Pathogens and the maintenance of genetic variation in an island population of the Seychelles warbler (*Acrocephalus sechellensis*)

Charlotte **(Charli) Sarah Davies**

A thesis submitted for the degree of Doctor of philosophy

School of Biological Sciences

University of East Anglia

March 2021

© This copy of the thesis has been supplied on condition that anyone who consults understood to recognise that its copyright rests with the author and that use of any information derived therefrom must be in accordance with current UK Copyright Law. In addition, any quotation or extract must include full attribution.

Abstract

The objective of this thesis was to investigate how evolutionary forces shape immunogenetic variation in small populations. To achieve this, I investigated functional variation at key genes in both the innate and adaptive immune systems in the Seychelles warbler (*Acrocephalus sechellensis)*, and how this differed spatio-temporally and in relation to individual traits and fitness components. First, I tracked evolution at one non-synonymous *TLR3* SNP over 25-years. Results showed a significant and consistent temporal decline in the minor *TLR3*^{C} allele frequency within the original Seychelles warbler population and in all four derived populations. Further investigations showed that positive selection – acting on both survival and reproduction, was driving these temporal changes in the Cousin population. I then investigated whether pre- or post-copulatory sexual selection was acting in relation to the *TLR*3 locus. I found evidence of both pre-copulatory assortative social pairing, and a post-copulatory bias against paternal inheritance of the *TLR3*^{C} haplotype. Thus, multiple mechanisms of selection appear be causing contemporary *TLR3* evolution in the Seychelles warbler. Lastly, I investigated whether host-microbiome coevolution may interact with functional immunogenetic variation in this species. Using next generation sequencing techniques, I characterised the MHC (both class I and class II), and gut microbiome (GM) variation existing in a subset of individuals. I found that presence of specific MHC alleles, but not MHC diversity, was associated with differences in GM diversity and composition. These results confirm variation in the host's immune system may play a role in shaping an individual's GM in a population of wild animals. Collectively, my results provide insight into how different mechanisms can interact to shape functional genetic variation in a natural population. They also raise some important, as yet unanswered, questions about the identity of the selective agents (pathogens) affecting the immunogenetic variation in this species.

Access Condition and Agreement

Each deposit in UEA Digital Repository is protected by copyright and other intellectual property rights, and duplication or sale of all or part of any of the Data Collections is not permitted, except that material may be duplicated by you for your research use or for educational purposes in electronic or print form. You must obtain permission from the copyright holder, usually the author, for any other use. Exceptions only apply where a deposit may be explicitly provided under a stated licence, such as a Creative Commons licence or Open Government licence.

Electronic or print copies may not be offered, whether for sale or otherwise to anyone, unless explicitly stated under a Creative Commons or Open Government license. Unauthorised reproduction, editing or reformatting for resale purposes is explicitly prohibited (except where approved by the copyright holder themselves) and UEA reserves the right to take immediate 'take down' action on behalf of the copyright and/or rights holder if this Access condition of the UEA Digital Repository is breached. Any material in this database has been supplied on the understanding that it is copyright material and that no quotation from the material may be published without proper acknowledgement.

Contents

Abstract ii
Contents ……………..……………... iii
Table of tables and figures vi
Chapter contributions ix
Acknowledgments x

Chapter 1: General introduction

1.1 Genetic variation ………... 2

1.1.1. Evolutionary forces shaping genetic variation

1.1.2. Genetic variation in small populations

1.2. Pathogens as a selective pressure ………………………………………………… 5

1.2.1. Pathogen-mediated balancing selection

1.2.2. Pathogen-mediated sexual selection

1.3. The gut microbiome (GM) as a selective pressure ……………………………… 7

1.4. Immunogenetic loci ………………………………………………..………………… 9

1.4.1. Innate immune genes: Toll-like Receptors (TLRS)

1.4.2. Adaptive immune genes: The Major Histocompatibility Complex (MHC)

1.5. Seychelles Warbler: a model system for evolutionary questions …..……..… 14

1.5.2. Life history

1.5.3. Genetic variation in the Seychelles warbler

1.6. Thesis outline …………………………………………………………………………… 17

1.7. References …………………………………………………………………………..… 17

Chapter 2: Contemporary evolution of the viral-sensing *TLR3* gene in an isolated vertebrate population

2.1. Abstract …………………………………………………………………………………... 36

2.2. Introduction ……………………………………………………………………………… 37

2.3. Methods ……………………………………………………………..…………………… 39

2.3.1. Study species and system

2.3.2. Molecular methods

2.3.3. Statistical analyses

2.4. Results ……………………………………………………………………………………..… 45

2.4.1. Temporal patterns of TLR3 *variation on Cousin*

2.4.2. Testing for contemporary selection on TLR3 *variation on Cousin*

2.4.3. Hardy-Weinberg Equilibrium in fledglings sampled on Cousin

2.4.4. Spatial and temporal TLR3 *variation across islands*
2.5. Discussion ………………………………………………………………………… 52
2.5.1. Conclusion
2.6. References ……………………………………………………………………..… 57
2.7. Supplementary material ……………………………………………………….. 66

Chapter 3: Sexual selection shapes variation in the viral-sensing *TLR3* gene in a natural population
3.1. Abstract …………………………………………………………………………... 71
3.2. Introduction ……………………………………………………………………… 72
3.3. Methods …………………………………………………………………………... 76
3.3.1. Study system
3.3.2. Molecular methods
3.3.3. Statistical analysis
3.4. Results …………………………………………………………………………… 81
3.4.1. Competitive ability: Physiological condition
3.4.2. Social mate choice: Likelihood of dominance gain likelihood
3.4.3. Social mate choice: Assortative mating and/or directional mate choice
3.4.4. Genetic mate choice: Offspring EPP likelihood
3.4.5. Dominant male annual EGP and WGP success
3.4.6. Post-copulatory: Mendelian expectations
3.5. Discussion ……………………………………………………………………….. 89
3.5.1. TLR3 *associated condition dependent cues and dominance gain*
3.5.2. TLR3 *mediated social mate choice through assortative mating*
3.5.3. TLR3 *dependent genetic mate choice*
3.5.4. Mendelian inheritance and post-copulatory selection
3.5.5. Conclusion
3.6. References ……………………………………………………………………..… 95

Chapter 4: Immunogenetic variation shapes the gut microbiome in a wild vertebrate population
4.1. Abstract ………………………………………………………………………..… 107
4.2. Introduction …………………………………………………………………….. 108
4.3. Methods ………………………………………………………………………… 111
4.3.1. Study species and sample collection
4.3.2. Molecular methods
4.3.3. Statistical analysis
4.4. Results …………………………………………………………………….…... 120

4.4.1. Seychelles warbler GM profile
4.4.2. MHC characteristics
4.4.3. The effect of MHC and other host variables on GM alpha diversity
4.4.4. The effect of host variables on GM
4.4.5. The influence of host variables on the abundance of specific ASVs
4.5. Discussion 128
4.5.1. MHC variation in the Seychelles warbler
4.5.2. Does MHC variation shape GM variation?
4.5.3. Can the GM drive the evolution of immune genes?
4.5.4. Effects of age, sex, and field period on the GM
4.5.5. Conclusion
4.6. References 138
4.7. Supplementary material 150

Chapter 5: General discussion
5.1. Thesis synthesis 159
5.2. Understanding Immunogenetic variation in the Seychelles warbler 159
5.3. Contemporary selection of *TLR3*.. 161
5.4. Is *TLR3* under sexual selection? 164
5.5. Could the gut microbiome be a selective pressure? 167
5.6. Conclusion 169
5.7. References 170

Table of tables and figures

Table 2.1: Time-dependent Cox Regression modelling to test the effects of *TLR3* genotype on bi-annual survival in the Seychelles warbler population (*n* = 517) on Cousin **46**

Table 2.2: Reproductive success in male and female Seychelles warblers in relation to *TLR3* genotype: **A**) Lifetime reproductive success for all birds, **B**) Reproductive success controlling for longevity for birds that survived to adulthood **49**

Table 2.3: Allelic differentiation of one *TLR3* SNP in the five isolated island populations of the Seychelles warbler between: **A)** two time points for the same island, and **B**) between different pairs of islands using the most recently sampled data **51**

Table S2.1: Information on translocations and follow up periods for five Seychelles Warbler populations **66**

Table S2.2: Time-dependent Cox Regression modelling to test the effects of *TLR3* allele presence on bi-annual survival in the Seychelles warbler population (*n* = 517) on Cousin **66**

Table S2.3: Reproductive success in male and female Seychelles warblers in relation to *TLR3* allele presence: **A**) Lifetime reproductive success for all birds, **B**) Reproductive success controlling for longevity for birds that survived to adulthood **67**

Table S2.4: Testing Hardy-Weinberg Equilibrium and inbreeding (F_{IS}) at one*TLR3* SNP from all Seychelles warblers first caught before 3 months of age from the Cousin population **68**

Table S2.5: Allele frequencies, sample size, F_{is}, and HWE, from one *TLR3* SNP across two time points in five Seychelles Warbler populations **68**

Table 3.1: Hypotheses for whether TLR3 genotype could be shaped by sexual selection, as well as positive selection in the Seychelles warbler. Table shows potential mechanisms tested, what effects were assessed, and whether any TLR3 related associations were found **75**

Table 3.2: Predictors of **A)** body mass and **B**) size for adult male and female Seychelles warblers **82**

Table 3.3: Likelihood of gaining a dominant breeding position before two years of age for male and female Seychelles warblers in relation to *TLR3* Genotype **84**

Table 3.5: Likelihood of offspring extra-pair paternity likelihood based on the social male and mother characteristics in the Seychelles warbler in relation to *TLR3* genotype **86**

Table 3.6: Annual extra- (EGP) and within- (WGP) group paternity gained by dominant male Seychelles warblers in relation to *TLR3* genotype **88**

Table 4.1: Primer sequences used for MHC sequencing **113**

Table S4.1: Testing repeatability of MHC-I (n = 26) and MHC-II (n = 24) genotyping for different dominant frequency thresholds while keeping minimum amplicon frequency at 0.3% **150**

Table S4.2: Testing repeatability of MHC-I (*n* = 26) and MHC-II (*n* = 24) genotyping for different minimum amplicon frequencies while keeping minimum dominant frequency threshold at 25% **150**

Table S4.3: The effect of host intrinsic and extrinsic variables on GM diversity in the Seychelles warbler (n = 195) for three metrics of alpha diversity: Shannon diversity, Chao 1(log transformed) and Faiths phylogenetic diversity (log transformed): **A**) including the presence/absence of MHC alleles or **B**) MHC diversity **151**

Table S4.4: PERMANOVA tests investigating the effect of host variables on the composition of the Seychelles warbler gut microbiome (*n* = 195) **152**

Fig 2.1: Allele frequency change at a nonsynonymous *TLR3* SNP in the Cousin population of the Seychelles warbler over 25 years (1993 - 2018) **45**

Fig 2.2: Effect of *TLR3* genotype on survival in the Seychelles warbler population on Cousin (*n* = 517) **47**

Fig 2.3: Effects of *TLR3* genotype on reproductive success in the Cousin population of the Seychelles warbler: **A**) Lifetime reproductive success (offspring surviving >3 months) for all birds (*n* = 487), **B**) Rate of reproduction (i.e., offspring surviving to >3 months/ longevity for focal birds that survived to adulthood (*n* = 323)) **48**

Fig 2.4: Change in the minor allele frequency (C) of the nonsynonymous *TLR3* SNP between two time points in the five isolated island populations of the Seychelles warbler **50**

Fig S2.1: Observed and expected values for *TLR3* genotypes in the Cousin population of the Seychelles warbler for **A**) all individuals first caught before 3 months of age (*n* = 591), **B**) individuals first caught before 3 months of age which survive to adulthood (*n* = 380) **69**

Fig 3.1: The effect of *TLR3* genotype on body condition (mass) **in A**) adult male (*n* = 97), and **B**) adult female (*n* = 120) Seychelles warbler **81**

Fig 3.2: Effect of *TLR3* genotype and natal group size on likelihood of gaining dominance by 2 years of age for male Seychelles warblers **83**

Fig 3.3: The association between *TLR3* genotype and social mate-choice in Seychelles warblers. **A**) Assortative social mating based on pair *TLR3* similarity (purple), or directional female social pair preference of **B**) $TLR3^{AA}$ females (dark green), **C**) $TLR3^{AC}$ females (light green), **D**) $TLR3^{CC}$ females (yellow), based on the number of $TLR3^{A}$ alleles in the social male **85**

Fig 3.4: The likelihood of offspring being extra-pair in relation to the parent's *TLR3* genotype in the Seychelles warbler: **A**) social males (*n* = 107), **B**) mothers (*n* = 137) **86**

Fig 3.5: Annual **A**) extra-group paternity (EGP), and **B**) within-group paternity (WGP) gained by dominant male Seychelles warblers in relation to their TLR3 genotype (n= 96).. **87**

Fig 3.6: Observed and expected values of offspring's *TLR3* genotype based on parental *TLR3* genotypes, when at least one parent is *TLR3* heterozygous (*n* = 173) in the Seychelles warbler **88**

Fig 3.7: Observed and expected values of whether offspring paternally inherit a TLR3^A or *TLR3*^C allele when genetic father is heterozygous ($TLR3^{AC}$), and the mother is homozygous ($TLR3^{AA}$ or $TLR3^{CC}$) (n = 52) in the Seychelles warbler **89**

Fig 4.1: The relative abundance (%) of bacterial **A)** phyla and **B)** classes in 281 faecal samples, collected from 224 Seychelles warblers **122**

Fig 4.2: Variation in MHC-I exon 3 and MHC-II exon 2 in 244 Seychelles Warblers **124**

Fig 4.3: Effects of the presence of specific MHC alleles, TLR3 genotype and genome-wide heterozygosity and other host variables on alpha diversity in 195 Seychelles warblers **125**

Fig 4.4: Beta diversity of Seychelles warbler gut microbiome composition according to the presence or absence of the MHC-I **A)** *Ase-ua7* allele, **B)** *Ase-ua11* allele, or **C)** *Ase-ua1/10* allele **127**

Fig 4.5: Differentially abundant ASV's (P_{adj} < 0.01) in the gut microbiome of Seychelles warblers, according to the presence/absence of the MHC-I alleles **A)** *Ase-ua7* **B)** *Ase-ua11* or **C)** *Ase-ua1/10* **129**

Fig S4.1: Prevalence and total abundance of all ASV's separated by phylum **153**

Fig S4.2: Sequencing repeatability of the gut microbiome tested using 37 faecal samples taken from Seychelles warblers, sequenced twice **154**

Fig S4.3: Beta diversity of Seychelles warbler gut microbiome composition in different age classes **155**

Fig S4.4: Differentially abundant ASV's (P_{adj} < 0.01) in the gut microbiome of Seychelles warblers between different age categories. **156**

Fig S4.5: Differentially abundant ASV's (P_{adj} < 0.01) in the gut microbiome of Seychelles warblers between seasons **157**

Chapter contributions

At time of submission, one of the three data chapters in this thesis has been published in a peer-reviewed journal. I am the lead author on all manuscripts, and I have been responsible for the largest contribution. Below, I provide co-author information, and detail the specific contributions I have made to each data chapter

Chapter 2: Davies CS, Taylor M, Hammers M, Burke T, Komdeur J, Dugdale H, Richardson DS. 2021. *Molecular Ecology.* **doi**:10.1111/mec.15914

I performed some of the lab work, compiled and analysed the data, and drafted the manuscript (80%)

Chapter 3: Davies CS, Burke T, Komdeur J, Dugdale H, Richardson DS (in prep)

I performed some of the lab work, compiled most of the data, analysed the data, and drafted the manuscript (80%)

Chapter 4: Davies CS, Worsley S, Maher K, Burke T, Komdeur J, Dugdale H, Richardson DS (in review: preprint **doi**: 10.21203/rs.3.rs-703361/v1)

I performed the field work, lab work (apart from the microbiome library prep), bioinformatics, compiled and analysed the data, and drafted the manuscript (80%)

Acknowledgments

First of all, many thanks to my supervisors David Richardson, Hannah Dugdale and Martin Taylor. You have been incredibly supportive, and helpful throughout the PhD. I've really enjoyed my time as a PhD student, which would not have been possible without your teaching and encouragement, and I am extremely grateful for all the work you did. Special thanks goes to Dave, for always being available when needed, and being an all-round excellent supervisor

As well as Dave and Hannah, I was also fortunate enough to have Terry Burke and Jan Komdeur as PIs in the Warbler group - thank-you for always being enthusiastic about my research, and for having unique points of view. Thank-you also to everyone who has been a member of the Wobbler group during my time: Sara, Alex, Sjouke, Marco, Michaela, Owen, Dave, Tara, Ellie, Janske, Monica, Tom, Sarah, Char, Martijn and Kat – the wobbler group was a great community to have been part of and I always looked forward to warbler meeting downtime thanks to all of you. Thank you also to the 100+ people who have helped collect data on the Seychelles warbler project over the last 35 years –this PhD would not have been possible without it!

I was lucky enough to spend a lot of time on Cousin Island for fieldwork, this would not have been half as fun without all the fabulous people. Kat Bebbington – for showing a non-birder how to do everything, and for all the burger nights. Charlotte Bartleet-Cross – I can never think of game of thrones (or mosquitos) without thinking of you and Cousin! Martijn Hammers – for always being enthusiastic, bringing great fieldwork snacks, and for being a great co-author. Dave and Terry – for being great fun in the field and bringing much needed energy near the end of the season. The A-team – Tom Brown and Arne Van Eerden – I wish I could properly translate the sound of the flycatcher on this page! The Island itself couldn't function without Nature Seychelles and all the fantastic work the wardens and managers do – special thanks to Jovani, Eric, Jan, Kara, and Cheryl for help making fieldwork on Cousin run efficiently and for non-warbler chats!

As well as fieldwork – I also spent some time in Sheffield at the NERC Biomolecular Analysis Facility (thanks for NBAF for funding this!). I would particularly like to thank Gavin Horsburgh for showing me how to do all the lab work, and always having planet rock on. Katy Maher for helping me with bioinformatics – and for help troubleshooting when things didn't go to plan! I am also grateful for the support of Deborah Dawson, Rachel Tucker and Celine Pagnier, who helped keep everything running smoothly.

Lastly, all the people at Sheffield University, and my air BnB hosts who made everything so much more fun.

UEA has been a great place – the department and city are incredibly friendly, and I have immensely enjoyed my time here. I particularly want to thank Marie-Elena for being the font of all molecular lab (and chicken) knowledge. Thanks also for Dave Wright for teaching me how to do molecular things, and Sarah Worsley for being my Qiime guru! I have made many friends at UEA – but I would like to add special mention to Becky, Ellen and Ents – you have all been amazing friends – I look forward to many more coffee (or cocktail) breaks in the future! Jen for early morning ringing, and late-night drinking. And Dave C – for being as enthusiastic about cheese and wine nights as I am. My fellow caterpillar office mates for always being happy to answer stupid questions. Thanks to everyone who came to the beehive pub quiz and biodrinks – especially Ryan, Mike, Kris, Jessie, Stewert, Jenn, Mabel and Alex - you all made everything so much more fun. For everyone on floor 01 – there are too many of you too list here – but you all made the PhD journey so much more memorable – it really is a fantastic department! Erin, Alicia, Becky, Tom, Agatha, Rob, and of course Dougle – thank-you for going through the PhD journey with me, and for always being up for a BBQ! Harriet and Finn – lockdown was better with people to share it with! Sarah B – thanks for being an awesome housemate and fellow plant enthusiast– lockdown was way more fun with drag race! And the Couplands – thanks for always having an open and welcoming house!

Lastly, I would like to thank my Mum for always being supportive in whatever I decide to do – even if its living half-way round the world studying obscure animals. And Dad - I’m really sorry you couldn’t be here to see me finally finish my PhD – you will always be missed. You’ve both helped so much to get me to where I am today, and I will always be grateful for that.

Chapter 1

General introduction

Seychelles warbler from Cousin Island

1.1. Genetic variation

Genetic variation underpins heritable variation in phenotypic traits and provides the substrate for evolution. Genetic variation within populations thus affords these populations with the ability to evolve in response to abiotic or biotic change – termed adaptive potential (Fisher 1930). New genetic variation ultimately arises by mutation (in the wider sense) as a result of point mutations, recombination, gene conversion, and gene or genome duplication (Lande 1995, Martinsohn et al. 1999). Once present, genetic variation can be erased through directional selection and/or drift – resulting in a loss of heterozygosity within individuals, and adaptive potential within populations (Lacy 1987, Lande 1988). Thus, resulting in reduced fitness at the individual, and eventually at the population level (Reed and Frankham 2003). Consequently, a reduction in genetic variation can have detrimental impacts upon the survival of an individual and the persistence of populations and species (Crnokrak and Roff 1999, Keller and Waller 2002, Bouzat 2010). Alternatively, once present, genetic variation can be maintained in a population by balancing selection – an umbrella term for any mechanism or combination of mechanisms by which genetic variation is maintained within a population (Dobzhansky 1955). Maintenance of genetic variation within a population, can play a crucial role in maintaining a populations adaptive potential. Therefore, understanding where and how evolutionary forces shape genetic variation, and the downstream consequences these have for individuals and populations, is important from both an evolutionary and conservation perspective (Frankham 1996).

1.1.1. Evolutionary forces shaping genetic variation

A number of evolutionary forces act separately and simultaneously to shape genetic variation in a population. These can be ‘neutral’ processes such as genetic drift or gene flow, (Wright 1931, Lande 1976, Lacy 1987). Genetic drift is when alleles are arbitrarily lost or driven to fixation within a population by chance (Wright 1931). Because of random sampling between generations not all alleles are passed down to the next generation, ultimately resulting in a decrease in genetic variation within a population (Kimura 1968, Lande 1976, Charlesworth 2009). Gene flow occurs when individuals migrate from one population to another and successfully reproduce, thus introducing alleles into the population (Slatkin 1985), resulting in changing allele frequencies (for example Chen et al. 2019). Migration can be random, or be phenotype specific - therefore unlike genetic drift, gene flow can be a non-random process (Chesser 1991). The random nature of genetic drift means that different alleles will increase or decrease in frequency in different populations, causing increasing genetic divergence between populations. In contrast,

gene flow homogenises allele frequencies between interconnected populations, thus leading to a reduction in genetic divergence between populations (Slatkin 1987).

Genetic variation can also be shaped by the deterministic process of selection (Lande 1976). Individuals with functional genetic variation will exhibit different phenotypes, which will then be selected for within a population (Fisher 1930), thus adaptation can occur (Darwin 1859). Natural selection is the process whereby traits which influence fitness are selected for, however if a specific trait influences competition over mates or their gametes then it is under sexual selection (Darwin 1896, Andersson 1994, Shuker 2010). Natural and sexual selection can shape genetic variation through various mechanisms. Directional or purifying selection can drive beneficial alleles to fixation and remove disadvantageous alleles, thus resulting in reduced genetic variation (reviewed in Mukherjee et al. 2009). Conversely, balancing selection maintains genetic variation via a combination of different potential mechanisms. First, the heterozygote advantage stipulates that individuals who are heterozygote have greater fitness than those who are homozygote, therefore maintaining genetic variation (Dobzhansky 1955). Such heterozygote advantage can occur by dominant selection - whereby the heterozygote will have the same fitness advantage as the fittest homozygote (Penn et al. 2002). Alternatively, if a heterozygous allele is under over-dominant selection then the heterozygote will benefit from the combined synergistic effect of the two different variants, and thus have a greater fitness advantage than the fittest homozygote (Hughes and Nei 1988). Second, rare-allele advantage (also known as frequency-dependent selection) occurs when rare alleles confer a greater fitness advantage, than do common alleles (Ayala and Campbell 1974). Rare-allele advantage is usually the result of co-evolution between competing species (e.g., predator- prey, Bond and Kamil 2002), or conspecifics (e.g., competition for resources, Fitzpatrick et al. 2007). Rare alleles can be created by new alleles that arise through mutation, or old alleles which are at a low frequency in the population. As these rare alleles increase in frequency in a population, their selective advantage is diluted. Consequently, these alleles are no longer novel and thus stop being selected for. Instead, possession of a different rare allele will have a selective advantage (Ayala and Campbell 1974). Lastly, the fluctuating selection hypothesis proposes that spatio-temporal variation in the selective pressure can maintain genetic diversity (Hill et al. 1991). For example, an individual carrying an allele conferring a fitness advantage at one particular time and/or place will have increased fitness, but the same allele could confer a reduced fitness advantage at a different time and/or place. This results in different subsets of alleles being selected for at different spatio-temporal points, thus increasing variation across sub-populations.

Overall, the potentially combined and interacting effect of one or more of these mechanisms of balancing selection will lead to multiple genetic variants being maintained in a population, thus counteracting the loss of diversity through drift and directional selection. Gene flow homogenises genetic variation between populations, and so has important implications for rare-allele, and fluctuating selection mechanisms of balancing selection. Lastly, genetic variation can be maintained by a trade-off between natural and sexual selection (Johnston et al. 2013, Slade et al. 2019).

1.1.2. Genetic variation in small populations

Problems arising from the loss of genetic variation are particularly exacerbated in small populations (Frankham 1996). By their very nature, populations of conservation concern are often small or fragmented. Consequently, they are more sensitive to sudden environmental and human-induced perturbations, such as habitat degradation, or the introduction of new pathogens (reviewed in Jamieson 2015). To effectively conserve small populations, it is important to know what, and how different evolutionary forces are shaping genetic variation in these population's (Lacy 1987, Sutton et al. 2011).

Populations which have undergone bottlenecks – (i.e., have been reduced down to a small number of individuals (Nei et al. 1975)) – often have reduced genetic variation, and therefore reduced individual and population level fitness (Hale and Briskie 2007). Reduced genetic variation in bottlenecked populations, can occur by several interacting means. Individuals in small populations may have no choice but to mate with relatives, due to the lack of non-kin mates, resulting in inbreeding. Inbreeding causes individual and population-wide negative fitness consequences (e.g., Walling et al. 2011) - termed inbreeding depression (Crnokrak and Roff 1999). This is due to increasing homozygosity resulting in the expression of deleterious recessive alleles (Lynch et al. 1995, Charlesworth and Charlesworth 1999), or loss of heterozygote advantage – where heterozygotes have greater fitness than either homozygous genotypes (Charlesworth and Charlesworth 1990). Secondly, founder effects (i.e., the reduction of genetic variation in a population due to establishment by a small number of individuals from a larger population (Mayr 1954)) determine the type and amount of genetic variation retained in a population, depending upon the genetic makeup, and amount of genetic diversity of the original founders of the population (e.g., Dlugosch and Parker 2008, Kolbe et al. 2012). Lastly, the amount of genetic variation present in a population directly affects its adaptive potential. Decreasing genetic variation will likely result in loss of beneficial functional genes or alleles. Populations with little, or reduced genetic variation

will be less able to adapt to changing circumstances and thus are more likely to go extinct (reviewed in Bouzat 2010).

When effective population sizes are reduced, such as in small populations, the role of genetic drift is exaggerated. The importance of genetic drift as the primary mechanism shaping genetic variation in small populations has been well documented (Lacy 1987, Miller and Lambert 2004, Grueber et al. 2013). The extent and role of genetic drift can be further confounded in situations where genetic variation may already be reduced due to inbreeding depression and founder effects (Nei et al. 1975, Frankham 1996).

