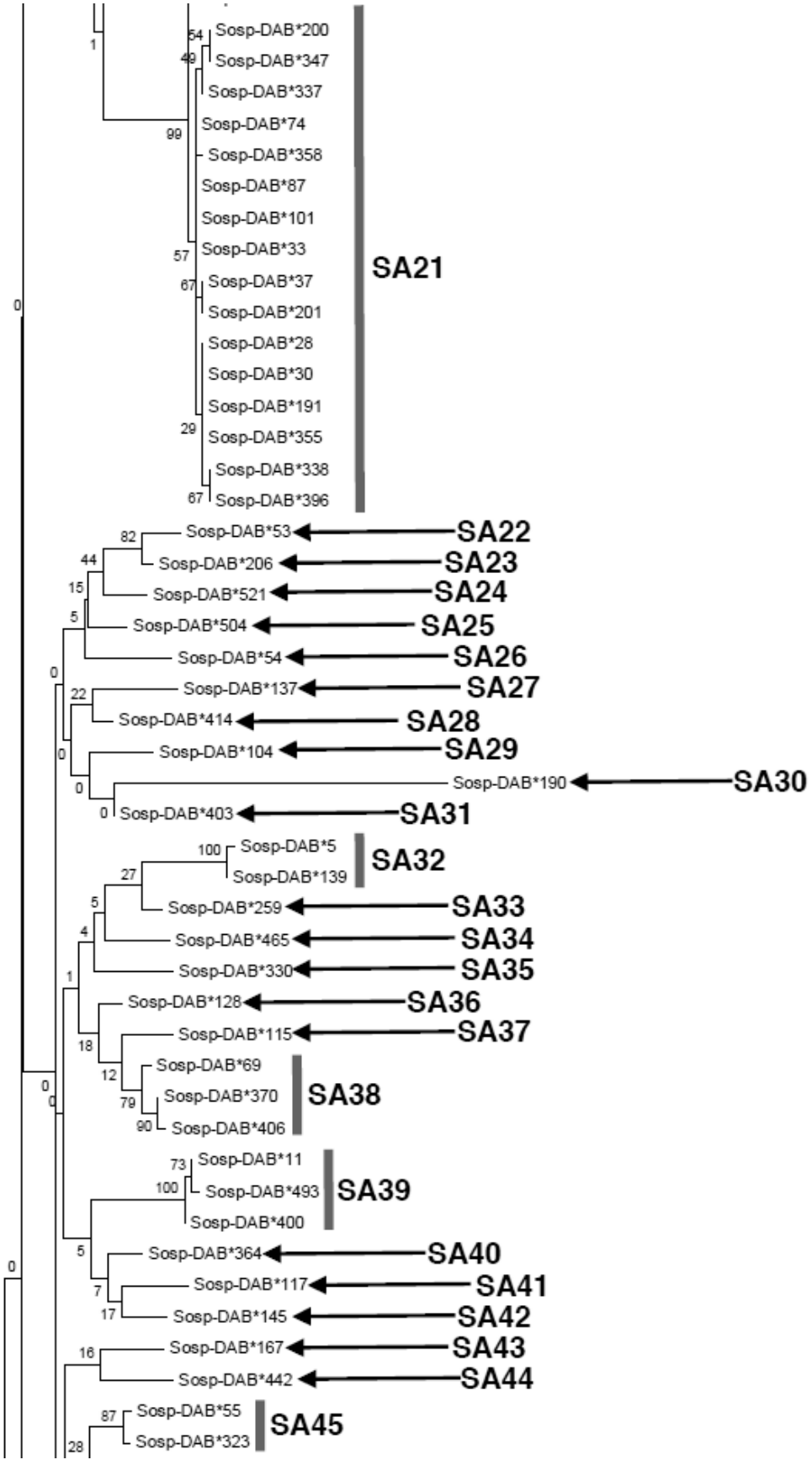
Peak	**# C's**	**Mol. Wt.**	**Acid:Alc % comp**	**Acid:Alc % comp**	**Acid:Alc % comp**	**Acid:Alc % comp**	**Acid:Alc % comp**	**Acid:Alc % comp**	**Acid:Alc % comp**
28a	28	424	C13:C15 32.6	C14:C14 32.4	C15:C13 28.5	C16:C12 6.5	C17:C11 0.0	C18:C10 0.0	C19:C9 0.0
28b*	28	424	C13:C15 13.4	C14:C14 26.5	C15:C13 45.2	C16:C12 12.3	C17:C11 2.6	C18:C10 0.0	C19:C9 0.0
29a	29	438	C13:C16 100.0	C14:C15 0.0	C15:C14 0.0	C16:C13 0.0	C17:C12 0.0	C18:C11 0.0	C19:C10 0.0
29b	29	438	C13:C16 83.6	C14:C15 11.9	C15:C14 1.5	C16:C13 2.8	C17:C12 0.0	C18:C11 0.0	C19:C10 0.2
30a	30	452	C13:C17 60.4	C14:C16 14.6	C15:C15 15.6	C16:C14 7.4	C17:C13 1.4	C18:C12 0.5	C19:C11 0.2
30c	30	452	C13:C17 83.2	C14:C16 13.1	C15:C15 2.8	C16:C14 0.4	C17:C13 0.2	C18:C12 0.2	C19:C11 0.2
31a	31	466	C13:C18 0.0	C14:C17 59.8	C15:C16 28.2	C16:C15 10.4	C17:C14 0.0	C18:C13 1.1	C19:C12 0.4
31b	31	466	C13:C18 0.0	C14:C17 0.0	C15:C16 0.0	C16:C15 78.7	C17:C14 0.0	C18:C13 12.2	C19:C12 9.1
31c	31	466	C13:C18	C14:C17	C15:C16	C16:C15	C17:C14	C18:C13	C19:C12

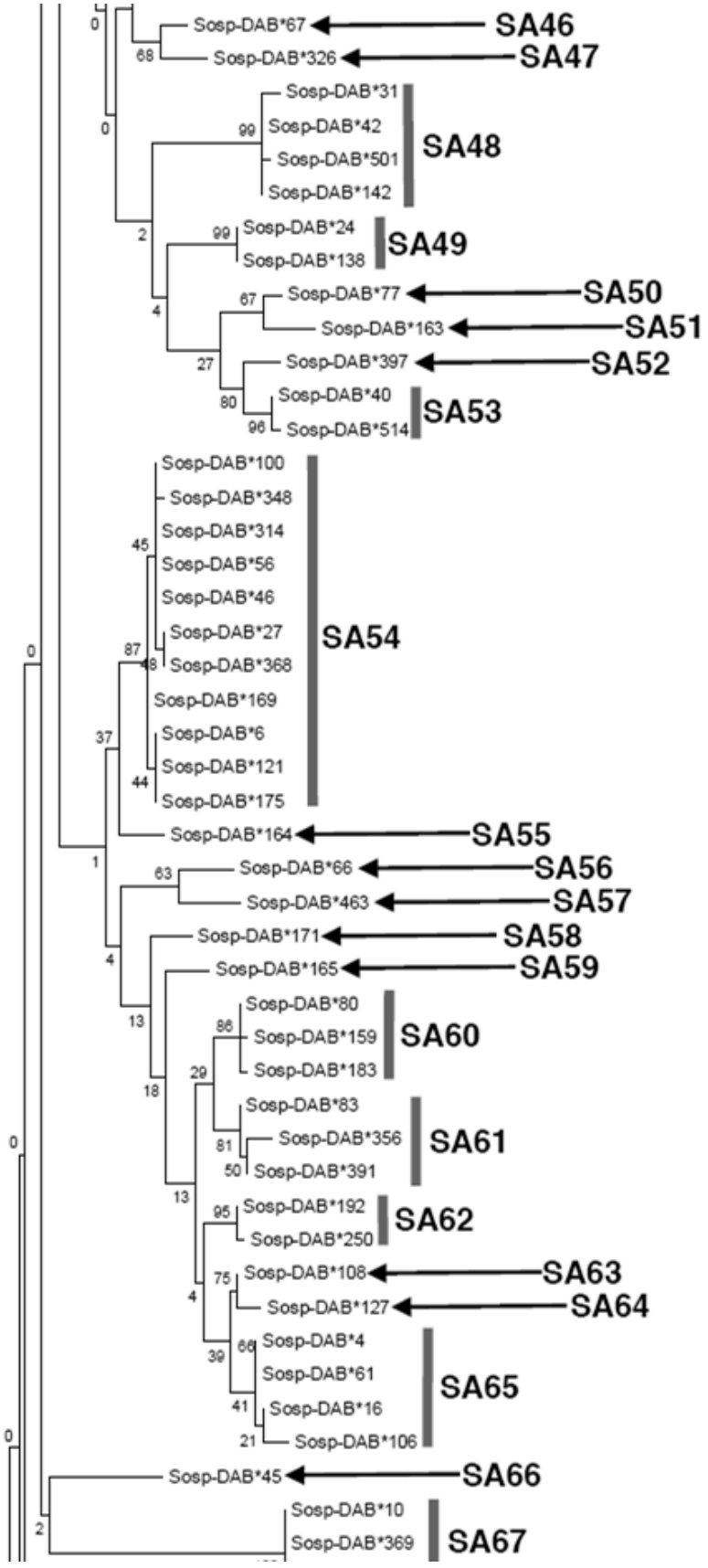
Peak	# C's	Mol. Wt.	Acid:Alc % comp	Acid:Alc % comp	Acid:Alc % comp	Acid:Alc % comp	Acid:Alc % comp	Acid:Alc % comp	Acid:Alc % comp
			33.7	50.5	13.6	1.2	0.5	0.3	0.3
32a	32	480	C13:C19	C14:C18	C15:C17	C16:C16	C17:C15	C18:C14	C19:C13
			35.4	29.9	28.3	3.7	1.8	0.8	0.0
32b	32	480	C13:C19	C14:C18	C15:C17	C16:C16	C17:C15	C18:C14	C19:C13
			0.0	0.0	0.0	85.0	0.0	15.0	0.0
32c*	32	480	C13:C19	C14:C18	C15:C17	C16:C16	C17:C15	C18:C14	C19:C13
			0.0	0.0	100.0	0.0	0.0	0.0	0.0
32d	32	480	C13:C19	C14:C18	C15:C17	C16:C16	C17:C15	C18:C14	C19:C13
			12.3	26.9	54.9	4.0	0.9	0.4	0.5
33a	33	494	C13:C20	C14:C19	C15:C18	C16:C17	C17:C16	C18:C15	C19:C14
			31.7	22.5	27.8	13.0	3.6	0.9	0.5
33b	33	494	C13:C20	C14:C19	C15:C18	C16:C17	C17:C16	C18:C15	C19:C14
			0.0	0.0	0.0	99.1	0.0	0.0	0.9