1.2. Pathogens as a selective pressure

Pathogens are a very diverse and adaptable group of organisms, constantly evolving to better their own fitness at the detriment of their host (Freeland 1976). By definition, pathogens detrimentally influence the survival and reproductive success of their host, resulting in disease and mortality (Daszak et al. 2000). Consequently, they are powerful drivers of evolutionary change (Haldane 1992). Hosts evolve defences to protect themselves from pathogens, while pathogens, in turn, evolve ways to overcome these defences, resulting in antagonistic coevolution between host and pathogen (Slade and McCallum 1992, Woolhouse et al. 2002). Two interacting underlying drivers of balancing selection are particularly important in the maintenance of genetic variation, that of pathogen-mediated balancing selection and sexual selection (Hedrick 1998, Bernatchez and Landry 2003, Spurgin and Richardson 2010, Ejsmond et al. 2014).

1.2.1. Pathogen-mediated balancing selection

Pathogen-mediated selection has often been put forward as an important driver of balancing selection (see 1.1.1 (Bernatchez and Landry 2003)), with considerable evidence that this maintains diversity at immune genes (Hedrick 2002, Ferrer-Admetlla et al. 2008, Eizaguirre et al. 2012). For example, individuals which are heterozygous at immune related genes, will have a greater likelihood of recognising and/or overcoming a pathogenic infection, and hence have increased resistance to pathogens than if homozygous (Doherty and Zinkernagel 1975). In this context, there is support for both dominant and over-dominant selection maintaining genetic diversity across metapopulations (Takahata and Nei 1990, Oliver et al. 2009), although dominant selection alone is thought to be unable to maintain high levels of diversity within individual populations (reviewed in Spurgin and Richardson 2010). Secondly, rare alleles can

confer greater resistance to pathogens (Phillips et al. 2018). As these rare alleles increase in frequency in a population, the pathogens they detect become under increasingly stronger selection to evolve to evade their effects. Consequently, these alleles will no longer confer resistance to the pathogen and thus stop being selected for. Instead, possession of a different rare allele will have a selective advantage. Overall, this results in a cyclical, co-evolutionary arms race between pathogen and alleles (Takahata and Nei 1990, Slade and McCallum 1992). Lastly, when variation in pathogen type or abundance changes depending on geographical or temporal variation, intensity of directional selection upon the hosts genotype will also change (Hedrick 2002) – consistent with fluctuating selection. Pathogen-mediated fluctuating selection can be driven by trade-offs, whereby resistance to pathogens comes at a cost to the hosts. When pathogens are present, hosts with resistant alleles are favoured, but when pathogens are not present the resistant allele will be detrimental to the host (Tellier and Brown 2011).

All three of the above proposed mechanisms, along with pathogen-mediated sexual selection (see 1.2.2 below) are able to act singularly or in concert together. Thus, gaining a definitive understanding of the relative importance of the different mechanisms in driving balancing selection is extremely difficult (reviewed in Spurgin and Richardson 2010). For example, individuals which are carrying rare alleles are also likely to be heterozygous as their low frequency in the population means it is unlikely for an individual to inherit two copies. Moreover, the intensity of selection a pathogen exerts can fluctuate due to differences across temporal and spatial scales, further confounding the importance that rare allele and/or heterozygote advantage confers at a given point. Studies which attempt to tease apart these mechanisms of pathogen-mediated selection do so by contrasting functional immuno-genetic variation with that expected under neutrality (e.g., Landry et al. 2001b), or by relating immuno-genetic variation with pathogen load (e.g., Bonneaud et al. 2006b).

1.2.2. Pathogen-mediated sexual selection

There is substantial evidence that pathogen-mediated sexual selection also acts (on top of the other mechanisms of balancing selection) to maintain variation at immune genes (reviewed in Apanius et al. 1997, Penn and Potts 1999). Pathogen mediated sexual selection can act prior to mating e.g., via inter-sexual mate-choice (Hamilton and Zuk 1982) and intra-sexual competition (e.g., Freeland 1976). Or post-mating e.g., through cryptic mate-choice (Eberhard 1996), sperm competition (Birkhead and Møller 1998), and sperm/egg incompatibility (Fernandez, Cooper et al. 1999). Mate-choice can

potentially help maintain immuno-genetic variation via three different routes. First, individuals may choose mates who carry ’good genes’ which confer resistance to current pathogens, thus selecting for rare ‘resistant’ alleles (reviewed in Neff and Pitcher 2005, Kempenaers 2007). The presence of sexually selected traits or ornaments may be important here, due to their capability to be used as an honest signal of condition (Hamilton and Zuk 1982). For example, those males carrying the alleles which confer best pathogen resistance should be in best condition and have most extravagant secondary sexual traits (Von Schantz et al. 1996, Moore and Wilson 2002, Eizaguirre et al. 2009, Dunn et al. 2013). Alternatively, the best condition may be conferred by heterozygosity at immune genes. So, choice for the best condition mate should also increase the heterozygosity of offspring (Landry et al. 2001, Reusch et al. 2001, Richardson et al. 2005, Brouwer et al. 2010). Lastly, being able to recognise which potential mates have dissimilar immune genes compared to themselves, can be beneficial for two reasons. Dissimilar mating can help alleviate the associated costs of inbreeding depression by preventing mating with close kin (Tregenza and Wedell 2000). Or, dissimilar (compatible) mating by mating can maximise offspring heterozygosity, and fitness (Mays Jr and Hill 2004). Either way, preference for dissimilar mates can occur both pre-mating and post-mating (Potts and Wakeland 1990, Gillingham et al. 2009, Løvlie et al. 2013) and odour cues have been suggested to be the primary mode of communication (Milinski 2006).

The relative roles of pathogen-mediated balancing selection and sexual selection in maintaining immuno-genetic diversity are hard to disentangle as they can work in concert, reinforcing host-pathogen co-evolution (Ejsmond et al. 2014). Alternatively, they can trade-off against each other together, both forces effectively cancelling each other out (Slade et al. 2019). For example, antagonistic pleiotropy between genes, results in trade-offs between reproduction and survival (Johnston et al. 2013). Therefore, unless all traits which could be under selection are measured, the functioning mechanism, or mechanisms could be missed or misinterpreted (Loiseau et al. 2008).

1.3. The gut microbiome (GM) as a selective pressure

Animals evolved in a microbial world and should be studied accordingly (Amato 2013). The complex microbial community that interacts with any given individual - comprised of bacteria, archaea, viruses, and microbial eukaryotes – is collectively known as the host’s microbiome. These microbiomes exist at the interface between animals and their environments, playing a fundamental role in many host biological processes (McFall-

Ngai et al. 2013). As the microbes within these communities can be pathogenic, mutualistic, or commensal, they can be beneficial or detrimental to the host, thus they can play an important role in host evolution (Kolodny et al. 2020). Thus, the microbiome can act as a selective pressure, shaping host phenotypes (reviewed in Koskella and Bergelson 2020, Rowe et al. 2020). Ultimately resulting in host-microbiome co-evolution, adaptation, or speciation (Greene et al. 2020, Moeller and Sanders 2020).

In the vertebrate gut, the gut microbiome (GM) contributes to many key biological processes from enabling nutrient uptake (Hooper et al. 2002), to pathogen defence (Pickard et al. 2017). However, despite the recent expansion in GM studies – particularly those focusing on bacteria, the relative importance that host genetic factors have in governing the diversity and composition of the GM is still unclear. Additionally, it has become apparent that captivity/unnatural environments radically affect an organism's microbiome (Clayton et al. 2016). By design, captive environments are carefully controlled, and so exposure to environmental and microbial complexity is reduced compared to wild populations. Thus, not all factors that govern the diversity and composition of the GM would be present – including selective factors shaping co-evolution of host-microbiome. The importance of environmental context is particularly crucial when considering selection of the hosts immune system which has evolved with a variety of pathogens (Maizels and Nussey 2013). Therefore, more studies on wild populations are needed, which incorporate the complexities of real life (reviewed in Amato 2013, Hird 2017).

A multitude of environmental and host factors appear to shape individual differences in GM in wild populations including diet (Maurice et al. 2015), sex (Stoffel et al. 2020), age (Videvall et al. 2018), and sociality (Raulo et al. 2021). Although the role of host genetics has been less well studied (Kurilshikov et al. 2017), particularly in wild systems, evidence does suggest that host genotype influences individual GM variation (Spor et al. 2011, Suzuki et al. 2019). Specifically, variation in key immune genes has been linked to GM characteristics (reviewed in Kurilshikov et al. 2017). Likewise, GM variation can directly affect various host traits (Suzuki 2017) including modulating the severity of infectious diseases (Villarino et al. 2016), host immune function (Round and Mazmanian 2009), and ultimately, hosts survival (Benskin et al. 2015), thus providing the potential for evolutionary adaptation. Therefore, if immune genes can regulate fitness through modulation of the GM, then the GM can also act as a driver of selection on immune genes (Koskella and Bergelson 2020). The same mechanisms that maintain genetic variation via pathogen mediated selection, i.e., balancing selection and sexual selection, could also shape host variation via the GM.

1.4. Immunogenetic loci

Hosts try to avoid pathogens through behavioural modifications and external barriers (reviewed in Loehle 1995). However, if this fails then an effective immune system is needed to protect and mitigate against pathogenic intrusions. Consequently, to effectively combat the myriad of pathogenic infections they may be exposed to, hosts have evolved a complex and flexible immune system (Boehm 2012). Thus, genes involved in controlling this immune system are good candidate genes for studying how evolutionary processes shape genetic variation (reviewed in Sommer 2005). In particular, immune receptor genes have a greater rate of nonsynonymous than synonymous substitution compared to other immune gene categories, indicating greater rates of positive selection (Ekblom et al. 2010).

In vertebrates, two main branches of the immune system work together to mount an immune response; the evolutionary ancient innate immune system, and the more recently evolved adaptive immune system (Delves and Roitt 2000). Following pathogenic infection, the innate immune system is the host's first response to pathogens, as such it enables a broad defence to be mounted against an assortment of organisms (Aderem and Ulevitch 2000). When pathogens are first encountered and recognised, the innate immune system triggers an immediate, non-specific response, preventing or curbing infections, before pathogen-specific acquired (adaptive) processes can be activated (Iwasaki and Medzhitov 2010). This immediate response can be particularly important when novel pathogen outbreaks occur (Bonneaud, Balenger et al. 2012). Pathogens are recognised through detection of pathogen-associated molecular patterns (PAMPs), which are highly conserved, non-specific, molecular structures integral to the survival of the pathogen, and which are thus difficult for them to alter (Smith et al. 2003). PAMPs are microbe specific, therefore there is a clear distinction between recognition of 'self' and 'non-self' (Medzhitov 2001, Akira et al. 2006), and are recognised by pattern recognition receptors. When triggered, these receptors activate the inflammatory response, mediate phagocytosis, and/or activate other immunity genes further down the immune cascade (Aderem and Ulevitch 2000, Ferrer-Admetlla et al. 2008). Currently, five major classes of pattern recognition receptors have been described (Takeuchi and Akira 2010). Of these, the Toll-like receptors (TLRs) are the most extensively studied (Netea et al. 2012).

1.4.1. Innate immune genes: Toll-like Receptors (TLRS)

TLRs are type 1 trans-membrane receptors, characterised by three major domains: an extracellular leucine-rich repeat domain; a short conserved transmembrane domain; and a highly conserved intracellular signalling Toll/IL-1 receptor domain (Medzhitov 2001). Once the extracellular domain recognises and binds to a specific PAMP, the intracellular domain initiates signalling pathways, thus triggering an inflammatory immune response (Aderem and Ulevitch 2000). When triggered, TLRs are also able to stimulate genes involved in the acquired immune response, bridging the two arms of the immune system, and affording the host a vast range of responses to a suit of pathogens (Akira et al. 2006, Iwasaki and Medzhitov 2010).

TLRs can be broadly separated into those that are localised on cell surface membranes (non-viral), or those localised on endosomes in intracellular vesicles (viral). More specifically, as vertebrate TLRs recognise various pathogen groups, they can be further divided into six-eight major subfamilies depending upon what PAMPs they bind to (Roach et al. 2005, Brownlie and Allan 2011, Liu et al. 2019). For example, TLR1 binds to bacterial lipoproteins, TLR3 binds to viral double stranded RNA (Barton 2007), TLR4 binds to bacterial lipopolysaccharide (Poltorak et al. 1998), TLR5 binds to bacterial flagellin (Brownlie and Allan 2011), and TLR7 binds to single stranded RNA (Roach et al. 2005).

While the majority of each TLR molecule is structurally conserved across individuals and species (Roach et al. 2005), variation in the leucine-rich repeat domain region which binds to PAMPs does occur, resulting in differential pathogen recognition (Werling et al. 2009). Indeed, specific polymorphisms within the leucine-rich repeat domain of TLR's have been associated with differential host immunity (Villaseñor-Cardoso and Ortega 2011), resulting in either increased resistance (Antonides et al. 2019, Armstrong et al. 2019), or susceptibility (Hawn et al. 2003, Schröder and Schumann 2005, Kloch et al. 2018) to various pathogens.

TLR genes were initially believed to have been relatively evolutionary conserved across species (Roach et al. 2005, Chapman et al. 2016). However, recent studies have shown evolutionary forces to be shaping TLR variation at the species, and population level (Bagheri and Zahmatkesh 2018, Velová et al. 2018, Khan et al. 2019). At the population level, several studies have found that genetic drift, rather than selection is important in shaping TLR variation, particularly in small and/or bottlenecked populations, e.g., (Grueber et al. 2013, Gonzalez-Quevedo et al. 2015). However, in other vertebrate populations greater TLR variation is present than expected (Grueber et al. 2012, Grueber et al. 2015). With evidence suggesting that various modes of selection are important in

shaping TLR variation (Alcaide and Edwards 2011, Grueber et al. 2014, Gilroy et al. 2017). Thus, indicating that pathogen-mediated balancing selection could be involved in the maintenance of TLR diversity (Downing et al. 2010, Croze et al. 2016). Indeed, there is evidence that past balancing selection (Ferrer-Admetlla et al. 2008, Quemere et al. 2015, Kloch et al. 2018) as well as purifying and/or positive selection (Areal et al. 2011, Grueber et al. 2014, Nelson-Flower et al. 2018, Lara et al. 2020) has affected TLR variation in a range of vertebrate populations. However, although studies have provided evidence that selection has shaped the evolution of TLR genes at some point in the past, few studies have investigated the fitness consequences, and contemporary evolution, of TLR variation in current populations, (but see Grueber et al. 2013, Hartmann et al. 2014, Bateson et al. 2016). Furthermore, none have investigating whether sexual selection can also drive TLR variation.

1.4.2. Adaptive immune genes: The Major Histocompatibility Complex (MHC)

Compared to the innate immune response, the adaptive immune response evolved relatively recently and is only found in jawed vertebrates (gnathostomes) (Flajnik and Kasahara 2001). It involves a specialised group of host cell molecules, encoded by MHC genes, which recognise and bind antigens (including pathogen derived peptides) and presents them to T- and B-cell receptors. Once the combined MHC-peptide complex is recognised as non-self, the T and B cells are activated, triggering an appropriate immune cascade (Delves and Roitt 2000). After such a reaction, the adaptive immune system produces memory cells, enhancing future immune response (an acquired response) if the host is re-infected with that particular pathogen (Medzhitov 2001).

MHC genes can be classed into two main sub-groups (but see Gruen and Weissman 1997): Class I (MHC-I) molecules are expressed on the surfaces of virtually all somatic cells and are associated with defence against intracellular pathogens. Whereas Class II (MHC-II) molecules are only found on the surface of specific antigen-presenting cells and are primarily associated with defence against extracellular pathogens (Hughes and Yeager 1998). Both classes of MHC molecules are comprised of a receptor region called the peptide-binding region attached to an immunoglobulin ‘stalk’ which acts to anchor the molecule to the cell surface. Similar to the leucine-rich repeat domains domain of the TLR, it is this peptide-binding region that is responsible for antigen recognition (Klein 1986). However, unlike the leucine-rich repeat domains, the MHC peptide-binding region is highly specific to particular antigens, although some overlap between peptides occurs (Altuvia and Margalit 2004).

Among immune genes, the diversity and evolution of MHC genes have been particularly well studied, due in part to their extreme variability and key role in pathogen detection (reviewed in Hedrick 2002, Piertney and Oliver 2005). Specific functional variants of MHC genes, particularly those of the peptide-binding region, confer differential pathogen recognition (Thursz et al. 1995, Sin et al. 2014, Biedrzycka et al. 2018), and hence can have a direct effect upon fitness (e.g., Hill et al. 1991, Loiseau et al. 2008, Brouwer et al. 2010, Lukasch et al. 2017). Several studies on wild populations have shown that MHC genes often harbour more variation than neutral loci, suggesting balancing selection is shaping variation at this locus (Richardson and Westerdahl 2003, Aguilar et al. 2004, Schad et al. 2004). Indeed, there is evidence from different studies to support the action that all three of the main hypothesised mechanisms of pathogen-mediated balancing selection - rare-allele advantage (Bonneaud et al. 2006, Alcaide et al. 2008), heterozygote advantage (Penn et al. 2002, Oliver et al. 2009), and fluctuating selection (Landry and Bernatchez 2001, Charbonnel and Pemberton 2005), or a combination of the three (Westerdahl et al. 2005) - act to maintain genetic variation at the MHC (Hedrick 1994).

Sexual selection is underpinned by phenotypic secondary sexual traits or cues, which are often condition dependent (Hamilton and Zuk 1982). Various secondary sexually selected traits have been found to be linked to MHC variation, including weapons (Ditchkoff et al. 2001), ornaments (Von Schantz et al. 1996, Dunn et al. 2013), colouration (Jäger et al. 2007, Setchell et al. 2009), and olfactory cues (Grieves et al. 2019, Grogan et al. 2019). Hence, MHC variation can provide a genetic foundation for studying intra- and inter-sexual selection (reviewed in Milinski 2006). MHC variation can be selected for as a result of intra-sexual selection through male-male competition (Setchell et al. 2010). However, the focus of most studies has been on the role of MHC in intersexual competition (see Kamiya et al. 2014, Brandies et al. 2018) to understand the possible genetic benefits of female-based mate choice. Such benefits can potentially occur via three non-mutually exclusive process (Penn and Potts 1999, Milinski 2006). Preference could exist for mates with specific MHC alleles or 'good genes', which confer resistance to current pathogens (Eizaguirre et al. 2009). Alternatively, preference could be for MHC heterozygous (or MHC diverse) mates, i.e., those that are heterozygote at any given loci, or across multiple duplicated MHC loci - hence overall MHC diversity - (Landry et al. 2001). Lastly, there may be choice for MHC compatible mates, i.e., mates are chosen based on MHC similarity or dissimilarity to themselves (reviewed in Tregenza and Wedell 2000, Neff and Pitcher 2005). Preference for compatible mates may take the form of disassortative mating, where preferred partners are those that are most dissimilar (e.g., Gillingham et al. 2009, Løvlie et al. 2013). At one end of the scale, disassortative

mating would prevent mating with close kin (who would have a similar MHC) - and thus avoid inbreeding depression in offspring (reviewed in Brown and Eklund 1994). However, it is thought that the benefits of disassortative mating could go beyond that to maximising offspring MHC diversity - and thus pathogen resistance overall (Penn, Damjanovich et al. 2002, Setchell et al. 2010, Evans et al. 2012). In contrast, MHC assortative mating has also been observed in some species (e.g., Bonneaud et al. 2006, Bos et al. 2009, Sin et al. 2015) – where it would counteract balancing selection (Slade et al. 2019). It is also important to note that various other studies have found no support for MHC-mediated sexual selection (Paterson and Pemberton 1997, Westerdahl 2004, Yu et al. 2018).

There are also potential drawbacks to high level of MHC variation in individuals. For example, MHC variation has been associated with almost all autoimmune diseases, although the mechanisms underlying this are still not fully understood (reviewed in Fernando et al. 2008). However, the primary reason why high MHC diversity may be detrimental is that individuals with increased MHC variation will be subject to greater costs occurred during T-cell maturation, as more T-cells which bind to self-peptide MHC molecules are killed, effectively reducing their pathogen recognising ability (Lawlor et al. 1990, Nowak et al. 1992). This has implications when considering the importance of MHC-based mate choice. It has been argued that individuals should choose mates to optimise MHC variation at some intermediate level rather than maximising offspring MHC variation (Nowak et al. 1992). Indeed, some evidence for this has been put forward in three-spined sticklebacks, *Gasterosteus aculeatus* (Reusch et al. 2001, Kalbe et al. 2009).

There are, however, complications with studying the MHC. Mechanisms such as gene conversion, recombination, and duplication, which are common at the MHC, have made it extremely difficult to identify, separate out, and characterise variation at individual loci (Spurgin and Richardson 2010). While recent expansion of next-generation sequencing studies provides efficient ways to characterise large number of individuals by amplicon sequencing (Biedrzycka et al. 2017), or in-depth analysis of MHC structure using long-read sequencing (He et al. 2020), there is still much to be done (reviewed in O'Connor et al. 2019). Immunogenetic loci which exhibit reduced complexity compared to the MHC (e.g., TLR genes - see 1.3.1 above), may provide more tractable candidates with which to study how evolutionary forces, including natural and sexual selection, shape variation in wild populations (reviewed in Acevedo-Whitehouse and Cunningham 2006).

1.5. Seychelles Warbler: a model system for evolutionary questions

The Seychelles warbler (*Acrocephalus sechellensis*) is a long-lived, insectivorous, passerine endemic to the Seychelles Archipelago in the Indian Ocean. Originally found across the archipelago (Spurgin et al. 2014), the species underwent an extreme bottleneck in the 1960's primarily due to habitat loss, with only ca 29 individuals remaining on Cousin Island (4°20'S, 55°40'E; 0.29 km^2, Crook 1960). In 1968 Cousin Island was purchased by a consortium led by the International Council for Bird Preservation (now BirdLife International), and an intensive conservation and habitat restoration programme began. Following successful habitat restoration, the population reached carrying capacity on Cousin (with ca. 320 adult birds present in ca 115 territories (Komdeur 1992, Brouwer et al. 2009)) in the 1980s, To allow continued population growth and expand the range of the species, four additional populations have since been successfully established (Komdeur 1994, Richardson et al. 2006, Wright et al. 2014): 29 birds were translocated to Aride in 1988 (4°12′S, 55°40'E, 0.68 km^2, Komdeur 1994), 29 birds to Cousine in 1990 (4°21′S, 55°39′E, 0.25km^2, Komdeur 1994), 58 birds to Denis in 2004 (3°48′S, 55°40′E, 1.42 km^2, Richardson et al. 2006), and 59 birds to Frégate in 2011 (4°35′S, 55°56′E, 2.19 km^2, Wright et al. 2014). Translocated individuals were selected based on sex, age, body condition and breeding experience, but without reference to genetic characteristics (Wright et al. 2014). Of the translocated populations, Aride and Cousine are now at carrying capacity (ca 1,850 individuals, and ca 210 individuals respectively (Wright et al. 2014)), while the populations on Denis and Frégate are still increasing (ca 424 birds in 2015, Doblas and McClelland 2015, and ca 141 birds in 2016, Johnson et al. 2018, respectively).

Regular monitoring of the original Cousin Island population has been conducted since 1986, intensifying from 1997 onwards (Komdeur 1992, Hammers et al. 2019). Nearly all birds on the island have been caught, blood sampled, and fitted with a unique combination of three colour rings, and a metal British Trust for Ornithology ring for easy individual identification in the field (Richardson et al. 2001). Seychelles warblers have a maximum lifespan of 19 years, with a median post-fledgling lifespan of 5.5 years (Hammers & Brouwer, 2017). Mortality is greatest in the first year, but adult annual survival is high (84%, Brouwer et al., 2006). The extensive monitoring of this, and the neighbouring islands has shown that there is virtually no inter-island dispersal to, or from, Cousin (0.1%, Komdeur et al. 2004)). Furthermore, annual resighting of individuals is high (98%, (Brouwer et al. 2010)), allowing all individuals to be followed from birth till death, and thus enabling accurate lifetime fitness component estimates to be calculated (e.g., survival, status, and reproductive output – see data chapters for specific details)

(Brouwer et al. 2006). The long-term fitness data, combined with a detailed pedigree (Sparks et al. 2020), and the collection of blood samples make the Cousin population an ideal model system with which to answer a range of evolutionary ecology questions (for example; evolution of cooperative breeding (Komdeur 1996, Richardson et al. 2002), inbreeding (Richardson et al. 2004), mate choice (Richardson et al. 2005, Raj Pant et al. 2020), and senescence (Hammers et al. 2013, Hammers et al. 2015) among others).

1.5.1. Life history

Seychelles warblers are facultative cooperative breeders, with socially monogamous primary breeders defending strict territories year-round (Komdeur 1992). Breeding can occur year-round, but is dependent on food availability, thus there are two primary breeding seasons: the major season, June-September, and a minor season from November-March (Komdeur 1996, Komdeur and Daan 2005). Females typically lay single-egg clutches, although two-three egg clutches occur occasionally (Komdeur 1991, Richardson et al. 2001). Both male and female breeders participate in territory defence (Kingma et al. 2017) and provide offspring care (Komdeur 1994). Habitat saturation means that the availability of good quality breeding territories is extremely limited, consequently sexually mature birds sometimes delay independent breeding (Komdeur 1992), and instead become subordinates in natal or non-natal territories (Eikenaar et al. 2007, Kingma et al. 2016). Overall ca 50% of territories contain subordinates (Hammers et al. 2019). Furthermore, ca 30% of subordinate birds become helpers, feeding, or incubating young (Komdeur 1992, Komdeur 1994, Richardson et al. 2002). Although ca 40% of female subordinate helpers gain reproductive success within the territory through joint nesting, subordinate male helpers rarely gain paternity (Richardson et al. 2002, Raj Pant et al. 2019). Consequently, the benefits of cooperative breeding for helpers include both direct breeding and indirect kin selected benefits (Richardson et al. 2003).

Habitat saturation, along with social fidelity (Komdeur 1992, Richardson et al. 2007) and longevity (Hammers and Brouwer 2017) constrain social mate choice on Cousin (Richardson et al. 2005). However, high levels of extra-pair paternity (EPP) occur in this species, with 41% of offspring fathered by a dominant male from outside the natal territory (Richardson et al. 2001, Raj Pant et al. 2019), thus allowing the potential for genetic mate choice (see below 1.5.2). Social males attempt to prevent EPP in their own group by mate guarding their social female (Komdeur et al. 1999, Komdeur et al. 2007), and to gain EPP with other females via extra-territorial forays (Komdeur et al. 1999).

1.5.2. Genetic variation in the Seychelles warbler

The Seychelles warbler harbours reduced genetic variation as a result of a past genetic bottleneck (Spurgin et al. 2014), although some variation does still exist at neutral loci across the genome (Richardson et al 2000). Importantly elevated levels of variation appear to have been maintained at certain loci (Richardson and Westerdahl 2003, Gilroy et al. 2017). Key immune genes have already been characterised in the Seychelles warbler, including TLR genes (Gilroy et al. 2017), Avian β-defensins (Gilroy et al. 2016), MHC class I exon 3 (MHC-I) (Richardson and Westerdahl 2003, Hansson and Richardson 2005), and MHC class II exon 2 (MHC-II) (Hutchings 2009). These studies identified signatures of past positive selection at two TLR loci -*TLR3* and *TLR15,* in a sub-sample of individuals (Gilroy et al. 2017). Of these, one of the three SNPs present in *TLR3* appeared particularly important as it was non-synonymous, found within the functionally important leucine-rich repeat domain region, and had a relatively high minor allele frequency (32%, n = 28). Likewise, (Richardson and Westerdahl 2003) found variation still existed at the MHC-I locus, and evidence that this has been maintained by selection. However, the degree of functional variation at the MHC-II in the Seychelles warbler remains unresolved, though there is evidence of past selection (Hutchings 2009). Previous attempts to examine MHC-II, using cloning and reference strand-mediated conformation analysis (RSCA) proved to be inefficient and failed to resolve individual characteristics, though there was evidence of at least 10 functional alleles (Hutchings 2009).

As well as signatures of past selection, genetic variation has been linked to differences in individual fitness in the Seychelles warbler. For example, relatively low individual genome-wide homozygosity compared to the population mean (as the result of inbreeding) is negatively correlated to condition (Bebbington et al. 2016) and reproductive success (Brouwer et al. 2007). Likewise, both MHC-I diversity, and a specific MHC-I allele, *Ase-ua4*, are positively associated with survival (Brouwer et al. 2010). Sexual selection also seems to be acting on MHC-I variation, with females more likely to gain extra-pair fertilisations if their social mate has reduced MHC diversity. When this occurs, females are then more likely to gain fertilisations with males who have greater MHC diversity than their social mates (Richardson et al. 2005). This leads to offspring with greater MHC diversity and increased survival (Richardson et al. 2005, Brouwer et al. 2010). However, whether TLR variation is under contemporary selection, and what factor is causing this selective pressure at the MHC genes, remains unknown.

1.6. Thesis outline

In this thesis, I investigate how evolutionary forces shape immunogenetic variation in the Seychelles warbler (*Acrocephalus sechellensis)*. In Chapter 2, focusing on a functional *TLR3* locus. I assess temporal and spatial changes in allele frequencies in the Cousin population, and four other derived populations. I then test whether selection could be driving any change on in the Cousin population. In Chapter 3, I test a suite of questions to determine whether pre- or post- copulatory sexual selection, as well as natural selection, may act to shape *TLR3* variation. In Chapter 4, I investigate whether host-microbiome co-evolution may play a role in maintaining functional immunogenetic variation. I use next-generation sequencing to characterise MHC-I, MHC-II, and the gut microbiome. I then investigate the potential role these key immune genes play in determining characteristics of the GM. Lastly, in Chapter 5, I discuss the findings and broad implications resulting from Chapters 2 - 5 and suggest directions for future research.

1.7. References

Acevedo-Whitehouse, K. and Cunningham, A. A. (2006). Is MHC enough for understanding wildlife immunogenetics? *Trends in Ecology & Evolution.* **21**(8): 433-438.

Aderem, A. and Ulevitch, R. J. (2000). Toll-like receptors in the induction of the innate immune response. *Nature.* **406**(6797): 782-787.

Aguilar, A., Roemer, G., Debenham, S., Binns, M., Garcelon, D. and Wayne, R. K. (2004). High MHC diversity maintained by balancing selection in an otherwise genetically monomorphic mammal. *Proceedings of the National Academy of Sciences of the United States of America.* **101**(10): 3490-3494.

Akira, S., Uematsu, S. and Takeuchi, O. (2006). Pathogen Recognition and Innate Immunity. *Cell.* **124**(4): 783-801.

Alcaide, M. and Edwards, S. V. (2011). Molecular evolution of the toll-like receptor multigene family in birds. *Molecular Biology and Evolution.* **28**(5): 1703-1715.

Alcaide, M., Edwards, S. V., Negro, J. J., Serrano, D. and Tella, J. L. (2008). Extensive polymorphism and geographical variation at a positively selected MHC class II B gene of the lesser kestrel (*Falco naumanni*). *Molecular Ecology.* **17**(11): 2652-2665.

Altuvia, Y. and Margalit, H. (2004). A structure-based approach for prediction of MHC-binding peptides. *Methods.* **34**(4): 454-459.

Amato Katherine, R. (2013). Co-evolution in context: The importance of studying gut microbiomes in wild animals. *Microbiome Science and Medicine.* **1**(1).

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

Antonides, J., Mathur, S., Sundaram, M., Ricklefs, R. and DeWoody, A. J. (2019). Immunogenetic response of the bananaquit in the face of malarial parasites. *BMC Evolutionary Biology.* **19**(1): 107.

Apanius, V., Penn, D., Slev, P. R., Ruff, L. R. and Potts, W. K. (1997). The nature of selection on the major histocompatibility complex. *Critical Reviews™ in Immunology.* **17**(2).

Areal, H., Abrantes, J. and Esteves, P. J. (2011). Signatures of positive selection in Toll-like receptor (TLR) genes in mammals. *BMC Evolutionary Biology.* **11**(1): 368.

Armstrong, C., Davies, R. G., González-Quevedo, C., Dunne, M., Spurgin, L. G. and Richardson, D. S. (2019). Adaptive landscape genetics and malaria across divergent island bird populations. *Ecology and evolution.* **9**(22): 12482-12502.

Ayala, F. J. and Campbell, C. A. (1974). Frequency-dependent selection. *Annual review of ecology and systematics.* **5**(1): 115-138.

Bagheri, M. and Zahmatkesh, A. (2018). Evolution and species-specific conservation of toll-like receptors in terrestrial vertebrates. *International Reviews of Immunology.* **37**(5): 217-228.

Barton, G. M. (2007). Viral recognition by Toll-like receptors. *Seminars in Immunology.* **19**(1): 33-40.

Bateson, Z. W., Hammerly, S. C., Johnson, J. A., Morrow, M. E., Whittingham, L. A. and Dunn, P. O. (2016). Specific alleles at immune genes, rather than genome-wide heterozygosity, are related to immunity and survival in the critically endangered Attwater's prairie-chicken. *Molecular Ecology.* **25**(19): 4730-4744.

Bebbington, K., Spurgin, L. G., Fairfield, E. A., Dugdale, H. L., Komdeur, J., Burke, T. and Richardson, D. S. (2016). Telomere length reveals cumulative individual and transgenerational inbreeding effects in a passerine bird. *Molecular Ecology.* **25**(12): 2949-2960.

Benskin, C. M. H., Rhodes, G., Pickup, R. W., Mainwaring, M. C., Wilson, K. and Hartley, I. R. (2015). Life history correlates of fecal bacterial species richness in a wild population of the blue tit Cyanistes caeruleus. *Ecology and evolution.* **5**(4): 821-835.

Bernatchez, L. and Landry, C. (2003). MHC studies in nonmodel vertebrates: what have we learned about natural selection in 15 years? *Journal of Evolutionary Biology.* **16**(3): 363-377.

Biedrzycka, A., Sebastian, A., Migalska, M., Westerdahl, H. and Radwan, J. (2017). Testing genotyping strategies for ultra-deep sequencing of a co-amplifying gene

family: MHC class I in a passerine bird. *Molecular ecology resources.* **17**(4): 642-655.

Biedrzycka, A., Wojciech, B., Adam, Ć., Wojciech, S., Tadeusz, Z., Magdalena, M., . . . Jacek, R. (2018). Blood parasites shape extreme major histocompatibility complex diversity in a migratory passerine. *Molecular Ecology.* **27**(11): 2594-2603.

Birkhead, T. R. and Møller, A. (1998). Sperm competition. *New York: Academic.*

Boehm, T. (2012). Evolution of Vertebrate Immunity. *Current Biology.* **22**(17): R722-R732.

Bond, A. B. and Kamil, A. C. (2002). Visual predators select for crypticity and polymorphism in virtual prey. *Nature.* **415**(6872): 609-613.

Bonneaud, C., Balenger, S. L., Zhang, J., Edwards, S. V. and Hill, G. E. (2012). Innate immunity and the evolution of resistance to an emerging infectious disease in a wild bird. *Molecular Ecology.* **21**(11): 2628-2639.

Bonneaud, C., Chastel, O., Federici, P., Westerdahl, H. and Sorci, G. (2006a). Complex Mhc-based mate choice in a wild passerine. *Proceedings of the Royal Society B: Biological Sciences.* **273**(1590): 1111-1116.

Bonneaud, C., Pérez-Tris, J., Federici, P., Chastel, O., Sorci, G. and Koella, J. (2006b). Major histocompatibility alleles associated with local resistance to malaria in a passerine. *Evolution.* **60**(2): 383-389.

Bos, D. H., Williams, R. N., Gopurenko, D., Bulut, Z. and Dewoody, J. A. (2009). Condition-dependent mate choice and a reproductive disadvantage for MHC-divergent male tiger salamanders. *Molecular Ecology.* **18**(15): 3307-3315.

Bouzat, J. L. (2010). Conservation genetics of population bottlenecks: the role of chance, selection, and history. *Conservation Genetics.* **11**(2): 463-478.

Brandies, P. A., Grueber, C. E., Ivy, J. A., Hogg, C. J. and Belov, K. (2018). Disentangling the mechanisms of mate choice in a captive koala population. *PeerJ.* **6**: e5438.

Brouwer, L., Barr, I., Van De Pol, M., Burke, T., Komdeur, J. and Richardson, D. S. (2010). MHC-dependent survival in a wild population: evidence for hidden genetic benefits gained through extra-pair fertilizations. *Molecular Ecology.* **19**(16): 3444-3455.

Brouwer, L., Komdeur, J. and Richardson, D. S. (2007). Heterozygosity–fitness correlations in a bottlenecked island species: a case study on the Seychelles warbler. *Molecular Ecology.* **16**(15): 3134-3144.

Brouwer, L., Richardson, D. S., Eikenaar, C. and Komdeur, J. (2006). The role of group size and environmental factors on survival in a cooperatively breeding tropical passerine. *Journal of Animal Ecology.* **75**(6): 1321-1329.

Brouwer, L., Tinbergen, J. M., Both, C., Bristol, R., Richardson, D. S. and Komdeur, J. (2009). Experimental evidence for density-dependent reproduction in a cooperatively breeding passerine. *Ecology.* **90**(3): 729-741.

Brown, J. L. and Eklund, A. (1994). Kin Recognition and the Major Histocompatibility Complex: An Integrative Review. *The American Naturalist.* **143**(3): 435-461.

Brownlie, R. and Allan, B. (2011). Avian toll-like receptors. *Cell and Tissue Research.* **343**(1): 121-130.

Chapman, J. R., Hellgren, O., Helin, A. S., Kraus, R. H. S., Cromie, R. L. and Waldenström, J. (2016). The Evolution of Innate Immune Genes: Purifying and Balancing Selection on β-Defensins in Waterfowl. *Molecular Biology and Evolution.* **33**(12): 3075-3087.

Charbonnel, N. and Pemberton, J. (2005). A long-term genetic survey of an ungulate population reveals balancing selection acting on MHC through spatial and temporal fluctuations in selection. *Heredity.* **95**(5): 377-388.

Charlesworth, B. (2009). Effective population size and patterns of molecular evolution and variation. *Nature Reviews Genetics.* **10**(3): 195-205.

Charlesworth, B. and Charlesworth, D. (1999). The genetic basis of inbreeding depression. *Genetical Research.* **74**(3): 329-340.

Charlesworth, D. and Charlesworth, B. (1990). Inbreeding depression with heterozygote advantage and its effect on selection for modifiers changing the outcrossing rate. *Evolution.* **44**(4): 870-888.

Chen, N., Juric, I., Cosgrove, E. J., Bowman, R., Fitzpatrick, J. W., Schoech, S. J., . . . Coop, G. (2019). Allele frequency dynamics in a pedigreed natural population. *Proceedings of the National Academy of Sciences.* **116**(6): 2158-2164.

Chesser, R. K. (1991). Influence of gene flow and breeding tactics on gene diversity within populations. *Genetics.* **129**(2): 573-583.

Clayton, J. B., Vangay, P., Huang, H., Ward, T., Hillmann, B. M., Al-Ghalith, G. A., . . . Knights, D. (2016). Captivity humanizes the primate microbiome. *Proceedings of the National Academy of Sciences.* **113**(37): 10376-10381.

Crnokrak, P. and Roff, D. A. (1999). Inbreeding depression in the wild. *Heredity.* **83**(3): 260-270.

Crook, J. H. (1960). *The present status of certain rare land birds of the Seychelles islands.*

Croze, M., Živković, D., Stephan, W. and Hutter, S. (2016). Balancing selection on immunity genes: review of the current literature and new analysis in Drosophila melanogaster. *Zoology.* **119**(4): 322-329.

Darwin, C. (1859). On the origin of the species by natural selection.

Darwin, C. (1896). *The descent of man and selection in relation to sex.* D. Appleton.

Daszak, P., Cunningham, A. A. and Hyatt, A. D. (2000). Emerging infectious diseases of wildlife--threats to biodiversity and human health. *Science.* **287**(5452): 443-449.

Delves, P. J. and Roitt, I. M. (2000). The Immune System. *New England Journal of Medicine.* **343**(2): 108-117.

Ditchkoff, S. S., Lochmiller, R. L., Masters, R. E., Hoofer, S. R. and Bussche, R. A. V. D. (2001). Major-histocompatibility-complex-associated variation in secondary sexual traits of white-tailed deer (Odocoileus virginianus): evidence for good-genes advertisement. *Evolution.* **55**(3): 616-625.

Dlugosch, K. and Parker, I. (2008). Founding events in species invasions: genetic variation, adaptive evolution, and the role of multiple introductions. *Molecular Ecology.* **17**(1): 431-449.

Doblas, L. L. and McClelland, S. (2015). Seychelles warbler population census on Denis Island.

Dobzhansky, T. G. (1955). A review of some fundamental concepts and problems of population genetics.

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

Downing, T., Lloyd, A. T., O'Farrelly, C. and Bradley, D. G. (2010). The Differential Evolutionary Dynamics of Avian Cytokine and TLR Gene Classes. *The Journal of Immunology.* **184**(12): 6993-7000.

Dunn, P. O., Bollmer, J. L., Freeman-Gallant, C. R. and Whittingham, L. A. (2013). MHC variation is related to a sexually selected ornament, survival, and parasite resistance in common yellowthroats. *Evolution.* **67**(3): 679-687.

Eberhard, W. (1996). *Female control: sexual selection by cryptic female choice.* Princeton University Press.

Eikenaar, C., Richardson, D. S., Brouwer, L. and Komdeur, J. (2007). Parent presence, delayed dispersal, and territory acquisition in the Seychelles warbler. *Behavioral Ecology.* **18**(5): 874-879.

Eizaguirre, C., Lenz, T. L., Kalbe, M. and Milinski, M. (2012). Rapid and adaptive evolution of MHC genes under parasite selection in experimental vertebrate populations. *Nature Communications.* **3**: 621.

Eizaguirre, C., Yeates, S. E., Lenz, T. L., Kalbe, M. and Milinski, M. (2009). MHC-based mate choice combines good genes and maintenance of MHC polymorphism. *Molecular Ecology.* **18**(15): 3316-3329.

Ejsmond, M. J., Radwan, J. and Wilson, A. B. (2014). Sexual selection and the evolutionary dynamics of the major histocompatibility complex. *Proceedings of the Royal Society of London B: Biological Sciences.* **281**(1796): 20141662.

Ekblom, R., French, L., Slate, J. and Burke, T. (2010). Evolutionary Analysis and Expression Profiling of Zebra Finch Immune Genes. *Genome Biology and Evolution.* **2**: 781-790.

Evans, M. L., Dionne, M., Miller, K. M. and Bernatchez, L. (2012). Mate choice for major histocompatibility complex genetic divergence as a bet-hedging strategy in the Atlantic salmon (<i>Salmo salar</i>). *Proceedings of the Royal Society B: Biological Sciences.* **279**(1727): 379-386.

Fernandez, N., Cooper, J., Sprinks, M., AbdElrahman, M., Fiszer, D., Kurpisz, M. and Dealtry, G. (1999). A critical review of the role of the major histocompatibility complex in fertilization, preimplantation development and feto-maternal interactions. *Human Reproduction Update.* **5**(3): 234-248.

Fernando, M. M. A., Stevens, C. R., Walsh, E. C., De Jager, P. L., Goyette, P., Plenge, R. M., . . . Rioux, J. D. (2008). Defining the Role of the MHC in Autoimmunity: A Review and Pooled Analysis. *PLoS Genetics.* **4**(4): e1000024.

Ferrer-Admetlla, A., Bosch, E., Sikora, M., Marquès-Bonet, T., Ramírez-Soriano, A., Muntasell, A., . . . Casals, F. (2008). Balancing Selection Is the Main Force Shaping the Evolution of Innate Immunity Genes. *The Journal of Immunology.* **181**(2): 1315-1322.

Fisher, R. A. (1930). *The genetical theory of natural selection.* Oxford, Oxford University Press.

Fitzpatrick, M. J., Feder, E., Rowe, L. and Sokolowski, M. B. (2007). Maintaining a behaviour polymorphism by frequency-dependent selection on a single gene. *Nature.* **447**(7141): 210-212.

Flajnik, M. F. and Kasahara, M. (2001). Comparative Genomics of the MHC: Glimpses into the Evolution of the Adaptive Immune System. *Immunity.* **15**(3): 351-362.

Frankham, R. (1996). Relationship of Genetic Variation to Population Size in Wildlife. *Conservation Biology.* **10**(6): 1500-1508.

Freeland, W. (1976). Pathogens and the evolution of primate sociality. *Biotropica.* **8**(1): 12-24.

Gillingham, M. A. F., Richardson, D. S., Løvlie, H., Moynihan, A., Worley, K. and Pizzari, T. (2009). Cryptic preference for MHC-dissimilar females in male red junglefowl, Gallus gallus. *Proceedings of the Royal Society B: Biological Sciences.* **276**(1659): 1083-1092.

Gilroy, D., van Oosterhout, C., Komdeur, J. and Richardson, D. S. (2016). Avian β-defensin variation in bottlenecked populations: the Seychelles warbler and other congeners. *Conservation Genetics.* **17**(3): 661-674.

Gilroy, D., van Oosterhout, C., Komdeur, J. and Richardson, D. S. (2017). Toll-like receptor variation in the bottlenecked population of the endangered Seychelles warbler. *Animal Conservation.* **20**(3): 235-250.

Gonzalez-Quevedo, C., Spurgin, L. G., Illera, J. C. and Richardson, D. S. (2015). Drift, not selection, shapes toll-like receptor variation among oceanic island populations. *Molecular Ecology.* **24**(23): 5852-5863.

Greene, L. K., Williams, C. V., Junge, R. E., Mahefarisoa, K. L., Rajaonarivelo, T., Rakotondrainibe, H., . . . Drea, C. M. (2020). A role for gut microbiota in host niche differentiation. *The ISME Journal.* **14**(7): 1675-1687.

Grieves, L. A., Gloor, G. B., Bernards, M. A. and MacDougall-Shackleton, E. A. (2019). Songbirds show odour-based discrimination of similarity and diversity at the major histocompatibility complex. *Animal Behaviour.* **158**: 131-138.

Grogan, K. E., Harris, R. L., Boulet, M. and Drea, C. M. (2019). Genetic variation at MHC class II loci influences both olfactory signals and scent discrimination in ring-tailed lemurs. *BMC Evolutionary Biology.* **19**(1): 171.

Grueber, C. E., Knafler, G. J., King, T. M., Senior, A. M., Grosser, S., Robertson, B., . . . Jamieson, I. G. (2015). Toll-like receptor diversity in 10 threatened bird species: relationship with microsatellite heterozygosity. *Conservation Genetics.* **16**(3): 595-611.

Grueber, C. E., Wallis, G. P. and Jamieson, I. G. (2013). Genetic drift outweighs natural selection at toll-like receptor (TLR) immunity loci in a re-introduced population of a threatened species. *Molecular Ecology.* **22**(17): 4470-4482.

Grueber, C. E., Wallis, G. P. and Jamieson, I. G. (2014). Episodic Positive Selection in the Evolution of Avian Toll-Like Receptor Innate Immunity Genes. *PLoS ONE.* **9**(3): e89632.

Grueber, C. E., Wallis, G. P., King, T. M. and Jamieson, I. G. (2012). Variation at Innate Immunity Toll-Like Receptor Genes in a Bottlenecked Population of a New Zealand Robin. *PLoS ONE.* **7**(9): e45011.

Gruen, J. R. and Weissman, S. M. (1997). Evolving Views of the Major Histocompatibility Complex. *Blood.* **90**(11): 4252-4265.

Haldane, J. B. S. (1992). Disease and evolution. *Current Science.* **63**(9): 599-604.

Hale, K. A. and Briskie, J. V. (2007). Decreased immunocompetence in a severely bottlenecked population of an endemic New Zealand bird. *Animal Conservation.* **10**(1): 2-10.

Hamilton, W. D. and Zuk, M. (1982). Heritable true fitness and bright birds: a role for parasites? *Science.* **218**(4570): 384-387.

Hammers, M. and Brouwer, L. (2017). Rescue behaviour in a social bird: removal of sticky ‘bird-catcher tree’seeds by group members. *Behaviour.* **154**(4): 403-411.

Hammers, M., Kingma, S. A., Bebbington, K., van de Crommenacker, J., Spurgin, L. G., Richardson, D. S., . . . Komdeur, J. (2015). Senescence in the wild: Insights from a long-term study on Seychelles warblers. *Experimental Gerontology.* **71**: 69-79.

Hammers, M., Kingma, S. A., Spurgin, L. G., Bebbington, K., Dugdale, H. L., Burke, T., . . . Richardson, D. S. (2019). Breeders that receive help age more slowly in a cooperatively breeding bird. *Nature Communications.* **10**(1): 1301.

Hammers, M., Richardson, D. S., Burke, T. and Komdeur, J. (2013). The impact of reproductive investment and early-life environmental conditions on senescence: support for the disposable soma hypothesis. *Journal of Evolutionary Biology.* **26**(9): 1999-2007.

Hansson, B. and Richardson, D. S. (2005). Genetic variation in two endangered Acrocephalus species compared to a widespread congener: estimates based on functional and random loci. Animal Conservation forum, Cambridge University Press.

Hartmann, S. A., Schaefer, H. M. and Segelbacher, G. (2014). Genetic depletion at adaptive but not neutral loci in an endangered bird species. *Molecular Ecology.* **23**(23): 5712-5725.

Hawn, T. R., Verbon, A., Lettinga, K. D., Zhao, L. P., Li, S. S., Laws, R. J., . . . Nachman, A. (2003). A common dominant TLR5 stop codon polymorphism abolishes flagellin signaling and is associated with susceptibility to legionnaires' disease. *Journal of Experimental Medicine.* **198**(10): 1563-1572.

He, K., Minias, P. and Dunn, P. O. (2020). Long-Read Genome Assemblies Reveal Extraordinary Variation in the Number and Structure of MHC Loci in Birds. *Genome Biology and Evolution.* **13**(2).

Hedrick, P. W. (1994). Evolutionary genetics of the major histocompatibility complex. *The American Naturalist.* **143**(6): 945-964.

Hedrick, P. W. (1998). Balancing selection and MHC. *Genetica.* **104**(3): 207-214.

Hedrick, P. W. (2002). Pathogen resistance and genetic variation at MHC loci. *Evolution.* **56**(10): 1902-1908.

Hill, A. V. S., Allsopp, C. E. M., Kwiatkowski, D., Anstey, N. M., Twumasi, P., Rowe, P. A., . . . Greenwood, B. M. (1991). Common West African HLA antigens are associated with protection from severe malaria. *Nature.* **352**(6336): 595-600.

Hird, S. M. (2017). Evolutionary Biology Needs Wild Microbiomes. *Frontiers in Microbiology.* **8**(725).

Hooper, L. V., Midtvedt, T. and Gordon, J. I. (2002). How host-microbial interactions shape the nutrient environment of the mammalian intestine. *Annu Rev Nutr.* **22**: 283-307.

Hughes, A. L. and Nei, M. (1988). Pattern of nucleotide substitution at major histocompatibility complex class I loci reveals overdominant selection. *Nature.* **335**(6186): 167-170.

Hughes, A. L. and Yeager, M. (1998). Natural selection at major histocompatibility complex loci of vertebrates. *Annual review of genetics.* **32**(1): 415-435.

Hutchings, K. (2009). *Parasite-mediated selection in an island endemic, the Seychelles warbler (Acrocephalus sechellensis)*. University of East Anglia.

Iwasaki, A. and Medzhitov, R. (2010). Regulation of adaptive immunity by the innate immune system. *Science.* **327**(5963): 291-295.

Jäger, I., Eizaguirre, C., Griffiths, S. W., Kalbe, M., Krobbach, C., Reusch, T., . . . Milinski, M. (2007). Individual MHC class I and MHC class IIB diversities are associated with male and female reproductive traits in the three-spined stickleback. *Journal of Evolutionary Biology.* **20**(5): 2005-2015.

Jamieson, I. G. (2015). Significance of population genetics for managing small natural and reintroduced populations in New Zealand. *New Zealand Journal of Ecology.* **39**(1): 1.

Johnson, T. F., Brown, T. J., Richardson, D. S. and Dugdale, H. L. (2018). The importance of post-translocation monitoring of habitat use and population growth: insights from a Seychelles Warbler (Acrocephalus sechellensis) translocation. *Journal of Ornithology.* **159**(2): 439-446.

Johnston, S. E., Gratten, J., Berenos, C., Pilkington, J. G., Clutton-Brock, T. H., Pemberton, J. M. and Slate, J. (2013). Life history trade-offs at a single locus maintain sexually selected genetic variation. *Nature.* **502**(7469): 93-95.

Kalbe, M., Eizaguirre, C., Dankert, I., Reusch, T. B. H., Sommerfeld, R. D., Wegner, K. M. and Milinski, M. (2009). Lifetime reproductive success is maximized with optimal major histocompatibility complex diversity. *Proceedings of the Royal Society B: Biological Sciences.* **276**(1658): 925-934.

Kamiya, T., O'Dwyer, K., Westerdahl, H., Senior, A. and Nakagawa, S. (2014). A quantitative review of MHC-based mating preference: the role of diversity and dissimilarity. *Molecular Ecology.* **23**(21): 5151-5163.

Keller, L. F. and Waller, D. M. (2002). Inbreeding effects in wild populations. *Trends in Ecology & Evolution.* **17**(5): 230-241.

Kempenaers, B. (2007). Mate Choice and Genetic Quality: A Review of the Heterozygosity Theory. Advances in the Study of Behavior, Academic Press. **37:** 189-278.

Khan, I., Maldonado, E., Silva, L., Almeida, D., Johnson, W. E., O'Brien, S. J., . . . Antunes, A. (2019). The vertebrate TLR supergene family evolved dynamically

by gene gain/loss and positive selection revealing a host–pathogen arms race in birds. *Diversity.* **11**(8): 131.

Kimura, M. (1968). Evolutionary Rate at the Molecular Level. *Nature.* **217**(5129): 624-626.

Kingma, S. A., Bebbington, K., Hammers, M., Richardson, D. S. and Komdeur, J. (2016). Delayed dispersal and the costs and benefits of different routes to independent breeding in a cooperatively breeding bird. *Evolution.* **70**(11): 2595-2610.

Kingma, S. A., Komdeur, J., Burke, T. and Richardson, D. S. (2017). Differential dispersal costs and sex-biased dispersal distance in a cooperatively breeding bird. *Behavioral Ecology.* **28**(4): 1113-1121.

Klein, J. (1986). *Natural history of the major histocompatibility complex.* Wiley.

Kloch, A., Wenzel, M. A., Laetsch, D. R., Michalski, O., Bajer, A., Behnke, J. M., . . . Piertney, S. B. (2018). Signatures of balancing selection in toll-like receptor (TLRs) genes – novel insights from a free-living rodent. *Scientific Reports.* **8**(1): 8361.

Kolbe, J. J., Leal, M., Schoener, T. W., Spiller, D. A. and Losos, J. B. (2012). Founder effects persist despite adaptive differentiation: a field experiment with lizards. *Science.* **335**(6072): 1086-1089.