33c*	33	494	C13:C20	C14:C19	C15:C18	C16:C17	C17:C16	C18:C15	C19:C14
			0.0	0.0	0.0	0.0	0.0	0.0	100.0
33d	33	494	C13:C20	C14:C19	C15:C18	C16:C17	C17:C16	C18:C15	C19:C14
			2.7	17.5	53.6	20.8	4.1	0.7	0.5
34a	34	508	C13:C21	C14:C20	C15:C19	C16:C18	C17:C17	C18:C16	C19:C15
			14.9	25.1	27.4	18.3	11.4	2.0	0.9
34b	34	508	C13:C21	C14:C20	C15:C19	C16:C18	C17:C17	C18:C16	C19:C15
			0.0	0.0	0.0	69.1	0.0	29.7	1.2
34c	34	508	C13:C21	C14:C20	C15:C19	C16:C18	C17:C17	C18:C16	C19:C15
			0.0	0.0	0.0	0.0	98.7	0.0	1.3
34d*	34	508	C13:C21	C14:C20	C15:C19	C16:C18	C17:C17	C18:C16	C19:C15
			0.9	4.4	42.3	24.9	25.8	1.3	0.4
35a	35	522	C13:C22	C14:C21	C15:C20	C16:C19	C17:C18	C18:C17	C19:C16
			7.0	13.4	35.3	21.7	16.3	5.5	0.9
35b	35	522	C13:C22	C14:C21	C15:C20	C16:C19	C17:C18	C18:C17	C19:C16
			13.2	0.0	0.0	47.8	13.1	22.9	3.0
35c	35	522	C13:C22	C14:C21	C15:C20	C16:C19	C17:C18	C18:C17	C19:C16