Kolodny, O., Callahan, B. J. and Douglas, A. E. (2020). The role of the microbiome in host evolution. *Philosophical Transactions of the Royal Society B: Biological Sciences.* **375**(1808): 20190588.

Komdeur, J. (1992). Importance of habitat saturation and territory quality for evolution of cooperative breeding in the Seychelles warbler. *Nature.* **358**(6386): 493-495.

Komdeur, J. (1994). Conserving the Seychelles warbler Acrocephalus sechellensis by translocation from Cousin Island to the islands of Aride and Cousine. *Biological Conservation.* **67**(2): 143-152.

Komdeur, J. (1994). Experimental evidence for helping and hindering by previous offspring in the cooperative-breeding Seychelles warbler Acrocephalus sechellensis. *Behavioral Ecology and Sociobiology.* **34**(3): 175-186.

Komdeur, J. (1996). Influence of helping and breeding experience on reproductive performance in the Seychelles warbler: a translocation experiment. *Behavioral Ecology.* **7**(3): 326-333.

Komdeur, J. (1996). Seasonal timing of reproduction in a tropical bird, the Seychelles warbler: a field experiment using translocation. *Journal of Biological Rhythms.* **11**(4): 333-346.

Komdeur, J., Burke, T. and Richardson, D. S. (2007). Explicit experimental evidence for the effectiveness of proximity as mate-guarding behaviour in reducing extra-pair fertilization in the Seychelles warbler. *Molecular Ecology.* **16**(17): 3679-3688.

Komdeur, J. and Daan, S. (2005). Breeding in the monsoon: semi-annual reproduction in the Seychelles warbler (Acrocephalus sechellensis). *Journal of Ornithology.* **146**(4): 305-313.

Komdeur, J., Kraaijeveld-Smit, F., Kraaijeveld, K. and Edelaar, P. (1999). Explicit experimental evidence for the role of mate guarding in minimizing loss of paternity in the Seychelles warbler. *Proceedings of the Royal Society of London. Series B: Biological Sciences.* **266**(1433): 2075-2081.

Komdeur, J., Piersma, T., Kraaijeveld, K., Kraaijeveld-Smit, F. and Richardson, D. S. (2004). Why Seychelles Warblers fail to recolonize nearby islands: unwilling or unable to fly there? *Ibis.* **146**(2): 298-302.

Koskella, B. and Bergelson, J. (2020). The study of host-microbiome (co)evolution across levels of selection. *Philosophical Transactions of the Royal Society B: Biological Sciences.* **375**(1808): 20190604.

Kurilshikov, A., Wijmenga, C., Fu, J. and Zhernakova, A. (2017). Host Genetics and Gut Microbiome: Challenges and Perspectives. *Trends in Immunology.* **38**(9): 633-647.

Lacy, R. C. (1987). Loss of genetic diversity from managed populations: interacting effects of drift, mutation, immigration, selection, and population subdivision. *Conservation Biology.* **1**(2): 143-158.

Lande, R. (1976). Natural Selection and Random Genetic Drift in Phenotypic Evolution. *Evolution.* **30**(2): 314-334.

Lande, R. (1988). Genetics and demography in biological conservation. *Science.* **241**(4872): 1455-1460.

Lande, R. (1995). Mutation and Conservation. *Conservation Biology.* **9**(4): 782-791.

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

Landry, C., Garant, D., Duchesne, P. and Bernatchez, L. (2001b). 'Good genes as heterozygosity': the major histocompatibility complex and mate choice in Atlantic salmon (Salmo salar). *Proceedings of the Royal Society of London. Series B: Biological Sciences.* **268**(1473): 1279-1285.

Lara, C. E., Grueber, C. E., Holtmann, B., Santos, E. S. A., Johnson, S. L., Robertson, B. C., . . . Nakagawa, S. (2020). Assessment of the dunnocks' introduction to New Zealand using innate immune-gene diversity. *Evolutionary Ecology.* **34**(5): 803-820.

Lawlor, D. A., Zemmour, J., Ennis, P. D. and Parham, P. (1990). Evolution of Class-I MHC Genes and Proteins: From Natural Selection to Thymic Selection. *Annual Review of Immunology.* **8**(1): 23-63.

Liu, G., Zhang, H., Zhao, C. and Zhang, H. (2019). Evolutionary History of the Toll-Like Receptor Gene Family across Vertebrates. *Genome Biology and Evolution.* **12**(1): 3615-3634.

Loehle, C. (1995). Social Barriers to Pathogen Transmission in Wild Animal Populations. *Ecology.* **76**(2): 326-335.

Loiseau, C., Zoorob, R., Garnier, S., Birard, J., Federici, P., Julliard, R. and Sorci, G. (2008). Antagonistic effects of a Mhc class I allele on malaria-infected house sparrows. *Ecology Letters.* **11**(3): 258-265.

Løvlie, H., Gillingham, M. A., Worley, K., Pizzari, T. and Richardson, D. S. (2013). Cryptic female choice favours sperm from major histocompatibility complex-dissimilar males. *Proceedings of the Royal Society of London B: Biological Sciences.* **280**(1769): 20131296.

Lukasch, B., Westerdahl, H., Strandh, M., Knauer, F., Winkler, H., Moodley, Y. and Hoi, H. (2017). Major histocompatibility complex genes partly explain early survival in house sparrows. *Scientific Reports.* **7**(1): 6571.

Lynch, M., Conery, J. and Burger, R. (1995). Mutation accumulation and the extinction of small populations. *American Naturalist.* **146**(4): 489-518.

Maizels, R. M. and Nussey, D. H. (2013). Into the wild: digging at immunology's evolutionary roots. *Nature Immunology.* **14**(9): 879-883.

Martinsohn, J. T., Sousa, A. B., Guethlein, L. A. and Howard, J. C. (1999). The gene conversion hypothesis of MHC evolution: a review. *Immunogenetics.* **50**(3-4): 168-200.

Maurice, C. F., Cl Knowles, S., Ladau, J., Pollard, K. S., Fenton, A., Pedersen, A. B. and Turnbaugh, P. J. (2015). Marked seasonal variation in the wild mouse gut microbiota. *The ISME Journal.* **9**(11): 2423-2434.

Mayr, E. (1954). Change of genetic environment and evolution.

Mays Jr, H. L. and Hill, G. E. (2004). Choosing mates: good genes versus genes that are a good fit. *Trends in Ecology & Evolution.* **19**(10): 554-559.

McFall-Ngai, M., Hadfield, M. G., Bosch, T. C. G., Carey, H. V., Domazet-Lošo, T., Douglas, A. E., . . . Wernegreen, J. J. (2013). Animals in a bacterial world, a new imperative for the life sciences. *Proceedings of the National Academy of Sciences.* **110**(9): 3229-3236.

Medzhitov, R. (2001). Toll-like receptors and innate immunity. *Nature Reviews Immunology.* **1**(2): 135-145.

Milinski, M. (2006). The Major Histocompatibility Complex, Sexual Selection, and Mate Choice. *Annual Review of Ecology, Evolution, and Systematics.* **37**(1): 159-186.

Miller, H. C. and Lambert, D. M. (2004). Genetic drift outweighs balancing selection in shaping post-bottleneck major histocompatibility complex variation in New Zealand robins (Petroicidae). *Molecular Ecology.* **13**(12): 3709-3721.

Moeller, A. H. and Sanders, J. G. (2020). Roles of the gut microbiota in the adaptive evolution of mammalian species. *Philosophical Transactions of the Royal Society B: Biological Sciences.* **375**(1808): 20190597.

Moore, S. L. and Wilson, K. (2002). Parasites as a Viability Cost of Sexual Selection in Natural Populations of Mammals. *Science.* **297**(5589): 2015-2018.

Mukherjee, S., Sarkar-Roy, N., Wagener, D. K. and Majumder, P. P. (2009). Signatures of natural selection are not uniform across genes of innate immune system, but purifying selection is the dominant signature. *Proceedings of the National Academy of Sciences.* **106**(17): 7073-7078.

Neff, B. D. and Pitcher, T. E. (2005). Genetic quality and sexual selection: an integrated framework for good genes and compatible genes. *Molecular Ecology.* **14**(1): 19-38.

Nei, M., Maruyama, T. and Chakraborty, R. (1975). The bottleneck effect and genetic variability in populations. *Evolution.* 1-10.

Nelson-Flower, M. J., Germain, R. R., MacDougall-Shackleton, E. A., Taylor, S. S. and Arcese, P. (2018). Purifying selection in the Toll-like receptors of song sparrows Melospiza melodia. *Journal of Heredity.* **109**(5): 501-509.

Netea, M. G., Wijmenga, C. and O'Neill, L. A. J. (2012). Genetic variation in Toll-like receptors and disease susceptibility. *Nature Immunolgy.* **13**(6): 535-542.

Nowak, M. A., Tarczy-Hornoch, K. and Austyn, J. M. (1992). The optimal number of major histocompatibility complex molecules in an individual. *Proc Natl Acad Sci U S A.* **89**(22): 10896-10899.

O'Connor, E. A., Westerdahl, H., Burri, R. and Edwards, S. V. (2019). Avian MHC Evolution in the Era of Genomics: Phase 1.0. *Cells.* **8**(10): 1152.

Oliver, M., Telfer, S. and Piertney, S. (2009). Major histocompatibility complex (MHC) heterozygote superiority to natural multi-parasite infections in the water vole (Arvicola terrestris). *Proceedings of the Royal Society of London B: Biological Sciences.* **276**(1659): 1119-1128.

Paterson, S. and Pemberton, J. M. (1997). No evidence for major histocompatibility complex–dependent mating patterns in a free–living ruminant population. *Proceedings of the Royal Society of London. Series B: Biological Sciences.* **264**(1389): 1813-1819.

Penn, D. J., Damjanovich, K. and Potts, W. K. (2002). MHC heterozygosity confers a selective advantage against multiple-strain infections. *Proceedings of the National Academy of Sciences.* **99**(17): 11260-11264.

Penn, D. J. and Potts, W. K. (1999). The evolution of mating preferences and major histocompatibility complex genes. *The American Naturalist.* **153**(2): 145-164.

Phillips, K. P., Cable, J., Mohammed, R. S., Herdegen-Radwan, M., Raubic, J., Przesmycka, K. J., . . . Radwan, J. (2018). Immunogenetic novelty confers a selective advantage in host–pathogen coevolution. *Proceedings of the National Academy of Sciences.* **115**(7): 1552-1557.

Pickard, J. M., Zeng, M. Y., Caruso, R. and Núñez, G. (2017). Gut microbiota: Role in pathogen colonization, immune responses, and inflammatory disease. *Immunol Rev.* **279**(1): 70-89.

Piertney, S. B. and Oliver, M. K. (2005). The evolutionary ecology of the major histocompatibility complex. *Heredity.* **96**(1): 7-21.

Poltorak, A., Smirnova, I., He, X., Liu, M.-Y., Van Huffel, C., Birdwell, D., . . . Thompson, P. (1998). Genetic and physical mapping of the*Lps*Locus: identification of the Toll-4 receptor as a candidate gene in the critical region. *Blood Cells, Molecules, and Diseases.* **24**(3): 340-355.

Potts, W. K. and Wakeland, E. K. (1990). Evolution of diversity at the major histocompatibility complex. *Trends in Ecology & Evolution.* **5**(6): 181-187.

Quemere, E., Galan, M., Cosson, J. F., Klein, F., Aulagnier, S., Gilot-Fromont, E., . . . Charbonnel, N. (2015). Immunogenetic heterogeneity in a widespread ungulate: the European roe deer (*Capreolus capreolus*). *Molecular Ecology.* **24**(15): 3873-3887.

Raj Pant, S., Hammers, M., Komdeur, J., Burke, T., Dugdale, H. L. and Richardson, D. S. (2020). Age-dependent changes in infidelity in Seychelles warblers. *Molecular Ecology.* **29**(19): 3731-3746.

Raj Pant, S., Komdeur, J., Burke, T. A., Dugdale, H. L. and Richardson, D. S. (2019). Socio-ecological conditions and female infidelity in the Seychelles warbler. *Behavioral Ecology.* **30**(5): 1254-1264.

Raulo, A., Allen, B. E., Troitsky, T., Husby, A., Firth, J. A., Coulson, T. and Knowles, S. C. L. (2021). Social networks strongly predict the gut microbiota of wild mice. *The ISME Journal.*

Reed, D. H. and Frankham, R. (2003). Correlation between fitness and genetic diversity. *Conservation Biology.* **17**(1): 230-237.

Reusch, T. B. H., Haberli, M. A., Aeschlimann, P. B. and Milinski, M. (2001). Female sticklebacks count alleles in a strategy of sexual selection explaining MHC polymorphism. *Nature.* **414**(6861): 300-302.

Richardson, D., Jury, F., Blaakmeer, K., Komdeur, J. and Burke, T. (2001). Parentage assignment and extra-group paternity in a cooperative breeder: The Seychelles warbler (Acrocephalus sechellensis). *Molecular Ecology.* **10**(9): 2263-2273.

Richardson, D. and Westerdahl, H. (2003). MHC diversity in two Acrocephalus species: the outbred Great reed warbler and the inbred Seychelles warbler. *Molecular Ecology.* **12**(12): 3523-3529.

Richardson, D. S., Bristol, R. and Shah, N. J. (2006). Translocation of the Seychelles warbler Acrocephalus sechellensis to establish a new population on Denis Island, Seychelles. *Conservation Evidence.* **3**: 54-57.

Richardson, D. S., Burke, T. and Komdeur, J. (2003). Sex-specific associative learning cues and inclusive fitness benefits in the Seychelles warbler. *Journal of Evolutionary Biology.* **16**(5): 854-861.

Richardson, D. S., Burke, T. and Komdeur, J. (2007). Grandparent helpers: the adaptive significance of older, postdominant helpers in the Seychelles warbler. *Evolution.* **61**(12): 2790-2800.

Richardson, D. S., Burke, T., Komdeur, J. and Dunn, P. (2002). Direct benefits and the evolution of female-biased cooperative breeding in Seychelles warblers. *Evolution.* **56**(11): 2313-2321.

Richardson, D. S., Komdeur, J. and Burke, T. (2004). Inbreeding in the Seychelles warbler: environment-dependent maternal effects. *Evolution.* **58**(9): 2037-2048.

Richardson, D. S., Komdeur, J., Burke, T. and von Schantz, T. (2005). MHC-based patterns of social and extra-pair mate choice in the Seychelles warbler. *Proceedings of the Royal Society B: Biological Sciences.* **272**(1564): 759-767.

Roach, J. C., Glusman, G., Rowen, L., Kaur, A., Purcell, M. K., Smith, K. D., . . . Aderem, A. (2005). The evolution of vertebrate Toll-like receptors. *Proceedings of the National Academy of Sciences of the United States of America.* **102**(27): 9577-9582.

Round, J. L. and Mazmanian, S. K. (2009). The gut microbiota shapes intestinal immune responses during health and disease. *Nat Rev Immunol.* **9**(5): 313-323.

Rowe, M., Veerus, L., Trosvik, P., Buckling, A. and Pizzari, T. (2020). The Reproductive Microbiome: An Emerging Driver of Sexual Selection, Sexual Conflict, Mating Systems, and Reproductive Isolation. *Trends in Ecology & Evolution.* **35**(3): 220-234.

Schad, J., Sommer, S. and Ganzhorn, J. U. (2004). MHC Variability of a Small Lemur in the Littoral Forest Fragments of Southeastern Madagascar. *Conservation Genetics.* **5**(3): 299-309.

Schröder, N. W. and Schumann, R. R. (2005). Single nucleotide polymorphisms of Toll-like receptors and susceptibility to infectious disease. *The Lancet infectious diseases.* **5**(3): 156-164.

Setchell, J. M., Charpentier, M. J., Abbott, K. M., Wickings, E. J. and Knapp, L. A. (2009). Is brightest best? Testing the Hamilton-Zuk hypothesis in mandrills. *International Journal of Primatology.* **30**(6): 825-844.

Setchell, J. M., Charpentier, M. J. E., Abbott, K. M., Wickings, E. J. and Knapp, L. A. (2010). Opposites attract: MHC-associated mate choice in a polygynous primate. *Journal of Evolutionary Biology.* **23**(1): 136-148.

Shuker, D. M. (2010). Sexual selection: endless forms or tangled bank? *Animal Behaviour.* **79**(3): e11-e17.

Sin, Y. W., Annavi, G., Dugdale, H. L., Newman, C., Burke, T. and MacDonald, D. W. (2014). Pathogen burden, co-infection and major histocompatibility complex variability in the European badger (Meles meles). *Molecular Ecology.* **23**(20): 5072-5088.

Sin, Y. W., Annavi, G., Newman, C., Buesching, C., Burke, T., Macdonald, D. W. and Dugdale, H. L. (2015). MHC class II-assortative mate choice in European badgers (Meles meles). *Molecular Ecology.* **24**(12): 3138-3150.

Slade, J. W. G., Watson, M. J. and MacDougall-Shackleton, E. A. (2019). Balancing balancing selection? Assortative mating at the major histocompatibility complex despite molecular signatures of balancing selection. *Ecology and evolution.* **9**(9): 5146-5157.

Slade, R. and McCallum, H. (1992). Overdominant vs. frequency-dependent selection at MHC loci. *Genetics.* **132**(3): 861.

Slatkin, M. (1985). Gene flow in natural populations. *Annual review of ecology and systematics.* **16**(1): 393-430.

Slatkin, M. (1987). Gene flow and the geographic structure of natural populations. *Science.* **236**(4803): 787-792.

Smith, K. D., Andersen-Nissen, E., Hayashi, F., Strobe, K., Bergman, M. A., Barrett, S. L. R., . . . Aderem, A. (2003). Toll-like receptor 5 recognizes a conserved site on flagellin required for protofilament formation and bacterial motility. *Nature Immunology.* **4**(12): 1247-1253.

Sommer, S. (2005). The importance of immune gene variability (MHC) in evolutionary ecology and conservation. *Frontiers in zoology.* **2**(1): 1.

Sparks, A. M., Spurgin, L. G., van der Velde, M., Fairfield, E. A., Komdeur, J., Burke, T., . . . Dugdale, H. (2020). Telomere heritability and parental age at conception effects in a wild avian population.

Spor, A., Koren, O. and Ley, R. (2011). Unravelling the effects of the environment and host genotype on the gut microbiome. *Nature Reviews Microbiology.* **9**(4): 279-290.

Spurgin, L. G. and Richardson, D. S. (2010). How pathogens drive genetic diversity: MHC, mechanisms and misunderstandings. *Proceedings of the Royal Society B: Biological Sciences.* **277**(1684): 979-988.

Spurgin, L. G., Wright, D. J., van der Velde, M., Collar, N. J., Komdeur, J., Burke, T. and Richardson, D. S. (2014). Museum DNA reveals the demographic history of the endangered Seychelles warbler. *Evolutionary Applications.* **7**(9): 1134-1143.

Stoffel, M. A., Acevedo-Whitehouse, K., Morales-Durán, N., Grosser, S., Chakarov, N., Krüger, O., . . . Hoffman, J. I. (2020). Early sexual dimorphism in the developing gut microbiome of northern elephant seals. *Molecular Ecology.* **29**(11): 2109-2122.

Sutton, J. T., Nakagawa, S., Robertson, B. C. and Jamieson, I. G. (2011). Disentangling the roles of natural selection and genetic drift in shaping variation at MHC immunity genes. *Molecular Ecology.* **20**(21): 4408-4420.

Suzuki, T. A. (2017). Links between Natural Variation in the Microbiome and Host Fitness in Wild Mammals. *Integrative and Comparative Biology.* **57**(4): 756-769.

Suzuki, T. A., Phifer-Rixey, M., Mack, K. L., Sheehan, M. J., Lin, D., Bi, K. and Nachman, M. W. (2019). Host genetic determinants of the gut microbiota of wild mice. *Molecular Ecology.* **28**(13): 3197-3207.

Takahata, N. and Nei, M. (1990). Allelic genealogy under overdominant and frequency-dependent selection and polymorphism of major histocompatibility complex loci. *Genetics.* **124**(4): 967-978.

Takeuchi, O. and Akira, S. (2010). Pattern recognition receptors and inflammation. *Cell.* **140**(6): 805-820.

Tellier, A. and Brown, J. K. (2011). Spatial heterogeneity, frequency-dependent selection and polymorphism in host-parasite interactions. *BMC Evolutionary Biology.* **11**(1): 1-15.

Thursz, M. R., Kwiatkowski, D., Allsopp, C. E., Greenwood, B. M., Thomas, H. C. and Hill, A. V. (1995). Association between an MHC class II allele and clearance of hepatitis B virus in the Gambia. *New England Journal of Medicine.* **332**(16): 1065-1069.

Tregenza, T. and Wedell, N. (2000). Genetic compatibility, mate choice and patterns of parentage: invited review. *Molecular Ecology.* **9**(8): 1013-1027.

Velová, H., Gutowska-Ding, M. W., Burt, D. W. and Vinkler, M. (2018). Toll-like receptor evolution in birds: gene duplication, pseudogenisation and diversifying selection. *Molecular Biology and Evolution.* **35**(9): 2170-2184.

Videvall, E., Song, S. J., Bensch, H. M., Strandh, M., Engelbrecht, A., Serfontein, N., . . . Cornwallis, C. K. (2018). The development of gut microbiota in ostriches and its association with juvenile growth. *bioRxiv.*

Villarino, N. F., LeCleir, G. R., Denny, J. E., Dearth, S. P., Harding, C. L., Sloan, S. S., . . . Schmidt, N. W. (2016). Composition of the gut microbiota modulates the severity of malaria. *Proceedings of the National Academy of Sciences.* **113**(8): 2235-2240.

Villaseñor-Cardoso, M. I. and Ortega, E. (2011). Polymorphisms of innate immunity receptors in infection by parasites. *Parasite Immunology.* **33**(12): 643-653.

Von Schantz, T., Wittzell, H., Goransson, G., Grahn, M. and Persson, K. (1996). MHC genotype and male ornamentation: genetic evidence for the Hamilton-Zuk model. *Proceedings of the Royal Society of London B: Biological Sciences.* **263**(1368): 265-271.

Walling, C. A., Nussey, D. H., Morris, A., Clutton-Brock, T. H., Kruuk, L. E. B. and Pemberton, J. M. (2011). Inbreeding depression in red deer calves. *BMC Evolutionary Biology.* **11**(1): 318.

Werling, D., Jann, O. C., Offord, V., Glass, E. J. and Coffey, T. J. (2009). Variation matters: TLR structure and species-specific pathogen recognition. *Trends in Immunology.* **30**(3): 124-130.

Westerdahl, H. (2004). No evidence of an MHC-based female mating preference in great reed warblers. *Molecular Ecology.* **13**(8): 2465-2470.

Westerdahl, H., Waldenström, J., Hansson, B., Hasselquist, D., von Schantz, T. and Bensch, S. (2005). Associations between malaria and MHC genes in a migratory songbird. *Proceedings of the Royal Society B: Biological Sciences.* **272**(1571): 1511-1518.

Woolhouse, M. E. J., Webster, J. P., Domingo, E., Charlesworth, B. and Levin, B. R. (2002). Biological and biomedical implications of the co-evolution of pathogens and their hosts. *Nature Genetics.* **32**(4): 569-577.

Wright, D. J., Shah, N. J. and Richardson, D. S. (2014). Translocation of the Seychelles warbler Acrocephalus sechellensis to establish a new population on Frégate Island, Seychelles. *Conserv Evid.* **11**: 20-24.

Wright, D. J., Spurgin, L. G., Collar, N. J., Komdeur, J., Burke, T. and Richardson, D. S. (2014). The impact of translocations on neutral and functional genetic diversity within and among populations of the Seychelles warbler. *Molecular Ecology.* **23**(9): 2165-2177.

Wright, S. (1931). Evolution in Mendelian populations. *Genetics.* **16**(2): 97-159.

Yu, L., Nie, Y., Yan, L., Hu, Y. and Wei, F. (2018). No evidence for MHC-based mate choice in wild giant pandas. *Ecology and evolution.* **8**(17): 8642-8651.

Chapter 2

Contemporary evolution of the viral-sensing *TLR3* gene in an isolated vertebrate population

Life finds a way – MC 1997

2.1. Abstract

Understanding where genetic variation exists, and how it influences fitness within populations is important from an evolutionary and conservation perspective. Signatures of past selection suggest that pathogen-mediated balancing selection is a key driver of immunogenetic variation, but studies tracking contemporary evolution are needed to help resolve the evolutionary forces and mechanism at play. Previous work in a bottlenecked population of Seychelles warblers (*Acrocephalus sechellensis*) show that functional variation has been maintained at the viral-sensing Toll-like receptor 3 (*TLR3)* gene, including one non-synonymous SNP, resulting in two alleles. Here, we characterise evolution at this *TLR3* locus over a 25-year period within the original remnant population of the Seychelles warbler, and in four other derived, populations. Results show a significant and consistent temporal decline in the frequency of the *TLR3*^{C} allele in the original population, and that similar declines in the *TLR3*^{C} allele frequency occurred in all the derived populations. Individuals (of both sexes) with the *TLR3*^{CC} genotype had lower survival, and males - but not females - that carry the *TLR3*^{C} allele had significantly lower lifetime reproductive success than those with only the *TLR3*^{A} allele. These results indicate that positive selection, caused by an as yet unknown agent, is driving *TLR3* evolution in the Seychelles warblers. No evidence of heterozygote advantage was detected. However, whether the positive selection observed is part of a longer-term pattern of balancing selection (through fluctuating selection or rare-allele advantage) cannot be resolved without tracking the *TLR3*^{C} allele in the populations over an extended period of time.

2.2. Introduction

Genetic variation is key to both the fitness of individuals and the persistence of populations (Reed and Frankham 2003). Loss of genetic variation can result in inbreeding depression, and loss of heterozygote advantage in individuals, and a reduction in the adaptive potential of the population, all of which may be especially detrimental in small or bottlenecked populations (Lande 1995). Therefore, understanding the factors and mechanisms that shape genetic variation within such populations is important from both an evolutionary and conservation perspective (Frankham 1996).

Various interacting evolutionary forces act to shape genetic variation within populations, either through 'neutral' processes such as genetic drift, or 'adaptive' processes such as selection (Wright 1931, Lande 1976). Determining the relative importance of these forces in shaping genetic diversity is key to understanding the adaptive potential of populations (Lacy 1987, Sutton et al. 2011). In small populations, genetic drift is usually predominant, resulting in a decrease in genetic variation across the genome (Robinson et al. 2016). Nevertheless, selection can also act on functional genes, either counteracting or reinforcing the effect of drift. Directional or purifying selection can push alleles to fixation, resulting in a reduction in genetic variation and reinforcing drift (Mukherjee et al. 2009). In contrast, balancing selection (caused by a suite of potential mechanisms) may maintain genetic variation and counteract the effect of drift (Hedrick 1998).

Pathogens can have considerable negative impact on the survival and reproductive success of individuals (Daszak et al. 2000) and are strong drivers of evolutionary change in natural populations (Haldane 1992). Consequently, immunogenetic loci - i.e., those involved in the detection and combating of pathogens – are excellent candidates in which to investigate the evolutionary forces underlying the maintenance of genetic variation (Sommer 2005, Croze et al. 2016). Indeed, pathogen-mediated selection is thought to be a key driver of balancing selection (Spurgin and Richardson 2010). Three non-mutually exclusive mechanisms driving pathogen-mediated selection have been proposed: heterozygote advantage (Doherty and Zinkernagel 1975), rare allele advantage (Slade and McCallum 1992), and fluctuating selection (Hill et al. 1991). These three mechanisms – along with other forces such as sexual selection – can act independently, in concert, or in trade-off with one other (Apanius et al. 1997, Spurgin and Richardson 2010, Ejsmond et al. 2014).