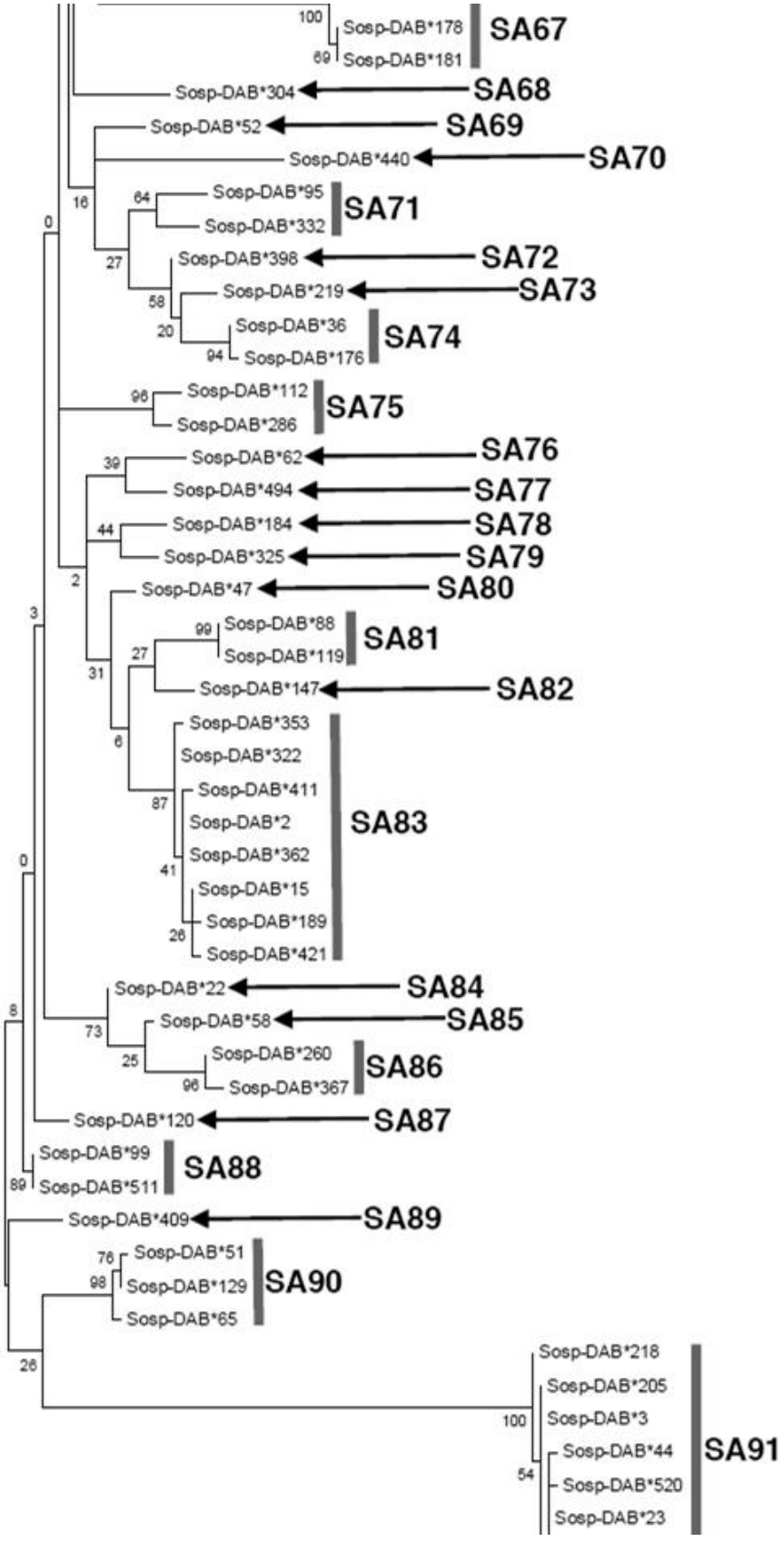
Peak	# C's	Mol. Wt.	Acid:Alc % comp	Acid:Alc % comp	Acid:Alc % comp	Acid:Alc % comp	Acid:Alc % comp	Acid:Alc % comp	Acid:Alc % comp
			10.3	0.0	0.0	0.0	89.7	0.0	0.0
35d	35	522	C13:C22	C14:C21	C15:C20	C16:C19	C17:C18	C18:C17	C19:C16
			0.1	1.6	13.3	30.1	43.7	10.4	0.9
36a	36	536	C13:C23	C14:C22	C15:C21	C16:C20	C17:C19	C18:C18	C19:C17
			2.7	8.2	26.3	32.8	21.3	6.4	2.2
36b	36	536	C13:C23	C14:C22	C15:C21	C16:C20	C17:C19	C18:C18	C19:C17
			0.0	0.0	0.0	0.0	0.0	100.0	0.0
36c	36	536	C13:C23	C14:C22	C15:C21	C16:C20	C17:C19	C18:C18	C19:C17
			0.0	0.0	0.0	0.0	0.0	0.0	100.0
36d	36	536	C13:C23	C14:C22	C15:C21	C16:C20	C17:C19	C18:C18	C19:C17
			0.0	0.3	3.8	13.7	56.8	19.0	6.5