Immunogenetic research on wild populations has focused mainly on receptor genes of the acquired immune system: in particular the exceptionally polymorphic major

histocompatibility complex (MHC) (reviewed in Piertney and Oliver 2005). However, high levels of diversity (Hedrick 1994), gene duplication (Bollmer et al. 2010), conversion, recombination (Miller and Lambert 2004), and epistasis (van Oosterhout 2009) makes it hard to tease apart the evolutionary forces driving MHC variation (Spurgin and Richardson 2010). In contrast, the genes involved in the innate immune response, while still often polymorphic, exhibit relatively lower complexity. Furthermore, the innate immune system is the host's first line of response to pathogens enabling a broad defence against an assortment of organisms (Aderem and Ulevitch 2000). Consequently, innate immune genes can be more tractable candidates with which to study the evolutionary forces shaping immunogenetic variation in wild populations (Acevedo-Whitehouse and Cunningham 2006).

Toll-Like Receptor (TLR) genes encode receptor molecules which bind to pathogen-associated molecular patterns - evolutionary conserved molecular motifs that are integral to the pathogen's survival (Medzhitov 2001). Once bound, the TLR molecule triggers a cascade of processes associated with the innate and adaptive immune responses (Akira et al. 2006). Vertebrate TLRs can be divided into six families, depending on the pathogen-associated molecular patterns they detect (Roach et al. 2005). For example, *TLR3* binds to viral dsRNA (Barton 2007), while *TLR5* binds to bacterial flagellin (Brownlie and Allan 2011). While the majority of the TLR structure is structurally conserved (Roach et al. 2005), there is variation in the leucine-rich repeat domain of TLR genes, resulting in functional variation at the binding site. Such TLR polymorphisms have been associated with resistance (Antonides et al. 2019), or susceptibility to specific pathogens (Kloch et al. 2018), or associated with increased survival (Grueber et al. 2013, Bateson et al. 2016). TLRs can evolve rapidly as a result of pathogen-mediated selection (Downing et al. 2010) and evidence of balancing selection at TLR genes has been reported for various taxa (e.g., Areal et al. 2011, Velová et al. 2018). Nevertheless, most of these studies only inferred past selection from sequence variation and could not determine if selection was still acting, or determine the specific mechanisms involved. Moreover, in various bottlenecked populations, genetic drift may override selection as the dominant evolutionary force shaping TLR variation (Grueber et al. 2013, Gonzalez-Quevedo et al. 2015).

Here, we investigate the contemporary evolution of TLR variation in a natural population of Seychelles warblers (*Acrocephalus sechellensis*). The last remaining population of this species on Cousin island underwent a bottleneck in the 1900s resulting in decreased genome-wide genetic variation (Spurgin et al. 2014). Extensive longitudinal monitoring and a lack of dispersal (Komdeur et al. 2004) means that virtually all individual warblers

on Cousin island are sampled, marked and tracked throughout their entire lives (Komdeur, 1992, Hammers et al. 2015). This allows for accurate measures of survival and reproductive success (Hammers et al. 2019). As part of a conservation programme, individuals have been translocated from Cousin to establish populations on four additional islands (Komdeur 1994, Richardson et al., 2006, Wright et al. 2014), allowing spatial TLR variation to be investigated. A previous study found that five of seven TLR loci examined in the contemporary population (2000-2008) of Seychelles warbler on Cousin Island were polymorphic and detected a signature of past positive selection at two loci, one of these being *TLR3* - a viral sensing TLR (Gilroy et al. 2017). One of the three SNPs at this *TLR3* locus was singled out for investigation because it is non-synonymous, found within the functionally important leucine-rich repeat domain region, and had a relatively high minor allele frequency (32%, *n* = 28). However, if and how balancing selection maintains variation at this locus has yet to be investigated.

We first assess how the frequency of this *TLR3* SNP has changed over 25-years in the Seychelles warbler on Cousin Island. We then test the role of selection in shaping *TLR3* variation in this population; specifically, if survival and reproductive success are associated with individual *TLR3* genotypes. Lastly, we compare patterns of *TLR3* evolution over time in, and between, the Cousin population and the newly established (translocated) populations. These analyses allow us to better understand which evolutionary forces shape immunogenetic variation in small populations of conservation concern.

2.3. Methods

2.3.1. Study species and system

The Seychelles warbler is a small (ca 15 g) insectivorous passerine endemic to the Seychelles. The species was distributed across the archipelago prior to human colonisation (Spurgin et al. 2014), but underwent a severe population reduction in the 1900s due to anthropogenic effects, with just ca 29 individuals remaining on Cousin Island (4°20’S, 55°40’E; 0.29 km^2) by the 1960s (Crook 1960). After intensive conservation, the population recovered to carrying capacity on Cousin (ca 320 adults present in ca 110 territories) by the 1980s (Brouwer et al. 2009, Komdeur 1992). Additional populations were established by translocations to four nearby islands (Table S2.1): Aride (29 birds in 1988), Cousine (29 birds in 1990), Denis (58 birds in 2004), and Frégate (59 birds in 2011) (Komdeur 1994, Richardson et al. 2006, Wright et al. 2014).

Founder individuals (all from Cousin) were selected based on sex, age, body condition, and breeding experience but without reference to genetic characteristics (Wright et al. 2014). Translocations to Aride and Cousine were undertaken before blood sampling became routine, whereas sampling of all the founders of the Denis and Frégate populations was undertaken (Wright et al. 2014). Of the translocated populations, two are now at carrying capacity (Aride: ca 1,850 individuals; Cousine: ca 210 individuals (Wright et al. 2014)), while the populations on the other islands are still increasing (Denis: ca 424 birds in 2015 (Doblas and McClelland 2015); Frégate: ca 141 birds in 2016 (Johnson et al. 2018)).

The Seychelles warbler on Cousin island has been monitored since 1986 (Komdeur 1992, Hammers et al. 2019). A comprehensive population census has taken place every year during the major breeding season (June–September), and – since 1997 – also during the minor breeding season (November–March) except in 2000–2002 and in 2006 (Brouwer et al. 2010). Individuals were recorded as present if caught, or observed, during the field season. The other populations have not been censused regularly and only sporadic census data are available.

The rate of annual resighting of individuals on Cousin is high (0.98, Brouwer et al. 2010) and there is virtually no inter-island dispersal (0.1%, Komdeur et al. 2004), thus enabling accurate survival estimates (Brouwer et al. 2006). Individuals can be confidently presumed dead if not seen for two consecutive breeding seasons; the date of death is assigned as the end of the last season in which a bird was observed (Hammers et al. 2013). Ages were rounded to the nearest 0.5 years. Adult annual survival is high (84%), with mortality being greatest in first-year birds (Brouwer et al. 2006). Median lifespan is 5.5 years post-fledging, and maximum lifespan is 19 years (Hammers and Brouwer 2017).

Females typically lay single-egg clutches (Richardson et al. 2001) and only occasionally two or three eggs (Komdeur 1991). They are facultatively cooperative breeders, with a socially monogamous dominant breeder pair defending strict territories year-round (Komdeur 1992). Some adult birds delay independent breeding and become subordinates (Kingma et al. 2016), and may help raise offspring (Komdeur 1992, Hammers et al. 2019). Although 44% of female subordinates gain reproductive success by co-breeding, male subordinates rarely gain paternity (Richardson et al. 2002, Raj Pant et al. 2019). Extra-pair paternity is frequent in this species (Richardson et al. 2001), with 41% of offspring fathered outside the natal territory (Raj Pant et al. 2019).

Individuals are caught either by mist-net, or as nestlings, and are aged based on hatch date, behaviour, and eye colour at first catch (for details see Komdeur 1992; Wright, 2014). Each bird is given a metal British Trust for Ornithology (BTO) ring and a unique combination of three colour rings (Richardson et al. 2001). Routine blood sampling began in 1993. Since 1997, >96% of the Cousin population has been ringed and blood sampled (Raj Pant et al. 2019). Samples (ca 25 µl) are collected by brachial venipuncture and stored in 0.8 ml of absolute ethanol at 4°C.

2.3.2. Molecular methods

Genomic DNA was extracted from blood using either a salt extraction technique (Richardson et al. 2001) or, since 2013, the DNeasy blood and tissue kit (Qiagen, Crawley, UK). Sex was determined via PCR (Griffiths et al. 1998). Individuals were genotyped at 30 polymorphic microsatellite loci (Richardson et al. 2001). Parentage assignment was carried out using MasterBayes 2.52 (Hadfield et al. 2006); for full details see Sparks et al. (2020). Parentage assignment was conducted for 1,966 offspring that hatched between 1993–2018, with 89% of fathers and 86% of mothers assigned at ≥80% accuracy. Standardised individual and maternal microsatellite heterozygosity (H_s) was calculated using the R package Genhet 3.1 (Coulon 2010). Two of the microsatellite loci were excluded from this heterozygosity analysis due to pooled alleles (see Sparks et al. 2020). Variation at exon 3 of the MHC class I loci had previously been screened in individuals from Cousin (1,148 individuals hatched between 1992–2009) (Richardson and Westerdahl 2003, Wright 2014).

Variation within the leucine-rich repeat domain of the *TLR3* exon 4 had previously been characterised; of the three SNPs found only one SNP was non-synonymous and had a minor allele frequency of >0.05 (Gilroy et al. 2017). This focal SNP is found at 198 bp in the Seychelles warbler *TLR3* reference sequence (NCBI accession number: KM657704.2), where the presence of an A or C nucleotide caused a change of amino acid from Lysine (+ charge) to Asparagine (polar). Variation at KM657704.2:g.198A>C (hereafter referred to as *TLR3* SNP) was genotyped in 1,647 individuals using the KASP genotyping technology by LGC Genomics, Hertfordshire.

2.3.3. Statistical analyses

Unless otherwise stated, all analyses were conducted in R 3.6.1.

2.3.3.1. Temporal patterns of TLR3 *variation on Cousin*

In total, 1,190 birds hatched on Cousin from four cohorts 1992–94, 1997–99, 2005–10, and 2016–18, were sequenced at the *TLR3* SNP. The earliest and latest of the sampled cohorts were used to assess temporal changes. In addition, the years 1997–99 and 2005–10 were selected; (i) to avoid hatch years in which translocations happened (2004, 2011), as the subsequent reduction in population density may have a positive effect on juvenile (<1 year) survival in that year (Brouwer et al. 2006), and, (ii) to focus on individuals with the most complete MHC and life-history data. Temporal allelic variation was analysed using a linear model (LM) and significance was assessed using the F-statistic. Frequency of *TLR3*^{C} in the sampled adult or juvenile population was the response variable, while year was the fixed factor.

2.3.3.2. Contemporary selection on TLR3 *variation on Cousin*

Survival: A mixed-effects Cox proportional hazards model in the package coxme 2.2-14 (Therneau 2019), was used to determine whether *TLR3* genotypes differed in survival. Model diagnostics using Schoenfeld's residuals confirmed that proportional hazards assumptions were met (Grambsch and Therneau 1994). Age at death was standardised to bi-annual levels corresponding to the major and minor seasons. Fieldwork was not conducted for four minor breeding seasons (2000–2002, 2006), so accurate bi-annual survival estimates could not be calculated for 77 individuals. Instead, the minimum date of death was assigned (i.e., the last season an individual was observed). Excluding these individuals did not qualitatively alter the results, so they were retained in the model. Birds first caught as an adult (>1 year, n = 21) were excluded to prevent any survivorship bias from including individuals that have already survived the first year of life, and because Seychelles warblers cannot be reliably aged past one year of age (Wright 2014). Individuals that were translocated to other islands (n = 39), and those still alive after the major 2018 breeding season (n = 42) were right-censored. Previous work has found that in low-quality seasons maternal heterozygosity affected offspring survival (Brouwer et al. 2007), and MHC diversity positively affected survival in juveniles, while individuals with the MHC class I allele (*Ase-ua4*) have a greater life expectancy (Brouwer et al. 2010). Due to these fitness component differences, and the fact that MHC-I has similar properties to TLR3 in that it primarily binds intracellular peptides, we also include MHC-I characteristics in subsequent analysis.*TLR3* genotype (*TLR3*^{AA}/*TLR3*^{AC}/*TLR3*^{CC}), MHC diversity (2–8 different alleles), presence of the *Ase-ua4* allele (Yes/No), individual heterozygosity (H_s), maternal heterozygosity (Maternal H_s), sex (Male/Female) and season in which born (Minor/Major) were included as fixed factors in the model, with hatch year included as a random factor. Individuals hatched on Cousin between 1997–99 or 2005–2010, for which these data were available, were included (n = 517). Cox

proportional hazards models in the package survival 2.44-1.1 (Therneau and Lumley 2015), without the random effects, were used to plot Kaplan–Meier survival curves.

Reproductive success: A zero-inflated generalised linear mixed model (GLMM) with a Poisson error structure was run using the package glmmTMB 0.2.3 (Brooks et al. 2017) to test whether lifetime reproductive success (LRS) was associated with *TLR3* variation. LRS was measured as the number of offspring that survived to independence (3 months) throughout an individual's lifespan. Both social and extra-pair offspring were included. Individuals that were translocated, or still alive after the minor 2018 season, were excluded due to incomplete data. Individuals first caught over one year of age, for which we did not have accurate age and longevity data, were also excluded. All other birds hatched on Cousin between 1997–99 and 2005–2010 were included (n = 487). *TLR3* genotype, MHC diversity, presence of the *Ase-ua4* allele, and individual H_s were fixed factors in the model, with year of hatch as a random factor to control for cohort effects. The sexes were modelled separately as it is likely that different factors and constraints act upon males and females.

As LRS is strongly correlated with longevity (GLMM, P<0.001, Table 2.2), and survival was strongly correlated with *TLR3* genotype (COXME, P = 0.026, Fig 2.2, Table 2.1), we tested if lifetime reproductive rate (defined as reproduction controlling for longevity) was associated with *TLR3* genotype. The model and dataset used was the same as used for LRS, except for two key differences: (i) Individuals which died before reaching adulthood (i.e., 1 year of age) were excluded from this analysis (resulting in n = 323), (ii) Age at death (i.e., longevity and longevity^2) were included as fixed factors. The inclusion of longevity, and the exclusion of non-adult individuals, allows reproductive success to be isolated from survival; thus, gaining a measure of the rate of reproduction during the individual's adult life.

For both LRS and rate of reproduction models all continuous factors were standardised (scaled and centred) using the package arm 1.10-1 (Gelman et al. 2018). Collinearity between fixed effects was tested using variance inflation factors. We used the package DHARMA 0.2.4 (Hartig 2017) to confirm that there was no over or under dispersion, residual spatial or temporal autocorrelation in the GLMM models. We used model averaging using the dredge function in the MUMIn package 1.43.6 (Barton and Barton 2019) to select plausible models. All models within 7 AICc of the top model were included in the averaged model, to get the final conditional model.

Selection coefficient: Mean values of LRS were calculated for each genotype from the raw data, relative fitness per *TLR3* genotype was calculated by dividing the mean for all three genotypes by the mean from the genotype with the greatest fitness. The dataset used was the same as that used for LRS – except that mean LRS was measured as the total number of offspring produced by an individual that survived to recruitment (>1 year) as this is a more accurate measure of genotype contribution to the next generation.

Hardy-Weinberg Equilibrium in young birds on Cousin: Deviation from Hardy-Weinberg Equilibrium (HWE) was tested using exact tests (Guo and Thompson 1992) based on allelic frequencies in Genepop 4.2 (Rousset 2008). *P* values were estimated with Markov chain algorithms (1,000 dememorizations, 100 batches, 1,000 iterations), and F_{IS} values are presented using Robertson and Hill estimates (Robertson and Hill 1984). First, all birds from Cousin first caught before 3 months of age (before independence) were tested ($n = 591$). Second, to determine if early-life mortality changed HWE proportions, this test was repeated including only individuals that survived until adulthood ($n = 361$). To determine if any deviation from HWE was caused by a temporal Wahlund-like effect (as in Pusack et al. 2014) we also re-ran the analysis separately for each hatch year.

2.3.3.3. Spatial and temporal TLR3 *variation across islands*

The earliest available samples from the source population, Cousin (120 birds caught in 1993 and 1994), were used to provide a proxy estimate of the initial *TLR3* diversity on Aride and Cousine (which were established in 1988 and 1990, i.e., before sampling took place). Samples from 56 of the 58 birds translocated to Denis, and all 59 birds translocated to Frégate were used to determine initial *TLR3* diversity on these islands. The most recent population samples were of 58 individuals caught in 2018 on Frégate, 158 individuals caught in 2015 on Denis, 54 individuals caught in 2012 and 2016 on Aride, 72 individuals caught in 2019 on Cousine, and 196 individuals caught in 2018 on Cousin (Table S2.1).

Genepop 4.2 (Rousset 2008) was used to test if the different island populations conformed to HWE (as above). We tested for temporal and spatial divergence in *TLR3* frequencies among populations using genic differentiation tests (Raymond and Rousset 1995) in Genepop 4.2 (Rousset 2008). Fisher's exact test and the Markov chain algorithm parameters were as above. First, we tested for differentiation between the initial (translocated or 1993–94 samples) and most recent samples from each population. Second, we tested for differentiation among populations using the most recent samples.

2.4. Results

In total, 1,608 out of 1,647 (0.98) samples were genotyped successfully at one *TLR3* SNP: 756/1608 (0.47) of these individuals had genotype $TLR3^{AA}$, 659/1608 (0.41) had $TLR3^{AC}$, and 193/1608 (0.12) had $TLR3^{CC}$.

2.4.1. Temporal patterns of TLR3 *variation on Cousin*

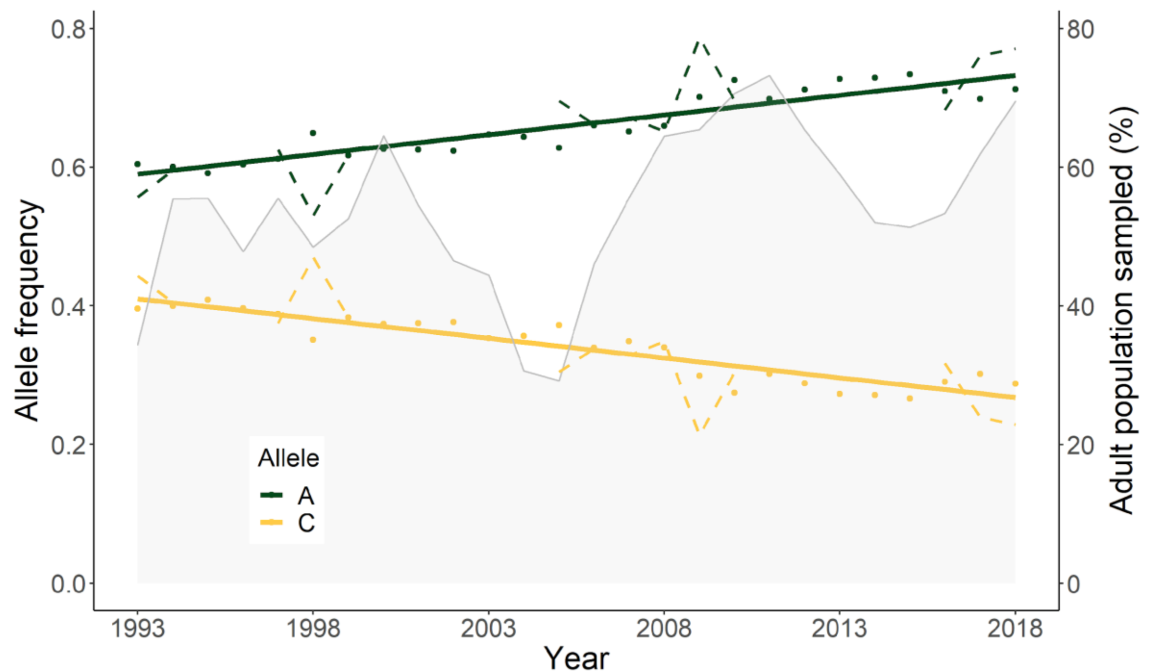


Fig 2.1: Allele frequency change at a nonsynonymous *TLR3* SNP in the Cousin population of the Seychelles warbler over 25 years (1993 - 2018). Points refer to *TLR3* allele frequencies in the adult population in a given year, the $TLR3^{A}$ allele in dark green, the $TLR3^{C}$ allele in yellow. Solid lines show linear regressions for the adult population. Dashed lines indicate frequencies in sampled individuals hatched in each year. The shaded grey area (right hand axis) shows the percentage of the adult population (mean: 310 individuals) screened in each year.

In the adult population on Cousin, the frequency of the minor $TLR3^{C}$ allele decreased significantly over time from 0.40 in 1993 to 0.29 in 2018, with a corresponding increase in the $TLR3^{A}$ allele (LM: $R^2 = 0.85$, $F_{1,24} = 140$, P <0.001, Fig 2.1). Likewise, the minor $TLR3^{C}$ allele also significantly decreased over time in the juvenile population from 0.44 in 1993 to 0.23 in 2018 (LM: $R^2 = 0.68$, $F_{1,12} = 28.7$, P <0.001, Fig 2.1).

2.4.2. Testing for contemporary selection on TLR3 *variation on Cousin*

There were significant differences in lifetime survival probabilities between *TLR3* genotypes. Individuals (first caught as juveniles) with the $TLR3^{CC}$ genotype had a 37% increased mortality risk compared to those with the $TLR3^{AC}$ or $TLR3^{AA}$ genotypes, with a median age of death of 1, 2, and 2.5 years respectively (COXME, *P* = 0.024, Fig 2.2, Table 2.1). Thus, individuals with at least one copy of the $TLR3^{A}$ allele had increased survival than those without (*P* = 0.025, Table S2.2). Independently – and as found previously in a smaller dataset (Brouwer et al. 2010) – individuals with the *Ase-ua4* MHC class I allele had a 25% lower risk of mortality than those without, corresponding to a median age of death at 3.5 years (compared to 2 years for those individuals without) (COXME, *P* = 0.028, Table 2.1). There was no significant effect of sex, H_s, maternal H_s, or MHC diversity on lifetime survival probability (Table 2.1), or of the season in which an individual hatched, although individuals hatched in the minor breeding season tended to have increased survival (COXME, *P* = 0.062, Table 2.1).

Table 2.1: Time-dependent Cox Regression modelling to test the effects of *TLR3* genotype on bi-annual survival in the Seychelles warbler population (*n* = 517) on Cousin. Coef = hazard rate; SE (coef) = standard error of the hazard rate; HR = hazard ratio. An HR >1 indicates increased hazard of mortality, and <1 indicates decreased hazard of mortality. Coefficient estimates are in reference to *TLR3* = ^{AA}, *Ase-ua4* = Present, Season born = Major, Sex = Female. Significant terms are in bold and underlined

Factor	coef	SE (coef)	HR	z	*P*
TLR3: ^{AC}	-0.01	0.10	0.99	-0.08	0.940
<u>*TLR3*: ^{CC}</u>	**<u>0.32</u>**	**<u>0.14</u>**	**<u>1.37</u>**	**<u>2.25</u>**	**<u>0.024</u>**
Individual H_s	-0.12	0.23	0.89	-0.52	0.600
<u>*Ase-ua4*</u>	**<u>-0.29</u>**	**<u>0.13</u>**	**<u>0.75</u>**	**<u>-2.20</u>**	**<u>0.028</u>**
MHC Diversity	-0.02	0.03	0.98	-0.77	0.440
Maternal H_s	-0.08	0.22	0.92	-0.37	0.710
Season born	-0.22	0.12	0.80	-1.86	0.062
Sex	-0.02	0.10	0.98	-0.19	0.850
Random effects	Variance	517 individuals			
Hatch year	0.015	9 hatch years			

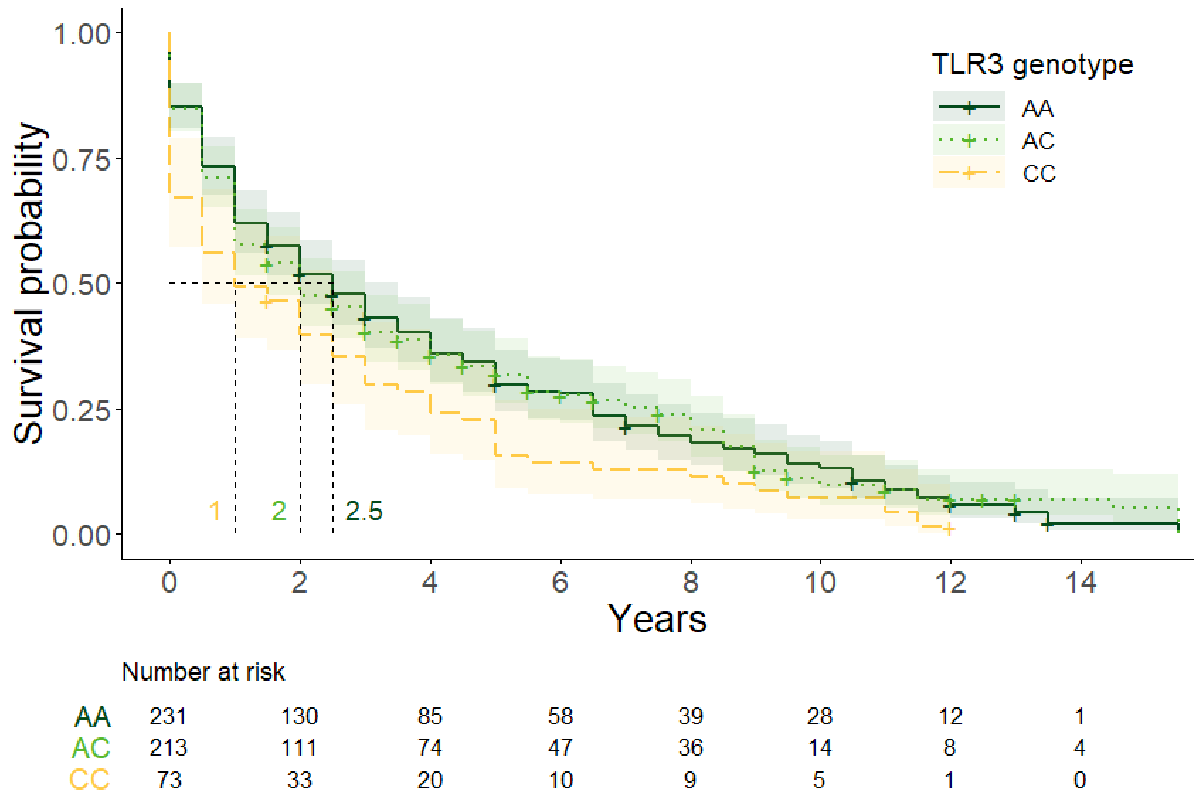


Fig 2.2: Effect of *TLR3* genotype on survival in the Seychelles warbler population on Cousin (n = 517). Lifetime survival probabilities classified into 6-month periods are shown for individuals with *TLR3*^{AA} (dark green, solid), *TLR3*^{AC} (light green, dotted) and *TLR3*^{CC} (yellow, dashed) genotypes. Shaded areas denote 95% confidence limits. Dotted vertical lines indicate median lifespan (in years) of each genotype. Translocated individuals and individuals still alive at the end of the study are right censored (indicated with the symbol '+').