37a	37	550	C13:C24	C14:C23	C15:C22	C16:C21	C17:C20	C18:C19	C19:C18
			0.9	2.4	18.5	26.0	40.8	8.5	2.8
37b	37	550	C13:C24	C14:C23	C15:C22	C16:C21	C17:C20	C18:C19	C19:C18
			0.0	0.0	0.0	0.0	0.0	89.9	10.1
37c	37	550	C13:C24	C14:C23	C15:C22	C16:C21	C17:C20	C18:C19	C19:C18
			100.0	0.0	0.0	0.0	0.0	0.0	0.0
37d	37	550	C13:C24	C14:C23	C15:C22	C16:C21	C17:C20	C18:C19	C19:C18
			0.0	0.0	0.0	7.9	37.3	38.6	16.1
38a	38	564	C13:C25	C14:C24	C15:C23	C16:C22	C17:C21	C18:C20	C19:C19
			0.3	0.4	4.1	24.7	51.1	15.6	3.8
38b	38	564	C13:C25	C14:C24	C15:C23	C16:C22	C17:C21	C18:C20	C19:C19
			1.4	0.0	0.0	0.0	0.0	98.6	0.0
38c	38	564	C13:C25	C14:C24	C15:C23	C16:C22	C17:C21	C18:C20	C19:C19
			0.0	0.0	1.0	0.0	0.0	42.2	56.7
39a	39	578	C13:C26	C14:C25	C15:C24	C16:C23	C17:C22	C18:C21	C19:C20