In males, individuals with different *TLR3* genotypes had significantly different LRS. Males with *TLR3*^{AA} had greater LRS than those with *TLR3*^{AC} (P <0.001, Table 2.2, Fig 2.3a) or *TLR3*^{CC} (P = 0.003, Table 2.2, Fig 2.3a), with *TLR3*^{AA} males producing on average twice the number of independent offspring (mean ± SEM: 1.40 ± 0.27) than either *TLR3*^{AC} (mean ± SEM: 0.63 ± 0.17), or *TLR3*^{CC} males (mean ± SEM: 0.70 ± 0.21) over their lifetime. There was no significant difference in LRS between *TLR3*^{AC} and *TLR3*^{CC} genotypes (P = 0.86) in males. Thus, males with at least one copy of the *TLR3*^{C} allele had reduced LRS than those without (P <0.001, Table S2.3). In contrast in females there was no association between *TLR3* genotype and LRS (Fig 2.3a). In males, LRS decreased with increasing MHC diversity (P = 0.047, Table 2.2), whereas in females LRS tended to increase with increasing MHC diversity, although this result was marginally non-significant (P = 0.064, Table 2.2). H_s and the presence of *Ase-ua4* did not predict LRS for either sex (Table 2.2).

As survival was strongly correlated with *TLR3* genotype, we also investigated whether *TLR3* genotypes predicted reproductive rate after controlling for parental survival – i.e., by including longevity and controlling for breeding ability (survival to recruitment into the adult population). In both sexes, individuals who lived longer (greater longevity) produced significantly more offspring (GLMM, Age P <0.001, Table 2.2). There was also evidence for a negative quadratic effect of longevity in both sexes (GLMM, Age^2 P <0.001, Table 2.2). Males of *TLR3*^{AA} genotype tended to produce more offspring (surviving >3 months; GLMM, P = 0.049, Table 2.2, Fig 2.3b) than those of *TLR3*^{AC} genotype, while *TLR3*^{AA} and *TLR3*^{AC} genotypes did not differ from *TLR3*^{CC} genotypes (P = 0.38 and 0.54, respectively). There was no association between the rate of reproduction and *TLR3* genotype or quadratic age in females. H_s, MHC diversity, and the presence of *Ase-ua4* did not predict reproductive rate for either sex (Table 2.2).

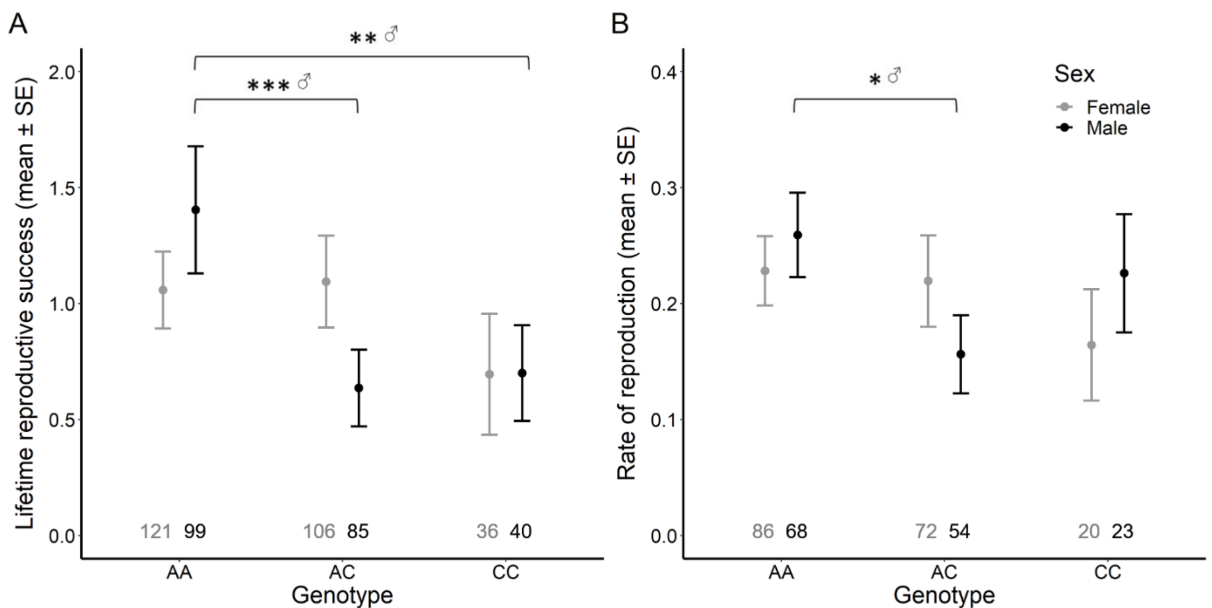


Fig 2.3: Effects of *TLR3* genotype on reproductive success in the Cousin population of the Seychelles warbler: **A**) Lifetime reproductive success (offspring surviving >3 months) for all birds; n = 487), **B**) Rate of reproduction (i.e., offspring surviving to >3 months/ longevity for focal birds that survived to adulthood; n = 323). Data are raw means and standard errors, with female data shown in light grey and males in black separated by genotype, with associated sample sizes at the bottom. *** P <0.001, ** P <0.01, * P <0.05.

The difference in LRS associated with *TLR3* variation equated to a selection coefficient of 0.34 against *TLR3*^{AC,} and 0.46 against *TLR3*^{CC} genotypes of both sexes, over ca 3 overlapping generations (assuming a generation time of 4 years (Spurgin et al. 2014)), when the selection coefficient of *TLR3*^{AA} genotype was set as 1.

Table 2: Reproductive success in male and female Seychelles warblers in relation to *TLR3* genotype: **A**) Lifetime reproductive success for all birds, **B**) Reproductive success controlling for longevity for birds that survived to adulthood. Zero-inflated GLMMs were used to generate conditional model-averaged values for all predictors featuring in the top model set ($\Delta AIC_c \leq 7$). Model-averaged estimates (β), their standard error (SE), adjusted SE, z value, *P* value, and relative importance (ω) are shown for all predictors featuring in the top model set ($\Delta AIC_c \leq 7$). Estimates are in reference to *TLR3* = ^{AA}, *Ase-ua4* = Present. *** $P < 0.001$, ** $P < 0.01$, * $P < 0.05$. Significant terms are in bold and underlined.

Response	Factor	**Male (A:** n=224**; B:** n=145**)**							**Female (A:** n=263**; B:** n=178**)**						
		ω	β	SE	Adj. SE	z	P		ω	β	SE	Adj. SE	z	p	
A) LRS - Count of offspring surviving > 3 months (independence)	Intercept		**1.10**	**0.21**	**0.21**	**5.16**	**<0.001**	***		**0.95**	**0.15**	**0.15**	**6.38**	**<0.001**	***
	zero-inflated intercept		**0.49**	**0.16**	**0.16**	**2.96**	**0.003**	**		**0.53**	**0.14**	**0.14**	**3.68**	**<0.001**	***
	TLR3: ^{AC}	**1.00**	**-0.69**	**0.19**	**0.19**	**3.63**	**<0.001**	***	0.15	0.06	0.15	0.15	0.40	0.693	
	TLR3: ^{CC}		**-0.74**	**0.25**	**0.25**	**2.97**	**0.003**	**		-0.16	0.26	0.26	0.62	0.536	
	Individual Hs	0.26	0.04	0.17	0.17	0.23	0.815		0.58	-0.24	0.14	0.14	1.64	0.101	
	MHC Diversity	**0.71**	**-0.29**	**0.14**	**0.14**	**1.99**	**0.047**	*	0.68	0.26	0.14	0.14	1.85	0.064	.
	Ase-ua4	0.29	0.12	0.23	0.23	0.53	0.599		0.43	0.20	0.16	0.16	1.23	0.217	
B) Reproduction - Count of offspring surviving >3 months (independence)	Intercept		0.00	0.15	0.15	0.03	0.979			**1.72**	**0.17**	**0.17**	**10.11**	**<0.001**	***
	zero-inflated intercept		**-3.43**	**1.05**	**1.06**	**3.25**	**0.001**	**		**-0.60**	**0.22**	**0.22**	**2.73**	**0.006**	**
	Longevity	**1.00**	**3.31**	**0.30**	**0.31**	**10.81**	**<0.001**	***	**1.00**	**3.33**	**0.31**	**0.31**	**10.76**	**<0.001**	***
	Longevity^2	**1.00**	**-1.33**	**0.22**	**0.22**	**6.02**	**<0.001**	***	0.25	-0.09	0.51	0.52	0.17	0.868	
	TLR3: ^{AC}	**0.49**	**-0.34**	**0.17**	**0.17**	**1.97**	**0.049**	*	0.01	0.03	0.30	0.30	0.09	0.927	
	TLR3: ^{CC}		-0.19	0.21	0.22	0.88	0.382			-0.22	0.48	0.48	0.47	0.640	
	Individual Hs	0.27	0.03	0.17	0.17	0.20	0.845		0.40	-0.32	0.28	0.28	1.15	0.248	
	MHC Diversity	0.25	-0.03	0.14	0.14	0.18	0.858		0.35	0.28	0.28	0.28	1.00	0.318	
	Ase-ua4	0.28	0.10	0.18	0.18	0.57	0.570		0.25	-0.07	0.33	0.33	0.20	0.838	

2.4.3. Hardy-Weinberg Equilibrium in fledglings sampled on Cousin

There was a significant deviation from HWE among fledglings (individuals <3 months of age) on Cousin, with a deficiency of heterozygotes (n = 591, F_{IS} = 0.12, P = 0.002, Table S2.4, Fig S2.1a). However, there was no deviation from HWE in those individuals that survived until adulthood (individuals >1 year, n = 380, F_{IS} = 0.08, P = 0.13 Fig S2.1b). Individuals caught <3 months of age were then separated into hatch year, and HWE was assessed for each year. The heterozygote deficiency was consistent across most years (indicated by a positive F_{IS}), but with limited power, only 2007 showed a significant deviation from HWE (n = 53, F_{IS} = 0.31, P = 0.04, Table S2.4).

2.4.4. Spatial and temporal TLR3 *variation across islands*

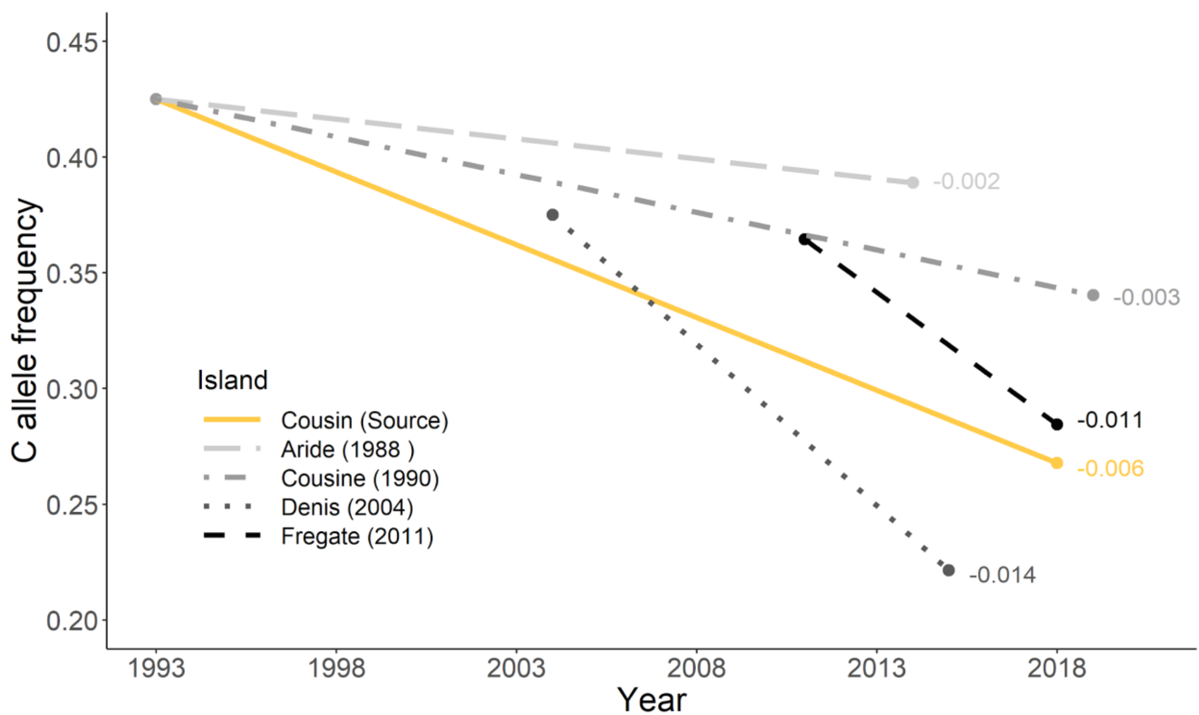


Fig 2.4: Change in the minor allele frequency (C) of the nonsynonymous *TLR3* SNP between two time points in the five isolated island populations of the Seychelles warbler. Points refer to $TLR3^{C}$ allele frequencies of all caught birds at each time point with lines added to emphasize the rate of change. The first time point for Cousin, Aride and Cousine is the 1993-94 Cousin source population (n = 120), whereas the first time points for Denis (2004, n = 56) and Frégate (2011, n = 59) Islands are the translocated individuals. The second time point indicates the most recent sampling event for each island: Cousin (2018, n = 196), Aride (2012 and 2016, n = 54), Cousine (2019, n = 72), Denis (2015, n = 158) and Frégate (2018, n = 58). The translocation year is indicated in the legend. Values represent annual change in frequency of $TLR3^{C}$ allele.

No significant deviation from HWE was observed in any of the different island populations, either pre- or post- translocation (Table S2.5). All populations showed the

same overall trend, with *TLR3*^{C} alleles decreasing in frequency over time (Fig 2.4), but the rate of change differed between islands (Table 2.3, Fig 2.4). As shown above for adults and juveniles, *TLR3*^{C} allele frequencies on Cousin were significantly lower for individuals caught in 2018 compared to 1993-94 (P <0.001; Fig 2.4; Table 2.3). Of the translocated populations, only Denis showed a significant decline in *TLR3*^{C} allele frequency between the initial and most recent sample (15 years difference; P = 0.002; Fig 2.4; Table 2.3). *TLR3* allele frequency temporal differences for Frégate (7 years difference), and between the oldest samples from the source population (Cousin) and the contemporary samples from Aride and Cousine (20 or 28-year difference respectively) were not significant (Fig 2.4; Table 2.3).

Table 2.3: Allelic differentiation of one *TLR3* SNP in the five isolated island populations of the Seychelles warbler between: **A)** two time points for the same island, and **B**) between different pairs of islands using the most recently sampled data. The first time point for Cousin, Aride and Cousine are from the 1993-94 Cousin source population, whereas the first time point for Denis and Frégate are from the translocated individuals. The second time point indicates the most recent sampling event for each island. Significant terms are in bold and underlined

	Population comparisons		χ^2	SE	*P*
A) Old vs recent population samples	**<u>Cousin (1993-94)</u>**	**<u>Cousin (2018)</u>**	**<u>19.44</u>**	**<u>0.00</u>**	**<u><0.001</u>**
	Cousin (1993-94)	Cousine 2019	4.51	0.01	0.105
	Cousin (1993-94)	Aride (2012/16)	1.13	0.01	0.568
	<u>Denis (Translocated)</u>	**<u>Denis (2015)</u>**	**<u>12.09</u>**	**<u>0.00</u>**	**<u>0.002</u>**
	Frégate (Translocated)	Frégate (2018)	3.07	0.01	0.216
B) Between most recent samples on different islands	Cousin (2018)	Cousine (2019)	4.51	0.01	0.105
	<u>Cousin (2018)</u>	**<u>Aride (2012/16)</u>**	**<u>7.66</u>**	**<u>0.00</u>**	**<u>0.022</u>**
	Cousin (2018)	Denis (2015)	3.69	0.01	0.158
	Cousin (2018)	Frégate (2018)	0.41	0.00	0.816
	Aride (2012/16)	Cousine (2019)	1.35	0.01	0.510
	<u>Aride (2012/16)</u>	**<u>Denis (2015)</u>**	**<u>13.74</u>**	**<u>0.00</u>**	**<u>0.001</u>**
	Aride (2012/16)	Frégate (2018)	4.28	0.00	0.118
	<u>Cousine (2019)</u>	**<u>Denis (2015)</u>**	**<u>9.41</u>**	**<u>0.00</u>**	**<u>0.009</u>**
	Cousine (2019)	Frégate (2018)	2.11	0.01	0.349
	Denis (2015)	Frégate (2018)	3.21	0.01	0.201

Focusing on the most recent samples, we found significant *TLR3* differentiation between Denis and Aride (*P* = 0.001; Table 2.3), Denis and Cousine (*P* = 0.009; Table 2.3), and Aride and Cousin (*P* = 0.022; Table 2.3). Denis had the lowest frequency of *TLR3*^{C} alleles (22%) while Aride had the highest (39%) (see Fig 2.4). All other pairwise comparisons were not significant (Table 2.1).

2.5. Discussion

We detected spatial and temporal changes in *TLR3* variation in the Seychelles warbler. On Cousin, we found a consistent decline in the minor allele frequency of the nonsynonymous *TLR3*^{C} allele in the adult population from 40% in 1993, to 29% in 2018 (Fig 2.1). Importantly, differential survival was associated with *TLR3* genotype; individuals with the *TLR3*^{CC} genotype had 37% increased mortality risk compared to those with *TLR3*^{AC} or *TLR3*^{AA} genotypes. Furthermore, males - but not females - with *TLR3*^{CC} or *TLR3*^{AC} genotypes had lower LRS than those with the*TLR3*^{AA} genotype (Fig 2.3a). Even when controlling for longevity, males with the *TLR3*^{AC} genotype had reduced reproduction compared to those with the *TLR3*^{AA} genotype (Fig 2.3b). Notably, the *TLR3* genotypes of nestlings/fledglings deviated from Hardy-Weinberg expectations. Lastly, although we found differences in the *TLR3* minor allele frequency among the island populations (Fig 2.4), they all showed the same pattern of a decrease in the minor allele frequency.

The temporal pattern in our data - with the *TLR3*^{C} allele declining in the population on Cousin over a 25-year period - could be driven by a number of evolutionary forces. However, the lack of migration in or out of Cousin (Komdeur et al. 2004), means it cannot be caused by gene flow. Importantly, our results show that individuals of either sex that were homozygous for *TLR3*^{C} had lower survival and that *TLR3*^{AC} males had a lower rate of reproduction. These differences in survival (and to a lesser degree reproductive rate) resulted, at least in males, in a considerable reduction in LRS; males with one or two copies of the *TLR3*^{C} allele had ca half the reproductive success of those with none (*TLR3*^{AC}: 0.63, *TLR3*^{CC}: 0.70, compared to *TLR3*^{AA}: 1.4 average independent offspring over their lifetime). These results indicate that selection is occurring and may explain the observed change in the *TLR3*^{C} allele frequency over time. Both *TLR3*^{AC} and *TLR3*^{CC} individuals had relatively large selection coefficients of 0.34 and 0.46 respectively. However, it should be noted that the added complication of overlapping generations in a relatively long-lived species could act to dilute the observed selective benefit of *TLR3*^{AA}

genotypes in the short term. While purifying selection in TLRs is the predominant selective mechanism in this multigene family (Alcaide and Edwards 2011), signatures of positive (or balancing) selection have been detected at the codon level in various wild vertebrate species (Areal et al. 2011, Khan et al. 2019, Liu et al. 2019). Indeed, previous work in the Seychelles warbler detected evidence of past positive selection at this *TLR3* locus (Gilroy et al. 2017). The present study now shows that this *TLR3* locus is under strong positive selection (through both survival and reproductive success differences) in the contemporary Cousin population.

Even if selection is acting upon the *TLR3* locus in the Seychelles warbler genetic drift will also occur. Other studies have shown that genetic drift can override the effect of selection in driving immune gene variation (Miller and Lambert 2004, Sutton et al. 2011, Quemere et al. 2015), including TLR variation (Grueber et al. 2013, Gonzalez-Quevedo et al. 2015). However, in the Seychelles warbler the temporal change in allele frequencies at the *TLR3* locus, aligned as it is with the differential fitness of the $TLR3^{C}$ allele, suggest that selection is currently the prevailing force acting upon this locus in this population. Furthermore, a previous study showed that neither neutral microsatellite diversity, nor functional MHC allelic richness, changed over a 18-year time period in the Cousin population, while the mean MHC diversity per individual increased over that time (Wright et al. 2014). This lack of a change at these other loci may suggest that the effect of genetic drift is limited in this already genetically depauperate (Richardson and Westerdahl 2003, Hansson and Richardson 2005) population over the timeframe observed here.

While various studies have linked TLR variation with pathogen infection (Tschirren et al. 2013, Quemere et al. 2015), few have found direct links between TLR variation and fitness in wild populations. In the pale-headed brushfinch (*Atlapetes pallidiceps*), decreased survival was associated with high overall TLR diversity (Hartmann et al. 2014), whilst in song sparrows (*Melospiza melodia*) there was no relationship between survival and TLR heterozygosity (Nelson-Flower et al. 2018), although in both cases the effect of specific alleles was not tested. In the Stewart Island robin (*Petroica australis rakiura*), early life mortality was reduced in individuals with the $TLR4^{BE}$ genotype, compared to other *TLR4* genotypes, despite it being a synonymous substitution (Grueber et al. 2013). Lastly, in Attwater's prairie-chicken (*Tympanuchus cupido attwateri*) the presence of a specific *TLR1B* allele was associated with reduced survival (Bateson et al. 2016). Like the latter two studies, we found the presence of a specific allele to confer differential survival; the $TLR3^{A}$ allele conferred a selective advantage via increased survival, predominantly in early life. Given the importance of *TLR3* as an innate immune

receptor (Barton 2007), and that the SNP investigated causes a functional difference in the binding region, it is likely that the survival differences seen here are due to differential pathogen recognition.

In this study, we also found some evidence of *TLR3* genotypes conferring differential reproductive success in male, but not female warblers. To our knowledge, this is the first-time variation at a TLR gene has been associated with reproductive success in a wild population. In vertebrates, longevity is generally strongly positively correlated with lifetime reproductive success (Clutton-Brock 1988), indeed we found longevity to be the greatest predictor of reproductive success in the Seychelles warbler. However, even after controlling for fitness effects associated with offspring genotype, ability to breed, and longevity we found an effect of male *TLR3* genotype. Combined with differential survival, this resulted in $TLR3^{AA}$ males having considerably greater overall LRS than other genotypes. This observed difference in the reproductive output of males, but not females, could be driven by male-male competition – with males in better condition (through differential immune response due to the *TLR3* variation) better able to outcompete others and gain more social or extra-group offspring. For example, specific alleles at both immune and non-immune loci have been associated with increased competitive ability and increased reproductive success in male vertebrates (Johnston et al. 2013, Sepil et al. 2013).

If female choice is occurring based on the *TLR3* variant in the Seychelles warbler this could explain how only male, and not female, individuals had differential reproduction associated with different *TLR3* genotype. Previous studies, on both the Seychelles warbler (Richardson et al. 2005, Wright et al 2016) and other vertebrate taxa, have focused on MHC-based female mate choice (reviewed in Milinski 2006, Kamiya et al. 2014). As we found a *TLR3* heterozygote deficiency in offspring it is possible that assortative mating could be taking place, whereby individuals' mate with individuals similar to themselves more frequently than expected by chance (Sin et al. 2015). Likewise, as *TLR3* heterozygous individuals do not have greater fitness than *TLR3* homozygous individuals, mate choice is unlikely to be based on *TLR3* heterozygosity. Further investigation should focus on 'good genes' or assortative mating as potential candidate mechanisms in driving the differential reproduction observed in this study.

A third possibility that could explain the pattern of reproductive success linked to TLR variation is that the heterozygote deficit in offspring is due to selection on those offspring. For example, males with $TLR3^{AA}$ genotypes are unable to produce $TLR3^{CC}$ offspring (whoever they breed with), so those males will never suffer from reduced reproductive

success caused by the greater mortality of *TLR3*^{CC} offspring, and thus will have greater LRS. Nonetheless, if this were the sole determinant of the differential reproductive success found in this study, one would expect an equivalent outcome for females. However, there was no effect of *TLR3* genotype on female overall LRS or rate of reproduction, despite females not differing from males in terms of survival linked to the *TLR3* variation. To differentiate between the three non-mutually exclusive mechanisms outlined above, future studies could determine if differences in competitive ability such as body condition and immune responses, and/or differential patterns of mating success are occurring based on this *TLR3* variation.

That there is contemporary positive selection acting upon the *TLR3* locus in the Seychelles warbler provides insight into the evolutionary mechanisms acting upon this important immune locus. The decline in the *TLR3*^{C} allele demonstrated in the current study only represents a snap-shot view of positive selection acting upon this locus. That a selective beneficial polymorphism does exist at this locus despite the considerable bottleneck this species has undergone (Richardson and Westerdahl 2003, Hansson and Richardson 2005), may indicate that balancing selection is acting on this locus over the long-term. Given the role this locus plays in the innate immune response, this is likely to be pathogen-mediated. Of the three main mechanisms by which balancing selection is thought to maintain immune variation (reviewed in Spurgin and Richardson 2010), our study shows that this is not caused by heterozygote advantage (Doherty and Zinkernagel 1975); *TLR3*^{AC} individuals did not gain greater LRS or have increased survival than the homozygote genotypes. The variation observed could potentially be driven by rare allele advantage (Slade and McCallum 1992), or fluctuating selection (Hill et al. 1991), or both. However, differentiating the relative importance of these two mechanisms in driving genetic variation, and separating them from other evolutionary mechanisms is complicated and beyond the scope of the present study (Spurgin and Richardson 2010).

In the present study, we identified a decrease in the *TLR3*^{C} allele frequency over time across all five island populations (Fig 2.4) though they did differ in rate of change. These temporal patterns of *TLR3*^{C} loss suggest that whatever selective agent is acting on Cousin is present on the other islands. Given their very close proximity, and similarity to Cousin - compared to the more isolated islands of Denis and Frégate - the weaker effect on Aride and Cousine is surprising as one may expect close and environmentally similar islands to contain similar pathogens. For example, Cousine (the closest island to Cousin) is the only island to have retained (after translocation) the single strain of the *Haemoproteus nucleocondensus* pathogen that is present in the original Cousin population (Fairfield et al. 2016). A similar pattern of spatio-temporal change in *TLR1LA* diversity between

translocated populations of the New Zealand South Island saddleback, *Philesturnus carunculatus*, was put down to the distribution of malaria parasites (Knafler et al. 2017). However, the distribution of the haemoproteus pathogen found in the Seychelles warbler (not on Aride, Denis or Frégate) means that this cannot be the selective agent here. Work is now needed to identify the pathogen responsible, and determine why the distribution, or impact of this pathogen, differs among the islands.

The avian *TLR3* is orthologous to mammalian *TLR3* and recognises viral dsRNA (including avian pox and influenza viruses) (Hutchens et al. 2008, Brownlie and Allan 2011, Chen et al. 2013). Therefore, it is likely that the selective agent is a virus. Despite this, we have found no obvious evidence of any viral illness in the Seychelles warbler in over thirty years of study. Furthermore, while viruses such avian pox are common in many parts of the world (van Riper III and Forrester 2007) there are no reports of this, or any other virus, circulating in the passerines in the Seychelles (Hutchings 2009). Influenza A has been reported in Procellariformes (petrels and shearwaters) in the Seychelles (Lebarbenchon et al. 2015), but whether this could be passed to the warblers is unknown. It is possible that we just do not see visible signs of a pathogen that is circulating in the warblers because of mild virulence or evolved host tolerance (Råberg 2014, Hammers et al. 2016). Furthermore, individuals may only show visible symptoms during the acute phase of infection when they are also least active, consequently they may be unlikely to be observed before recovery or death (LaPointe et al. 2009).