Appendix C


Sosp-DAB*21
Sosp-DAB*473
Sosp-DAB*168
Sosp-DAB*484
Sosp-DAB*19
Sosp-DAB*197
Sosp-DAB*18
Sosp-DAB*199
Sosp-DAB*444
SA1
Sosp-DAB*134
Sosp-DAB*321
Sosp-DAB*177
Sosp-DAB*29
Sosp-DAB*522
SA2
Sosp-DAB*389
SA3
Sosp-DAB*48
SA4
Sosp-DAB*275
SA5
Sosp-DAB*132
SA6
Sosp-DAB*313
SA7
Sosp-DAB*320
SA8
Sosp-DAB*371
Sosp-DAB*68
Sosp-DAB*160
SA9
Sosp-DAB*78
Sosp-DAB*303
SA10
Sosp-DAB*20
Sosp-DAB*59
Sosp-DAB*94
SA11
Sosp-DAB*474
Sosp-DAB*85
Sosp-DAB*405
Sosp-DAB*1
Sosp-DAB*8
Sosp-DAB*346
SA12
Sosp-DAB*79
SA13
Sosp-DAB*393
SA14
Sosp-DAB*123
Sosp-DAB*144
SA15
Sosp-DAB*354
SA16
Sosp-DAB*70
SA17
Sosp-DAB*109
SA18
Sosp-DAB*76
Sosp-DAB*172
SA19
Sosp-DAB*318
SA20
Sosp-DAB*352


Sosp-DAB*200
Sosp-DAB*347
Sosp-DAB*337
Sosp-DAB*74
Sosp-DAB*358
Sosp-DAB*87
Sosp-DAB*101
Sosp-DAB*33
Sosp-DAB*37
Sosp-DAB*201
Sosp-DAB*28
Sosp-DAB*30
Sosp-DAB*191
Sosp-DAB*355
Sosp-DAB*338
Sosp-DAB*396
SA21
Sosp-DAB*53
SA22
Sosp-DAB*206
SA23
Sosp-DAB*521
SA24
Sosp-DAB*504
SA25
Sosp-DAB*54
SA26
Sosp-DAB*137
SA27
Sosp-DAB*414
SA28
Sosp-DAB*104
SA29
Sosp-DAB*190
SA30
Sosp-DAB*403
SA31
Sosp-DAB*5
Sosp-DAB*139
SA32
Sosp-DAB*259
SA33
Sosp-DAB*465
SA34
Sosp-DAB*330
SA35
Sosp-DAB*128
SA36
Sosp-DAB*115
SA37
Sosp-DAB*69
Sosp-DAB*370
Sosp-DAB*406
SA38
Sosp-DAB*11
Sosp-DAB*493
Sosp-DAB*400
SA39
Sosp-DAB*364
SA40
Sosp-DAB*117
SA41
Sosp-DAB*145
SA42
Sosp-DAB*167
SA43
Sosp-DAB*442
SA44
Sosp-DAB*55
Sosp-DAB*323
SA45


Sosp-DAB*67
SA46
Sosp-DAB*326
SA47
Sosp-DAB*31
Sosp-DAB*42
Sosp-DAB*501
Sosp-DAB*142
SA48
Sosp-DAB*24
Sosp-DAB*138
SA49
Sosp-DAB*77
SA50
Sosp-DAB*163
SA51
Sosp-DAB*397
SA52
Sosp-DAB*40
Sosp-DAB*514
SA53
Sosp-DAB*100
Sosp-DAB*348
Sosp-DAB*314
Sosp-DAB*56
Sosp-DAB*46
Sosp-DAB*27
Sosp-DAB*368
Sosp-DAB*169
Sosp-DAB*6
Sosp-DAB*121
Sosp-DAB*175
SA54
Sosp-DAB*164
SA55
Sosp-DAB*66
SA56
Sosp-DAB*463
SA57
Sosp-DAB*171
SA58
Sosp-DAB*165
SA59
Sosp-DAB*80
Sosp-DAB*159
Sosp-DAB*183
SA60
Sosp-DAB*83
Sosp-DAB*356
Sosp-DAB*391
SA61
Sosp-DAB*192
Sosp-DAB*250
SA62
Sosp-DAB*108
SA63
Sosp-DAB*127
SA64
Sosp-DAB*4
Sosp-DAB*61
Sosp-DAB*16
Sosp-DAB*106
SA65
Sosp-DAB*45
SA66
Sosp-DAB*10
Sosp-DAB*369
SA67