Even if there are no virulent pathogens currently in the populations, maintaining immunogenetic variation could have important consequences for the future success of this species. If selection continues, the SNP investigated here will go to fixation, and potentially important immunogenetic variation will be lost in the system. This is particularly important given the reduced diversity already present at this, and other innate immune genes, in the Seychelles warbler (Gilroy et al. 2016, Gilroy et al. 2017). The innate immune response is often the organism's first line of defence against pathogens and plays an important role in the evolution to novel disease outbreaks (Bonneaud et al. 2012). Thus, knowing the underlying variation present, and understanding the mechanisms driving evolutionary change at these key functional sites could be important for future species conservation. This is important in small populations and/or those of conservation concern which often have reduced genetic variation. Managing genetic variation in such populations could be important for their adaptive potential, while monitoring pathogen presence may be important to identify and control disease outbreaks - both of which may be crucial for the populations long term survival.

2.5.1. Conclusion

We found strong evidence that selection – acting through both survival and (to a lesser degree) reproduction, was associated with *TLR3* locus variation in the contemporary Cousin population. This suggests that an unknown pathogen is present in the Seychelles warbler population, driving evolution at this *TLR3* locus. It is possible that this current positive selection may be part of a much longer-term pattern of balancing selection, but only further monitoring will be able to determine this.

2.6. References

Acevedo-Whitehouse, K., and Cunningham, A. A. (2006). Is MHC enough for understanding wildlife immunogenetics? *Trends in Ecology and Evolution,* **21**(8), 433-438.

Aderem, A., and Ulevitch, R. J. (2000). Toll-like receptors in the induction of the innate immune response. *Nature,* **406**(6797), 782-787.

Akira, S., Uematsu, S., and Takeuchi, O. (2006). Pathogen Recognition and Innate Immunity. *Cell,* **124**(4), 783-801.

Alcaide, M., and Edwards, S. V. (2011). Molecular evolution of the toll-like receptor multigene family in birds. *Molecular Biology and Evolution,* **28**(5), 1703-1715.

Antonides, J., Mathur, S., Sundaram, M., Ricklefs, R., and DeWoody, A. J. (2019). Immunogenetic response of the bananaquit in the face of malarial parasites. *BMC Evolutionary Biology*, **19**(1), 107.

Apanius, V., Penn, D., Slev, P. R., Ruff, L. R., and Potts, W. K. (1997). The nature of selection on the major histocompatibility complex. *Critical Reviews™ in Immunology,* **17**(2).

Areal, H., Abrantes, J., and Esteves, P. J. (2011). Signatures of positive selection in Toll-like receptor (TLR) genes in mammals. *BMC Evolutionary Biology,* **11**(1), 368.

Barton, G. M. (2007). Viral recognition by Toll-like receptors. *Seminars in Immunology,* **19**(1), 33-40.

Barton, K., and Barton, M. K. (2019). Package ‘MuMIn’ (Version R package version 1.43.6).

Bateson, Z. W., Hammerly, S. C., Johnson, J. A., Morrow, M. E., Whittingham, L. A., and Dunn, P. O. (2016). Specific alleles at immune genes, rather than genome-wide heterozygosity, are related to immunity and survival in the critically endangered Attwater's prairie-chicken. *Molecular Ecology,* **25**(19), 4730-4744.

Bollmer, J. L., Dunn, P. O., Whittingham, L. A., and Wimpee, C. (2010). Extensive MHC Class II B Gene Duplication in a Passerine, the Common Yellowthroat (Geothlypis trichas). *Journal of Heredity,* **101**(4), 448-460.

Bonneaud, C., Balenger, S. L., Zhang, J., Edwards, S. V., and Hill, G. E. (2012). Innate immunity and the evolution of resistance to an emerging infectious disease in a wild bird. *Molecular Ecology,* **21**(11), 2628-2639.

Brooks, M. E., Kristensen, K., van Benthem, K. J., Magnusson, A., Berg, C. W., Nielsen, A., . . . Bolker, B. M. (2017). glmmTMB balances speed and flexibility among packages for zero-inflated generalized linear mixed modelling. *The R journal,* **9**(2), 378-400.

Brouwer, L., Barr, I., Van De Pol, M., Burke, T., Komdeur, J., and Richardson, D. S. (2010). MHC-dependent survival in a wild population: evidence for hidden genetic benefits gained through extra-pair fertilizations. *Molecular Ecology,* **19**(16), 3444-3455.

Brouwer, L., Komdeur, J., and Richardson, D. S. (2007). Heterozygosity–fitness correlations in a bottlenecked island species: a case study on the Seychelles warbler. *Molecular Ecology,* **16**(15), 3134-3144.

Brouwer, L., Richardson, D. S., Eikenaar, C., and Komdeur, J. (2006). The role of group size and environmental factors on survival in a cooperatively breeding tropical passerine. *Journal of Animal Ecology,* **75**(6), 1321-1329.

Brouwer, L., Tinbergen, J. M., Both, C., Bristol, R., Richardson, D. S., and Komdeur, J. (2009). Experimental evidence for density-dependent reproduction in a cooperatively breeding passerine. *Ecology,* **90**(3), 729-741.

Brownlie, R., and Allan, B. (2011). Avian toll-like receptors. *Cell and Tissue Research,* **343**(1), 121-130.

Chen, S., Cheng, A., and Wang, M. (2013). Innate sensing of viruses by pattern recognition receptors in birds. *Veterinary Research,* **44**(1), 82.

Clutton-Brock, T. H. (Ed.) (1988). *Reproductive success: studies of individual variation in contrasting breeding systems*. Chicago IL: University of Chicago Press.

Coulon, A. (2010). genhet: an easy-to-use R function to estimate individual heterozygosity. *Molecular ecology resources,* **10**(1), 167-169.

Crook, J. H. (1960). *The present status of certain rare land birds of the Seychelles islands*.

Croze, M., Živković, D., Stephan, W., and Hutter, S. (2016). Balancing selection on immunity genes: review of the current literature and new analysis in Drosophila melanogaster. *Zoology,* **119**(4), 322-329.

Daszak, P., Cunningham, A. A., and Hyatt, A. D. (2000). Emerging infectious diseases of wildlife--threats to biodiversity and human health. *Science,* **287**(5452), 443-449.

Doblas, L. L., and McClelland, S. (2015). *Seychelles warbler population census on Denis Island*.

Doherty, P. C., and Zinkernagel, R. M. (1975). Enhanced immunological surveillance in mice heterozygous at the H-2 gene complex. *Nature,* **256**(5512), 50-52.

Downing, T., Lloyd, A. T., O'Farrelly, C., and Bradley, D. G. (2010). The Differential Evolutionary Dynamics of Avian Cytokine and TLR Gene Classes. *The Journal of Immunology,* **184**(12), 6993-7000.

Ejsmond, M. J., Radwan, J., and Wilson, A. B. (2014). Sexual selection and the evolutionary dynamics of the major histocompatibility complex. *Proceedings of the Royal Society of London B: Biological Sciences,* **281**(1796), 20141662.

Fairfield, E. A., Hutchings, K., Gilroy, D. L., Kingma, S. A., Burke, T., Komdeur, J., and Richardson, D. S. (2016). The impact of conservation-driven translocations on blood parasite prevalence in the Seychelles warbler. *Scientific Reports,* **6**.

Frankham, R. (1996). Relationship of Genetic Variation to Population Size in Wildlife. *Conservation Biology,* **10**(6), 1500-1508.

Gelman, A., Su, Y., Masanao, Y., Zheng, T., and Dorie, V. (2018). arm: Data Analysis Using Regression and Multilevel/Hierarchical Models, version 1.10-1. In.

Gilroy, D., van Oosterhout, C., Komdeur, J., and Richardson, D. S. (2016). Avian β-defensin variation in bottlenecked populations: the Seychelles warbler and other congeners. *Conservation Genetics,* **17**(3), 661-674.

Gilroy, D., van Oosterhout, C., Komdeur, J., and Richardson, D. S. (2017). Toll-like receptor variation in the bottlenecked population of the endangered Seychelles warbler. *Animal Conservation,* **20**(3), 235-250.

Gonzalez-Quevedo, C., Spurgin, L. G., Illera, J. C., and Richardson, D. S. (2015). Drift, not selection, shapes toll-like receptor variation among oceanic island populations. *Molecular Ecology,* **24**(23), 5852-5863.

Grambsch, P. M., and Ttherneau, T. M. (1994). Proportional hazards tests and diagnostics based on weighted residuals. *Biometrika, 81*(3), 515-526.

Griffiths, R., Double, M. C., Orr, K., and Dawson, R. J. G. (1998). A DNA test to sex most birds. *Molecular Ecology,* **7**(8), 1071-1075.

Grueber, C. E., Wallis, G. P., and Jamieson, I. G. (2013). Genetic drift outweighs natural selection at toll-like receptor (TLR) immunity loci in a re-introduced population of a threatened species. *Molecular Ecology,* **22**(17), 4470-4482.

Guo, S. W., and Thompson, E. A. (1992). Performing the exact test of Hardy-Weinberg proportion for multiple alleles. *Biometrics*, 361-372.

Hadfield, J., Richardson, D., and Burke, T. (2006). Towards unbiased parentage assignment: combining genetic, behavioural and spatial data in a Bayesian framework. *Molecular Ecology, **15***(12), 3715-3730.

Haldane, J. B. S. (1992). Disease and evolution. *Current Science, **63***(9), 599-604.

Hammers, M., and Brouwer, L. (2017). Rescue behaviour in a social bird: removal of sticky ‘bird-catcher tree ‘seeds by group members. *Behaviour, **154***(4), 403-411.

Hammers, M., Kingma, S. A., Bebbington, K., van de Crommenacker, J., Spurgin, L. G., Richardson, D. S., . . . Komdeur, J. (2015). Senescence in the wild: Insights from a long-term study on Seychelles warblers. *Experimental Gerontology,* **71**, 69-79.

Hammers, M., Kingma, S. A., Spurgin, L. G., Bebbington, K., Dugdale, H. L., Burke, T., . . . Richardson, D. S. (2019). Breeders that receive help age more slowly in a cooperatively breeding bird. *Nature Communications,* **10**(1), 1301.

Hammers, M., Richardson, D. S., Burke, T., and Komdeur, J. (2013). The impact of reproductive investment and early-life environmental conditions on senescence: support for the disposable soma hypothesis. *Journal of Evolutionary Biology,* **26**(9), 1999-2007.

Hansson, B., and Richardson, D. S. (2005). *Genetic variation in two endangered Acrocephalus species compared to a widespread congener: estimates based on functional and random loci.* Paper presented at the Animal Conservation forum.

Hartig, F. (2017). DHARMa: residual diagnostics for hierarchical (multi-level/mixed) regression models. *R package version 0.2.4, 5*.

Hartmann, S. A., Schaefer, H. M., and Segelbacher, G. (2014). Genetic depletion at adaptive but not neutral loci in an endangered bird species. *Molecular Ecology,* **23**(23), 5712-5725.

Hedrick, P. W. (1994). Evolutionary genetics of the major histocompatibility complex. *The American Naturalist,* **143**(6), 945-964.

Hedrick, P. W. (1998). Balancing selection and MHC. *Genetica,* **104**(3), 207-214.

Hill, A. V. S., Allsopp, C. E. M., Kwiatkowski, D., Anstey, N. M., Twumasi, P., Rowe, P. A., . . . Greenwood, B. M. (1991). Common West African HLA antigens are associated with protection from severe malaria. *Nature,* **352**(6336), 595-600.

Hutchens, M., Luker, K. E., Sottile, P., Sonstein, J., Lukacs, N. W., Núñez, G., . . . Luker, G. D. (2008). *TLR3* increases disease morbidity and mortality from vaccinia infection. *Journal of immunology (Baltimore, Md. : 1950),* **180**(1), 483-491.

Hutchings, K. (2009). *Parasite-mediated selection in an island endemic, the Seychelles warbler (Acrocephalus sechellensis).* University of East Anglia,

Johnson, T. F., Brown, T. J., Richardson, D. S., and Dugdale, H. L. (2018). The importance of post-translocation monitoring of habitat use and population growth:

insights from a Seychelles Warbler (Acrocephalus sechellensis) translocation. *Journal of Ornithology,* **159**(2), 439-446

Johnston, S. E., Gratten, J., Berenos, C., Pilkington, J. G., Clutton-Brock, T. H., Pemberton, J. M., and Slate, J. (2013). Life history trade-offs at a single locus maintain sexually selected genetic variation. *Nature,* **502(**7469), 93-95.

Kamiya, T., O'Dwyer, K., Westerdahl, H., Senior, A., and Nakagawa, S. (2014). A quantitative review of MHC-based mating preference: the role of diversity and dissimilarity. *Molecular Ecology,* **23**(21), 5151-5163.

Khan, I., Maldonado, E., Silva, L., Almeida, D., Johnson, W. E., O'Brien, S. J., . . . Antunes, A. (2019). The vertebrate TLR supergene family evolved dynamically by gene gain/loss and positive selection revealing a host–pathogen arms race in birds. *Diversity,* **11**(8), 131.

Kingma, S. A., Bebbington, K., Hammers, M., Richardson, D. S., and Komdeur, J. (2016). Delayed dispersal and the costs and benefits of different routes to independent breeding in a cooperatively breeding bird. *Evolution,* **70**(11), 2595-2610.

Kloch, A., Wenzel, M. A., Laetsch, D. R., Michalski, O., Bajer, A., Behnke, J. M., . . . Piertney, S. B. (2018). Signatures of balancing selection in toll-like receptor (TLRs) genes – novel insights from a free-living rodent. *Scientific Reports,* **8**(1), 8361.

Knafler, G. J., Grueber, C. E., Sutton, J. T., and Jamieson, I. G. (2017). Differential patterns of diversity at microsatellite, MHC, and TLR loci in bottlenecked South Island saddleback populations. *New Zealand Journal of Ecology,* **41**(1), 98-106.

Komdeur, J. (1992). Importance of habitat saturation and territory quality for evolution of cooperative breeding in the Seychelles warbler. *Nature,* **358**(6386), 493-495.

Komdeur, J. (1994). Conserving the Seychelles warbler Acrocephalus sechellensis by translocation from Cousin Island to the islands of Aride and Cousine. *Biological Conservation,* **67**(2), 143-152.

Komdeur, J., Piersma, T., Kraaijeveld, K., Kraaijeveld-Smit, F., and Richardson, D. S. (2004). Why Seychelles Warblers fail to recolonize nearby islands: unwilling or unable to fly there? *Ibis,* **146**(2), 298-302.

Lacy, R. C. (1987). Loss of genetic diversity from managed populations: interacting effects of drift, mutation, immigration, selection, and population subdivision. *Conservation Biology,* **1**(2), 143-158.

Lande, R. (1976). Natural Selection and Random Genetic Drift in Phenotypic Evolution. *Evolution, 30*(2), 314-334.

Lande, R. (1995). Mutation and Conservation. *Conservation Biology, 9*(4), 782-791.

LaPointe, D. A., Hofmeister, E. K., Atkinson, C. T., Porter, R. E., and Dusek, R. J. (2009). Experimental infection of Hawaii amakihi (*Hemignathus virens*) with West Nile

virus and competence of a co-occurring vector, Culex quinquefasciatus: potential impacts on endemic Hawaiian avifauna. *Journal of wildlife diseases,* **45**(2), 257-271.

Lebarbenchon, C., Jaeger, A., Feare, C., Bastien, M., Dietrich, M., Larose, C., . . . Dellagi, K. (2015). Influenza A Virus on Oceanic Islands: Host and Viral Diversity in Seabirds in the Western Indian Ocean. *PLOS Pathogens,* ***11***(5), e1004925.

Liu, G., Zhang, H., Zhao, C., and Zhang, H. (2019). Evolutionary History of the Toll-Like Receptor Gene Family across Vertebrates. *Genome Biology and Evolution,* ***12***(1), 3615-3634.

Medzhitov, R. (2001). Toll-like receptors and innate immunity. *Nature Reviews Immunology****, 1***(2), 135-145.

Milinski, M. (2006). The Major Histocompatibility Complex, Sexual Selection, and Mate Choice. *Annual Review of Ecology, Evolution, and Systematics,* ***37***(1), 159-186.

Miller, H. C., and Lambert, D. M. (2004). Gene duplication and gene conversion in class II MHC genes of New Zealand robins (Petroicidae). *Immunogenetics,* ***56***(3), 178-191.

Miller, H. C., and Lambert, D. M. (2004). Genetic drift outweighs balancing selection in shaping post-bottleneck major histocompatibility complex variation in New Zealand robins (Petroicidae). *Molecular Ecology,* ***13***(12), 3709-3721.

Mukherjee, S., Sarkar-Roy, N., Wagener, D. K., and Majumder, P. P. (2009). Signatures of natural selection are not uniform across genes of innate immune system, but purifying selection is the dominant signature. *Proceedings of the National Academy of Sciences,* ***106***(17), 7073-7078.

Nelson-Flower, M. J., Germain, R. R., MacDougall-Shackleton, E. A., Taylor, S. S., and Arcese, P. (2018). Purifying selection in the Toll-like receptors of song sparrows Melospiza melodia. *Journal of Heredity,* ***109***(5), 501-509.

Piertney, S. B., and Oliver, M. K. (2005). The evolutionary ecology of the major histocompatibility complex. *Heredity,* ***96***(1), 7-21.

Pusack, T. J., Christie, M. R., Johnson, D. W., Stallings, C. D., and Hixon, M. A. (2014). Spatial and temporal patterns of larval dispersal in a coral-reef fish metapopulation: evidence of variable reproductive success. *Molecular Ecology,* ***23***(14), 3396-3408.

Quemere, E., Galan, M., Cosson, J. F., Klein, F., Aulagnier, S., Gilot-Fromont, E., . . . Charbonnel, N. (2015). Immunogenetic heterogeneity in a widespread ungulate: the European roe deer (Capreolus capreolus). *Molecular Ecology,* **24**(15), 3873-3887.

Råberg, L. (2014). How to Live with the Enemy: Understanding Tolerance to Parasites. *PLOS Biology,* **12**(11), e1001989.

Raj Pant, S., Komdeur, J., Burke, T. A., Dugdale, H. L., and Richardson, D. S. (2019). Socio-ecological conditions and female infidelity in the Seychelles warbler. *Behavioral Ecology,* **30**(5), 1254-1264.

Raymond, M., and Rousset, F. (1995). An Exact Test for Population Differentiation. *Evolution,* **49**(6), 1280-1283.

Reed, D. H., and Frankham, R. (2003). Correlation between fitness and genetic diversity. *Conservation Biology,* **17**(1), 230-237.

Richardson, D. S., and Westerdahl, H. (2003). MHC diversity in two Acrocephalus species: the outbred Great reed warbler and the inbred Seychelles warbler. *Molecular Ecology,* **12**(12), 3523-3529.

Richardson, D. S., Komdeur, J., Burke, T., and von Schantz, T. (2005). MHC-based patterns of social and extra-pair mate choice in the Seychelles warbler. *Proceedings of the Royal Society B: Biological Sciences*, **272**(1564), 759-767.

Richardson, D. S., Bristol, R., and Shah, N. J. (2006). Translocation of the Seychelles warbler Acrocephalus sechellensis to establish a new population on Denis Island, Seychelles. *Conservation Evidence,* **3**, 54-57.

Richardson, D. S., Burke, T., Komdeur, J., and Dunn, P. (2002). Direct benefits and the evolution of female-biased cooperative breeding in Seychelles warblers. *Evolution,* **56**(11), 2313-2321.

Roach, J. C., Glusman, G., Rowen, L., Kaur, A., Purcell, M. K., Smith, K. D., . . . Aderem, A. (2005). The evolution of vertebrate Toll-like receptors. *Proceedings of the National Academy of Sciences of the United States of America,* **102**(27), 9577-9582.

Robertson, A., and Hill, W. G. (1984). Deviations from Hardy-Weinberg proportions: sampling variances and use in estimation of inbreeding coefficients. *Genetics,* **107**(4), 703-718.

Robinson, Jacqueline A., Ortega-Del Vecchyo, D., Fan, Z., Kim, Bernard Y., vonHoldt, Bridgett M., Marsden, Clare D., . . . Wayne, Robert K. (2016). Genomic Flatlining in the Endangered Island Fox. *Current Biology,* **26**(9), 1183-1189.

Rousset, F. (2008). genepop'007: a complete re-implementation of the genepop software for Windows and Linux. *Molecular ecology resources,* **8**(1), 103-106.

Sepil, I., Lachish, S., and Sheldon, B. C. (2013). Mhc-linked survival and lifetime reproductive success in a wild population of great tits. *Molecular Ecology,* **22**(2), 384-396.

Sin, Y. W., Annavi, G., Newman, C., Buesching, C., Burke, T., Macdonald, D. W., and Dugdale, H. L. (2015). MHC class II-assortative mate choice in European badgers (Meles meles). *Molecular Ecology,* **24**(12), 3138-3150.

Slade, R., and McCallum, H. (1992). Overdominant vs. frequency-dependent selection at MHC loci. *Genetics,* **132**(3), 861.

Sommer, S. (2005). The importance of immune gene variability (MHC) in evolutionary ecology and conservation. *Frontiers in zoology,* **2**(1), 1.

Sparks, A. M., Spurgin, L. G., van der Velde, M., Fairfield, E. A., Komdeur, J., Burke, T., . . . Dugdale, H. (2020). Telomere heritability and parental age at conception effects in a wild avian population.

Spurgin, L. G., and Richardson, D. S. (2010). How pathogens drive genetic diversity: MHC, mechanisms and misunderstandings. *Proceedings of the Royal Society B: Biological Sciences,* **277**(1684), 979-988.

Spurgin, L. G., Wright, D. J., van der Velde, M., Collar, N. J., Komdeur, J., Burke, T., and Richardson, D. S. (2014). Museum DNA reveals the demographic history of the endangered Seychelles warbler. *Evolutionary Applications,* **7**(9), 1134-1143.

Sutton, J. T., Nakagawa, S., Robertson, B. C., and Jamieson, I. G. (2011). Disentangling the roles of natural selection and genetic drift in shaping variation at MHC immunity genes. *Molecular Ecology,* **20**(21), 4408-4420.

Therneau, T. (2019). Mixed effects Cox models (Version R package version 2.2-14).

Therneau, T. M., and Lumley, T. (2015). Package ‘survival’. *R Top Doc, 128*.

Tschirren, B., Andersson, M., Scherman, K., Westerdahl, H., Mittl, P. R., and Råberg, L. (2013). Polymorphisms at the innate immune receptor TLR2 are associated with Borrelia infection in a wild rodent population. *Proceedings of the Royal Society B: Biological Sciences,* **280**(1759), 20130364.

van Oosterhout, C. (2009). A new theory of MHC evolution: beyond selection on the immune genes. *Proceedings of the Royal Society B: Biological Sciences,* **276**(1657), 657-665.

van Riper III, C., and Forrester, D. J. (2007). Avian pox. *Infectious diseases of wild birds*, 131-176.

Wright, D. J. (2014). *Evolutionary and conservation genetics of the Seychelles warbler (Acrocephalus sechellensis).* University of East Anglia.

Wright, D. J., Shah, N. J., and Richardson, D. S. (2014). Translocation of the Seychelles warbler *Acrocephalus sechellensis* to establish a new population on Frégate Island, Seychelles. *Conservation Evidence,* **11**, 20-24.

Wright, D. J., Spurgin, L. G., Collar, N. J., Komdeur, J., Burke, T., and Richardson, D. S. (2014). The impact of translocations on neutral and functional genetic diversity within and among populations of the Seychelles warbler. *Molecular Ecology,* **23**(9), 2165-2177.

Wright, D. J., Brouwer, L., M.-E. Mannarelli, M.-E., Burke, T., Komdeur, J. and Richardson, D. S. (2016). Social pairing of Seychelles warblers under reduced

constraints: MHC, neutral heterozygosity, and age. *Behavioral Ecology* **27**(1), 295-303.

Wright, S. (1931). Evolution in Mendelian populations. *Genetics,* **16**(2), 97-159.

2.7. Supplementary material

Table S2.1: Information on translocations and follow up periods for five Seychelles Warbler populations. Cousin Island is the source population, for which we used samples collected in 1993/94 for the first time point, the other four Islands were established by translocation. The follow up period indicates the most recent sampling event for each island, and the number of individuals caught in that period.

Island	Translocation/ sample year	Translocated individuals (M:F)	Island size (km²)	Carrying capacity	Population at carrying capacity?	Follow up year	Follow up sample size
Cousin	1993/4	120 (65:55)	0.29	320	Yes	2018	196
Aride	1988	29 (16:13)	0.68	1,850	Yes	2012/16	54
Cousine	1990	29 (15:14)	0.25	210	Yes	2019	72
Denis	2004	58 (24:34)	1.42	500*	Increasing	2015	158
Frégate	2011	59 (23:36)	2.19	500*	Increasing	2018	58

*2,000 with habitat regeneration

Table S2.2: Time-dependent Cox Regression modelling to test the effects of *TLR3* allele presence on bi-annual survival in the Seychelles warbler population (*n* = 517) on Cousin. Coef = hazard rate; SE (coef) = standard error of the hazard rate; HR = hazard ratio. An HR >1 indicates increased hazard of mortality, and <1 indicates decreased hazard of mortality. Coefficient estimates are in reference to $TLR3^{A}$ allele =Present, $TLR3^{C}$ allele = Present, *Ase-ua4* = Present, Season born=Major, Sex= Female.
Significant terms are in bold and underlined.

Factor	coef	SE (coef)	HR	z	*P*
$TLR3^{A}$ allele	**-0.33**	**0.15**	**0.72**	**-2.24**	**0.025**
$TLR3^{C}$ allele	-0.01	0.10	0.99	-0.08	0.940
Individual H_s	-0.12	0.23	0.89	-0.52	0.600
MHC Diversity	-0.02	0.03	0.98	-0.78	0.440
Ase-ua4	**-0.29**	**0.13**	**0.75**	**-2.20**	**0.028**
Maternal H_s	-0.08	0.22	0.92	-0.37	0.710
Season born	-0.22	0.12	0.80	-1.86	0.062
Sex	-0.02	0.10	0.98	-0.19	0.850
Random effects	Variance	517 individuals			
Hatch year	0.015	9 hatch years			

Table S2.3: Reproductive success in male and female Seychelles warblers in relation to *TLR3* allele presence: **A**) Lifetime reproductive success for all birds, **B**) Reproductive success controlling for longevity for birds that survived to adulthood. Zero-inflated GLMMs were used to generate conditional model-averaged values for all predictors featuring in the top model set ($\Delta AIC_c \leq 7$). Model-averaged estimates (β), their standard error (SE), adjusted SE, z value, *P* value, and relative importance (ω) are shown for all predictors featuring in the top model set ($\Delta AIC_c \leq 7$). Estimates are in reference to *TLR3*^{A} allele =Present, *TLR3*^{C} allele = Present, *Ase-ua4* = Present. Hatch year is a random effect for all models, except for Female B which was run as an GLM. Significant terms are in bold and underlined.