Sosp-DAB*178
Sosp-DAB*181
SA67
Sosp-DAB*304
SA68
Sosp-DAB*52
SA69
Sosp-DAB*440
SA70
Sosp-DAB*95
Sosp-DAB*332
SA71
Sosp-DAB*398
SA72
Sosp-DAB*219
SA73
Sosp-DAB*36
Sosp-DAB*176
SA74
Sosp-DAB*112
Sosp-DAB*286
SA75
Sosp-DAB*62
SA76
Sosp-DAB*494
SA77
Sosp-DAB*184
SA78
Sosp-DAB*325
SA79
Sosp-DAB*47
SA80
Sosp-DAB*88
Sosp-DAB*119
SA81
Sosp-DAB*147
SA82
Sosp-DAB*353
Sosp-DAB*322
Sosp-DAB*411
Sosp-DAB*2
Sosp-DAB*362
Sosp-DAB*15
Sosp-DAB*189
Sosp-DAB*421
SA83
Sosp-DAB*22
SA84
Sosp-DAB*58
SA85
Sosp-DAB*260
Sosp-DAB*367
SA86
Sosp-DAB*120
SA87
Sosp-DAB*99
Sosp-DAB*511
SA88
Sosp-DAB*409
SA89
Sosp-DAB*51
Sosp-DAB*129
Sosp-DAB*65
SA90
Sosp-DAB*218
Sosp-DAB*205
Sosp-DAB*3
Sosp-DAB*44
Sosp-DAB*520
Sosp-DAB*23
SA91

Appendix C.1 MHC class II exon 2 maximum likelihood amino acid tree for 32 male song sparrows. Evolutionary history was inferred using the maximum likelihood method based on the WAG amino acid substitution model. The tree with the highest log likelihood is shown. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (500 replicates) are shown next to the branches. The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. The analysis involved 186 amino acid sequences (Sosp-DAB*#). Each thick vertical line indicates a superallele (SA) that contains more than one sequence, and each black arrow indicates a single sequence that belonged to its own SA.

Appendix C.2 Ranked set of candidate models predicting song sparrow song repertoire size, after excluding a potentially influential data point with low repertoire size and high MHC diversity (upper left corner of figure 4.1). Predictors were number of MHC superalleles and squared number of MHC superalleles.

Model	**df**	**logLik**	**AICc**	**ΔAICc**	**Model Weight**
Null: intercept only	2	-55.7	115.8	0	0.50
Quadratic: #MHC + $(\#MHC)^2$	4	-53.5	116.5	0.69	0.35
Linear: #MHC	3	-55.6	118.1	2.38	0.15

Appendix C.3 Model-averaged parameter estimates for predictors of song sparrow song repertoire size, after excluding a potentially influential data point with low repertoire size and high MHC diversity. Models were weighted as indicated in Appendix C.2 and fully averaged. Repertoire size increased with number of MHC superalleles and decreased with squared number of MHC superalleles.

Parameter	**Estimate ± SE**	**95% CI**
Intercept	2.94 ± 7.03	-11.02, 16.90
#MHC	1.92 ± 0.93	0.02, 3.82
$(\#MHC)^2$	-0.074 ± 0.036	-0.15, -0.0001

Appendix D

All studies followed the ethical guidelines from the Canadian Council on Animals Care, which was reviewed and approved by the Animal Use Subcommittee (AUS) at the University of Western Ontario. Below is the approval email from the AUS. Animal Use Protocol #: 2016-017.

AUP Number: 2016-017
PI Name: Macdougallshackleton, Elizabeth
AUP Title: Mating Signals, Gene Flow, And Disease Resistance In Songbirds

Approval Date: 06/09/2016

Official Notice of Animal Use Subcommittee (AUS) Approval: Your new Animal Use Protocol (AUP) entitled "Mating Signals, Gene Flow, And Disease Resistance In Songbirds

" has been APPROVED by the Animal Use Subcommittee of the University Council on Animal Care. This approval, although valid for four years, and is subject to annual Protocol Renewal.2016-017::1

1. This AUP number must be indicated when ordering animals for this project.
2. Animals for other projects may not be ordered under this AUP number.
3. Purchases of animals other than through this system must be cleared through the ACVS office. Health certificates will be required.

The holder of this Animal Use Protocol is responsible to ensure that all associated safety components (biosafety, radiation safety, general laboratory safety) comply with institutional safety standards and have received all necessary approvals. Please consult directly with your institutional safety officers.

Submitted by: Copeman, Laura
on behalf of the Animal Use Subcommittee
University Council on Animal Care

The University *of* Western Ontario
Animal Use Subcommittee / University Council on Animal Care
Health Sciences Centre, ● London, Ontario ● CANADA – N6A 5C1
PH: 519-661-2111 ext. 86768 ● FL 519-661-2028
Email: auspc@uwo.ca ● http://www.uwo.ca/animal/website/

Curriculum Vitae

Name:	Joel William Gordon Slade
Post-secondary Education and Degrees:	Dalhousie University Halifax, Nova Scotia, Canada 2005-2010, B.Sc.
	Saint Mary's University Halifax, Nova Scotia, Canada 2010-2012, M.Sc.
	University of Western Ontario London, Ontario, Canada 2013-2018, Ph.D.
Honours and Awards:	Society of Canadian Zoologists Cas C. Lyndsey Prize 2016
	QEII Graduate Scholarship in Science and Technology 2014, 2015, 2016
	American Ornithologists' Union Hesse Research Award 2015
	Society of Canadian Ornithologists Taverner Award 2015
	Natural Sciences and Engineering Research Council of Canada Canada Graduate Scholarship (NSERC CGS-M) 2011-2012
	Saint Mary's University Graduate Award 2010, 2011
	Dalhousie University Biology Teaching Assistant Award 2009
	Dalhousie University David Andrew Dougall Memorial Bursary 2008

Related Work Experience:

Research Associate
BEACON Center, Michigan State University
2017-present

Supported Course Redesign Graduate Assistant
University of Western Ontario
2017

Teaching Assistant
University of Western Ontario
2013, 2014, 2016

Sessional Instructor
University of Western Ontario
2015

Learning Development Graduate Fellowship
University of Western Ontario
2014-2015

Teaching Assistant
Saint Mary's University
2010-2012

Research Assistant
Dalhousie University
Fisheries Oceanography Lab: 2008, 2009
Plant Programmed Cell Death Lab: 2008

Publications:

Slade JWG, Watson MJ, MacDougall-Shackleton EA. 2017. Birdsong signals individual diversity at the major histocompatibility complex. *Bio Lett.* 13:20170430.

Slade JWG, Sarquis-Adamson Y, Gloor GB, Lachance M-A, MacDougall-Shackleton EA. 2017. Population differences at MHC do not explain enhanced resistance of song sparrows to local parasites. *J Hered.* 108:127-134.

Slade JWG, Watson MJ, Kelly TR, Gloor GB, Bernards MA, MacDougall-Shackleton EA. 2016. Chemical composition of preen wax reflects major histocompatibility complex similarity in songbirds. *Proc R Soc B.* 283:20161966.