*** $P < 0.001$, ** $P < 0.01$, * $P < 0.05$.

Response	Factor	Male (**A:** n=224; **B:** n=145)							Female (**A:** n=263; **B:** n=178)						
		ω	β	SE	Adj. SE	z	p		ω	β	SE	Adj. SE	z	p	
A) LRS - Count of offspring surviving > 3 months (independence)	Intercept		**1.08**	**0.26**	**0.26**	**4.20**	**<0.001**	***		**0.88**	**0.22**	**0.22**	**3.94**	**<0.001**	***
	zero-inflated intercept		**0.49**	**0.16**	**0.16**	**2.96**	**0.003**	**		**0.53**	**0.14**	**0.14**	**3.67**	**<0.001**	***
	TLR3^{A} allele	0.26	0.05	0.29	0.29	0.18	0.857		0.33	0.20	0.25	0.25	0.78	0.433	
	TLR3^{C} allele	**1.00**	**-0.70**	**0.17**	**0.17**	**4.12**	**<0.001**	***	0.27	0.03	0.15	0.15	0.23	0.822	
	Individual Hs	0.26	0.04	0.17	0.17	0.26	0.792		0.58	-0.24	0.14	0.14	1.65	0.098	.
	MHC Diversity	**0.71**	**-0.28**	**0.14**	**0.14**	**1.97**	**0.049**	*	0.68	0.26	0.14	0.14	1.86	0.063	.
	Ase-ua4	0.28	0.12	0.23	0.23	0.52	0.605		0.43	0.20	0.16	0.16	1.25	0.213	
B) Reproduction - Count of offspring surviving >3 months (independence)	Intercept		0.07	0.21	0.21	0.32	**0.748**			**1.66**	**0.30**	**0.30**	**5.53**	**<0.001**	***
	zero-inflated intercept		**-3.36**	**1.00**	**1.01**	**3.34**	**0.001**	***		**-0.60**	**0.22**	**0.22**	**2.73**	**0.006**	**
	Longevity	**1.00**	**3.29**	**0.30**	**0.30**	**10.80**	**<0.001**	***	**1.00**	**3.32**	**0.31**	**0.31**	**10.79**	**<0.001**	***
	Longevity^{2}	**1.00**	**-1.33**	**0.22**	**0.22**	**6.04**	**<0.001**	***	0.24	-0.09	0.51	0.52	0.16	0.870	
	TLR3^{A} allele	0.27	-0.09	0.26	0.26	0.37	0.715		0.27	0.24	0.46	0.47	0.52	0.606	
	TLR3^{C} allele	0.72	-0.30	0.16	0.16	1.93	0.054	.	0.24	-0.01	0.29	0.29	0.04	0.969	
	Individual Hs	0.25	0.04	0.16	0.16	0.28	0.783		0.40	-0.33	0.28	0.28	1.17	0.243	
	MHC Diversity	0.24	-0.04	0.14	0.14	0.27	0.785		0.35	0.28	0.28	0.28	1.00	0.318	
	Ase-ua4	0.27	0.10	0.18	0.18	0.58	0.561		0.24	-0.07	0.33	0.33	0.21	0.833	

Table S2.4: Testing Hardy-Weinberg Equilibrium and inbreeding (F_{IS}) at one*TLR3* SNP from all Seychelles warblers first caught before 3 months of age from the Cousin population. Data are separated by hatch year. At the bottom are the totals for the hatch years combined. Significant terms are in bold and underlined

Hatch Year	*n* (%)	*TLR3*^{A}	*TLR3*^{C} allele	F_{IS}	SE	*P*
1992	-	-	-	-	-	-
1993	-	-	-	-	-	-
1994	-	-	-	-	-	-
1997	31 (52%)	0.61	0.39	0.07	0	1
1998	50 (75%)	0.55	0.45	0.24	0.003	0.1
1999	88 (82%)	0.61	0.39	-0.16	0.004	0.17
2005	83 (75%)	0.66	0.34	0.09	0.004	0.46
2006	89 (82%)	0.66	0.34	0.08	0.004	0.49
2007	**53 (76%)**	**0.65**	**0.35**	**0.31**	**0.001**	**0.04**
2008	38 (68%)	0.63	0.37	0.11	0.001	0.72
2009	37 (45%)	0.81	0.19	0.13	0.002	0.58
2010	23 (33%)	0.61	0.39	0.30	0.002	0.21
2016	13 (25%)	0.77	0.23	0.18	0.002	0.52
2017	47 (53%)	0.72	0.28	0.16	0.003	0.29
2018	39 (66%)	0.77	0.23	0.29	0.002	0.09
All	**591**	**0.65**	**0.35**	**0.12**	**0.001**	**0.004**

Table S2.5: Allele frequencies, sample size, F_{is}, and HWE, from one *TLR3* SNP across two time points in five Seychelles Warbler populations. The first time point for Cousin, Aride and Cousine are from the 1993-94 Cousin source population whereas the first time point for Denis and Frégate are from the translocated individuals. The second time point indicates the most recent sampling event for each island.

Island (Year)	*n*	*TLR3*^{A}	*TLR3*^{C}	F_{is}	SE	*P*
Cousin (1993-94)	120	0.59	0.41	0.12	0.005	0.11
Cousin (2018)	196	0.74	0.26	0.05	0.004	0.58
Aride (2012/16)	54	0.61	0.39	0.08	0.002	0.77
Cousine (2019)	68	0.67	0.33	-0.02	0.000	1.00
Denis (2004-translocated)	56	0.63	0.38	-0.06	0.002	0.78
Denis (2015)	158	0.78	0.22	-0.02	0.002	0.82
Frégate (2011-translocated)	59	0.64	0.36	0.02	0.000	1.00
Frégate (2018)	56	0.71	0.29	-0.13	0.003	0.51

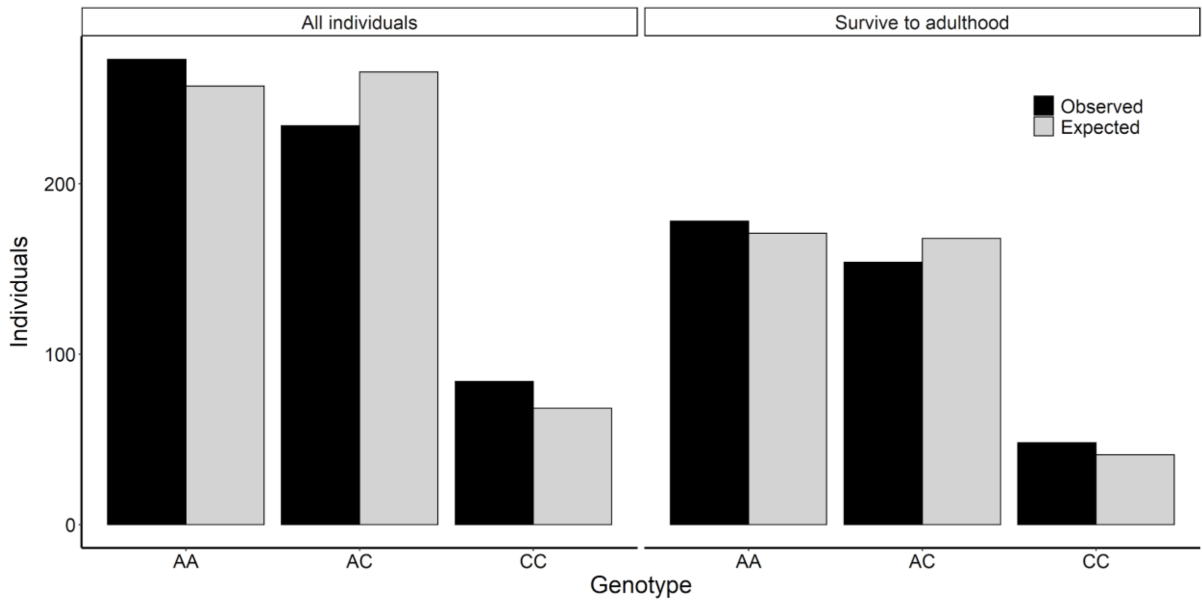


Fig S2.1: Observed and expected values for *TLR3* genotypes in the Cousin population of the Seychelles warbler for **A**) all individuals first caught before 3 months of age (*n* = 591), **B**) individuals first caught before 3 months of age which survive to adulthood (*n* = 380). Observed values are shown in black and expected values shown in light grey, separated by genotype.

Chapter 3

Sexual selection shapes variation in the viral-sensing *TLR3* gene in a natural population

Life of a warbler chick

3.1. Abstract

Sexual selection can be a powerful evolutionary force shaping genetic variation. However, very few studies have investigated the role sexual selection may take in shaping variation at genes involved in the innate immune system, despite evidence that pathogen-mediated balancing selection is acting at these loci. Previous work in a bottlenecked population of Seychelles warblers (*Acrocephalus sechellensis*) indicated that viability selection is driving contemporary evolution at one non-synonymous SNP in the viral-sensing Toll-like Receptor 3 (*TLR3)* gene. Here, we investigate whether pre- and post-copulatory sexual selection is also acting at this locus. Results show that *TLR3* genotype was not related to dominance gain, or extra-pair success, but $TLR3^{CC}$ males had reduced annual within-group paternity, potentially explaining their reduced reproductive success. Secondly, we found evidence of *TLR3*-based assortative social pairing; $TLR3^{AA}$ females were more likely - and $TLR3^{AC}$ females less likely - to pair with $TLR3^{AA}$ males, although $TLR3^{CC}$ females exhibited random pairing. Third, we found evidence of inheritance bias against paternal – but not maternal - $TLR3^{C}$ haplotypes (post-copulatory mechanism). The latter two results could explain the deficiency of $TLR3^{AC}$ heterozygote offspring found in this population. Thus, both pre- and post-copulatory sexual selection, as well as naturally selection, may act to shape *TLR3* evolution in the Seychelles warbler. However, as of yet, both the phenotypic cues that could mediate *TLR3*-dependent social mate choice, and the mechanistic basis behind the paternally inherited allelic bias in offspring are unknown and warrant further investigation.

3.2. Introduction

Sexual selection – acting through a range of mechanisms – can be a powerful force shaping genetic variation in natural populations (Andersson 1994). Pre-copulatory mechanisms can act via intra-sexual competition – either for the quantity or quality of mates, or resources that influence mating success (Clutton-Brock 2007, Rosvall 2011) – or via inter-sexual competition, through mate choice based on specific traits in the opposite sex (reviewed in Edward 2014). Similarly, post-copulatory mechanisms can occur either pre-fertilization via cryptic mate-choice (Eberhard 1996) and sperm competition (Birkhead and Møller 1998), or post-fertilisation via sperm/egg incompatibility and differential offspring investment (Møller and Thornhill 1998, Birkhead and Pizzari 2002). Preference for certain traits can provide direct advantages through enhanced resources or parental care, or through indirect advantages, such as acquiring heritable traits which increase offspring fitness (Andersson 1994). Disassortative mating can help avoid inbreeding, and increase heterozygosity of offspring, whilst assortative mating can aid adaptation to local selective pressures (Jiang et al. 2013).

Variation in the fitness-dependent traits which sexual selection acts upon may be pathogen-mediated (Hamilton and Zuk 1982). Such traits can be specific heritable structural or ornamental traits (Johnston et al. 2013), or more general traits such as body size or immunocompetence (Wilkinson et al. 2015), and may be determined by genome-wide variation, or by specific loci. Indeed, variation at immunity related genes contributes greatly to difference in individual survival and vigour and, in turn, may be maintained by pathogen-mediated balancing selection (Bernatchez and Landry 2003, Ferrer-Admetlla et al. 2008, Spurgin and Richardson 2010, Eizaguirre et al. 2012). Thus, it is not surprising if immune genes also play a role in sexual selection (reviewed in Edwards and Hedrick 1998, Milinski 2006).

Indeed, immunogenetic variation is shaped by intra-sexual selection through male–male competition (Setchell et al. 2010), and by inter-sexual competition (Penn and Potts 1999) often with a focus on female mate choice (Kamiya et al. 2014, Whitcomb et al. 2014, Brandies et al. 2018). In vertebrates, research on pathogen-mediated sexual selection has mainly focused on the Major Histocompatibility Complex (MHC) genes (e.g., Løvlie et al. 2013, Ejsmond et al. 2014) which play a key role in detecting pathogens and initiating an adaptive immune response (Piertney and Oliver 2005). Polymorphisms in these genes confer differential pathogen resistance and susceptibility (Worley et al. 2010, Westerdahl et al. 2013, Sin et al. 2014), and thus directly affect fitness (Hill et al. 1991). However, with some notable exceptions (e.g., Emlen et al. 2012, Khila et al. 2012,

Johnston et al. 2013), there are few examples of other genes that underlie sexual selection (reviewed in Wilkinson et al. 2015). This is despite the fact that other immune genes, particularly those of the innate immune system, are under pathogen-mediated balancing selection (Ferrer-Admetlla et al. 2008, Areal et al. 2011, Grueber et al. 2014, Velová et al. 2018).

Toll-Like Receptor (TLR) genes, of the innate immune system, can evolve rapidly as a result of pathogen-mediated selection (Downing et al. 2010, Khan et al. 2019, Liu et al. 2019). Vertebrate TLRs can be divided into six families (Roach et al. 2005), each encoding a molecule that binds to a specific evolutionary conserved structure integral to the functioning of the relevant pathogen group (Medzhitov 2001). For example, *TLR4* binds to bacterial lipopolysaccharide (Poltorak et al. 1998), while *TLR3* binds to viral dsRNA (Barton 2007). Once bound, the TLR molecule triggers a cascade of processes associated with the innate and adaptive immune responses (Akira et al. 2006). While the majority of the TLR structure is conserved, variation within the pathogen associated molecular pattern binding region confers differential pathogen resistance and susceptibility (Schröder and Schumann 2005, Kloch et al. 2018, Antonides et al. 2019), increasing survival likelihood (Grueber et al. 2013, Bateson et al. 2016). Unlike the MHC, where mechanisms such as gene duplication, recombination and gene conversion (Hedrick 1994, Miller and Lambert 2004, Bollmer et al. 2010), result in high complexity, TLR genes show relatively simple, independent, evolution (Alcaide and Edwards 2011). Thus, TLR genes provide more tractable candidate genes to study how evolutionary forces shape variation in wild populations (reviewed in Acevedo-Whitehouse and Cunningham 2006). However, whether TLR gene variation may also drive sexual selection is unknown.

The Seychelles warbler (*Acrocephalus sechellensis*) harbours reduced genetic variation as a result of a past genetic bottleneck (Spurgin et al. 2014). However, variation has been maintained at key immune loci, including MHC class I genes (Richardson and Westerdahl 2003, Hansson and Richardson 2005), MHC class II genes (Chapter 4), and some TLR genes (Gilroy et al. 2017) in the Cousin population. In particular, one non-synonymous SNP in the *TLR3* gene is under strong and consistent contemporary selection – being associated with both differential survival and reproductive success (Chapter 2). Specifically, individuals (of both sexes) with the *TLR3*^{CC} genotype have reduced survival, and males - but not females - that carry at least one copy of the *TLR3*^{C} allele have reduced reproductive success than those with two copies of the *TLR3*^{A} allele. Furthermore, a heterozygote deficiency was observed at the *TLR3* locus in young offspring (< 3 months; Chapter 2). However, whether sexual selection plays a role in the

differential reproductive success observed in relation to *TLR3* variation in the Seychelles warbler remains to be resolved.

The isolated population of Seychelles warblers on Cousin Island provides an excellent system for investigating evolutionary questions (e.g. Hammers et al. 2015, Spurgin et al. 2018, Hammers et al. 2019): extensive monitoring of all individuals over their lifetimes, combined with the lack of inter-island dispersal (Komdeur et al. 2004), allows exceptionally accurate measures of fitness components (Chapter 2), and a wealth of longitudinal data (Hammers et al. 2015, Raj Pant et al. 2020). Habitat saturation, combined with social fidelity (Komdeur 1992, Richardson et al. 2007) and considerable longevity (Hammers and Brouwer 2017) act to constrain social mate choice on Cousin Island (Richardson et al. 2005). However, extra-pair paternity (EPP) occurs frequently (Richardson et al. 2001), with 41% of offspring fathered by extra-pair males from outside the natal territory (Raj Pant et al. 2019) and has been linked to MHC characteristics (Richardson et al. 2005) and male age (Raj Pant et al. 2020). Consequently, both social and extra-pair mate choice may be important in this system.

Various mechanisms of sexual selection may drive the observed differential male reproductive success (and offspring heterozygote deficiency), associated with the *TLR3* locus in the Seychelles warbler (Chapter 2). Here we test a suite of questions linked to pre- and post-copulatory sexual selection (see Table 1): **1)** Is *TLR3* genotype associated with variation in size or body condition – which may indicate a difference in competitive ability? As $TLR3^{CC}$ individuals had reduced survival compared to $TLR3^{AA}$/$TLR3^{AC}$ individuals, we predict that $TLR3^{CC}$ individuals will be smaller/ in worse condition than their counterparts. **2)** Is the ability to gain a dominant position in a breeding territory (a prerequisite for male reproduction) linked to TLR3 genotype? **3)** Is social mate choice related - either via, **a)** assortative mating, or **b)** a directional preference, to *TLR3* genotype? Assortative mating in relation to the *TLR3* genotype could potentially explain the *TLR3* heterozygote deficiency found in young offspring (Chapter 2). **4)** Is the likelihood of offspring being extra-pair (EPP) dependent on either the social male's *TLR3* genotype or that of the mothers? **5)** Is the annual within-group (WGP), or extra-group paternity (EGP) success of the social male partner associated with *TLR3* genotype? Lastly, **6)** does the frequency of observed offspring *TLR3* genotypes deviate from mendelian expectations given their genetic parents? Unless otherwise stated, we predict that only males show variation based on *TLR3* genotype in relation the above questions, as only males had decreased reproductive success (Chapter 2). Furthermore, we predict that either $TLR3^{AC}$ or $TLR3^{CC}$ males would be disadvantaged compared to $TLR3^{AA}$ males.

Table 1: Hypotheses for whether TLR3 genotype could be shaped by sexual selection, as well as positive selection in the Seychelles warbler. Table shows potential mechanisms tested, what effects were assessed, and whether any TLR3 related associations were found.

Mechanism	Does TLR3 genotype effect	TLR3 effects assesed for	Any TLR3 differences?	Evidence
Competitive ability	*Physiological condition*	a) Body mass - Male	**Yes** - *TLR3* ^{AC} males tended to be heavier than *TLR3* ^{AA} males	Fig 1; Table 2
		b) Body mass - Female	**Yes** - *TLR3* ^{AC} females are heavier than *TLR3* ^{CC}	
		c) Size - Male	**No**	
		d) Size - Female	**No**	
Social mate choice preferences	*Dominance gain*	a) Male - by 2 years	**No** - But interaction between TLR3 genotype and natal group size	Fig 2; Table 3
		b) Female - by 2 years	**No**	
	Assortative mating		**Yes**	Fig 3a
	Mate choice preferences	a) Female *TLR3* ^{AA}	**Yes** - Pair with males with more A alleles	Fig 3b-d
		b) Female *TLR3* ^{AC}	**Yes** - Pair with males with fewer A alleles	
		c) Female *TLR3* ^{CC}	**No**	
Genetic mate choice preference	*Offspring EGP liklihood*	a) Social male	**No**	Fig 4; Table 4
		b) Mother	**Yes** - *TLR3* ^{AC} mothers tended to have fewer EPP offspring	
	Annual EGP	a) Dominant male	**No**	Fig 5a; Table 5
	Annual WGP	a) Dominant male	**Yes** - *TLR3* ^{CC} males have reduced WPP	Fig 5b; Table 5
Post-copulatory	*Mendelian pattern*	a) Offspring genotype	**No**	Fig 6
		b) Paternal inheritance	**Yes** - Offspring more likely to inherit paternal TLR3 ^{A} alleles	Fig 7

3.3. Methods

3.3.1. Study system

The Seychelles warbler is a small, insectivorous, passerine, endemic to the Seychelles archipelago. Individuals can live for a maximum of 19 years, with a median post-fledgling lifespan of 5.5 years (Hammers and Brouwer 2017). Most territories consist of a socially monogamous breeding pair defending a strict year-round territory (Komdeur 1992). Breeders of both sexes engage in territory defence (Kingma et al. 2017) and contribute to offspring care (Komdeur 1994b). Females typically lay single-egg clutches, although two or three eggs clutches occasionally occur (Richardson et al. 2001). Seychelles warblers can reproduce from eight months of age (Komdeur 1997), however due to habitat saturation some adult birds delay independent breeding and become subordinates (Komdeur 1992), often, but not always, on their natal territory (Eikenaar et al. 2007, Kingma et al. 2016a). One to five sexually mature subordinates exist on *ca* 50% of territories per year (Hammers et al. 2019) and may help to raise offspring on that territory (thus becoming helpers) (Komdeur 1992, Komdeur 1994b, Richardson et al. 2002). Although *ca* 40% of female subordinates gain reproductive success by co-breeding, very few offspring are sired by a subordinate male (Richardson et al. 2002, Raj Pant et al. 2019). However, EPP often occurs (Richardson et al., 2001, Raj Pant et al., 2019). MHC class I variation has been linked to individual fitness: while social mate choice is not MHC-dependent (Richardson et al. 2005, Wright et al. 2016) gaining extra-pair fertilisations is (Richardson et al. 2005), resulting in greater MHC diversity and survival in offspring (Brouwer et al. 2010).

The Seychelles warbler population on Cousin Island (4°20’S,55°40’E; 0.29km^2) has been monitored since 1985 (Komdeur 1992) but with increasing intensity from 1997 onwards (Hammers et al. 2019). Individuals are first caught in the nest, or as dependent juveniles on the natal territory using mist-nets (for details see Kingma et al. 2016b). Each individual is given a British Trust for Ornithology (BTO) metal ring and a unique combination of three colour rings (Richardson et al. 2001). Age is calculated based on hatch date, and eye colour (Komdeur 1992, Wright 2014a). Furthermore, *ca* 30% of the adult population are also recaptured each year (Brown et al. 2021). Morphometric measurements including body mass (to the nearest 0.1g), and right tarsus (to the nearest 0.1mm), are taken. Blood samples (25 µl) are collected by brachial venipuncture and stored in 0.8 ml of absolute ethanol at 4°C.

The Cousin Island population has a carrying capacity of around 320 adult individuals, existing in ca 115 territories (Komdeur 1992, Sparks et al. 2020). Each year the identity and status of individuals in each territory are identified through comprehensive population censuses conducted during the major (June-September), and minor (November-March) breeding seasons, except for the 2000-2 and 2006 minor seasons (Brouwer et al. 2010, Hammers et al. 2019). Group membership and status are identified based on a combination of the long-term demographic data, and field observations including foraging location, interactions, and proximity between individuals (Richardson et al. 2002). There is virtually no migration to or from Cousin (0.1%, (Komdeur et al. 2004)), and the rate of annual resighting of individuals is high (0.98 ± 0.01, (Brouwer et al. 2010)), thus enabling accurate survival estimates. Individuals are recorded as present if observed during the field season and presumed dead if not seen for two consecutive breeding seasons (Hammers et al. 2013). Territory quality is calculated based on insect abundance, vegetation cover, and territory size (as in Komdeur 1992, van de Crommenacker et al. 2011a).

3.3.2. Molecular methods

Genomic DNA was extracted from blood using either a salt extraction technique (Richardson et al. 2001), or the DNeasy blood and tissue kit (Qiagen). Sex was determined via PCR (Griffiths et al. 1998), and individuals were genotyped at 30 polymorphic microsatellite loci (Richardson et al. 2001, Sparks et al. 2020). Standardized individual microsatellite heterozygosity (Hs) was calculated using the R package Genhet 3.1 (Coulon 2010) (excluding two microsatellite loci as they were pooled alleles (see Sparks et al., 2020)), and parentage assigned using MasterBayes 2.52 (Hadfield et al. 2006); for full details see Sparks et al. (2020). A non-synonymous SNP within the leucine-rich repeat domain of *TLR3* exon 4 had previously been screened for individuals as part of another study; for details see Chapter 2. Variation at exon 3 of the MHC class I loci had also been screened for previously; for details see (Richardson and Westerdahl 2003, Wright et al. 2014b, Wright et al. 2016).

3.3.3. Statistical analysis

Unless otherwise stated, all analyses were conducted in R 3.6.1. MHC-I diversity has previously been linked to extra-pair paternity likelihood (Richardson et al. 2005), reproductive success (Chapter 2), and offspring survival (Brouwer et al. 2010). Additionally, the presence of the MHC-1 allele, *Ase-ua4*, is also linked to increased

survival (Brouwer et al. 2010, Chapter 2). Therefore, these variables were also included in analyses.

3.3.3.1. Competitive ability: Physiological condition

Linear mixed models (LMM) with a Gaussian error structure were constructed to determine whether individual condition or size were associated with *TLR3* variation. These analyses used measurements from all birds hatched between 1997-1999 and 2005-2010, for which we had *TLR3* and MHC-I genotypes. To reduce the confounding effects of development, senescence and season, individuals for which we did not have accurate age data were excluded, as were measurements from birds under one year old (juveniles), those potentially undergoing senescence (over six years of age) and those taken during the minor field season, or while carrying an egg. The sexes were modelled separately, as males are bigger than females (van de Crommenacker et al. 2011b). Either body mass (g), as a metric of condition, or right tarsus length (mm), as a metric of size, were used as the dependent variable. *TLR3* genotype ($TLR3^{AA}$, $TLR3^{AC}$ or $TLR3^{CC}$) was included as a fixed factor, along with MHC-1 diversity, *Ase-ua4* (present or absent), individual Hs, catch period (am or pm), and age (months). Tarsus length was also included as a covariate in the body condition model to correct for structural size differences between individuals. Individual identity and catch year were included as random factors to control for year effects.

3.3.3.2. Social mate choice: Likelihood of gaining dominance

Generalised linear mixed models (GLMM) with a binomial error structure were constructed to determine whether dominance gain was associated with either male or female *TLR3* genotype. Likelihood of gaining dominance by two years of age was used as the predictor, as the mean ± SE age at first dominance was 1.8 ± 0.04 years in this species (this study). Birds hatched between 1997-1999 and 2005-2010, for which we had full *TLR3,* and MHC-I information were included. Individuals which died before two years of age, and those first caught over one year of age, for which we did not have accurate age data, were excluded. *TLR3* genotype ($TLR3^{AA}$, $TLR3^{AC}$ or $TLR3^{CC}$), individual Hs (continuous), MHC-I diversity (continuous), and *Ase-ua4* (present, absent) were included as fixed factors, with hatch year included as a random factor to control for cohort effects. We also included season hatched (Major, Minor), natal group size (continuous), and natal territory quality (continuous), as these factors could affect dispersal likelihood and dominance gain (Eikenaaret al. 2007, Kingma et al. 2016a). The sexes were modelled separately as females, but not males, can co-breed (Richardson et al. 2001), thus it is likely that different factors and constraints related to gaining

dominance act upon males and females. As all males with the *Ase-ua4* allele gained dominance, this allele was excluded from the male analysis, to aid model convergence.

3.3.3.4. Social mate choice: Randomisation tests

First, to test for assortative mating, a *TLR3* pair similarity value was calculated for all social male/female pairs (observed and potential pairings), with zero denoting no shared alleles, 0.5 denoting one shared allele, and 1 denoting that both alleles were shared. Second, to test for directional preferences by the social female based on social male *TLR3* genotype, the number of *TLR3*^{A} alleles the social male had (zero, one, two) was calculated for all observed and potential pairings, and separate randomisation tests run for each female *TLR3* genotype. For both questions, randomisation tests were conducted using Monte Carlo simulations following (Santos et al. 2016, Santos et al. 2017) in Python 3.8.5. Observed means were calculated for all real pairs, then a distribution was created with all potential pairs re-assembled randomly 10,000 times with replacement. Potential pairs were calculated from the pool of available males and females for each year, based on which individuals gained dominance that year. For both questions the female was considered the chooser. Two-tailed *P* values were calculated by comparing the proportion of times the simulated means were further away from the distribution mean than the observed mean, of these only the lowest value is presented here.

3.3.3.5. Genetic mate choice: Offspring EPP likelihood

GLMMs with a binomial error structure were constructed to determine whether offspring EPP likelihood was associated with either the dominant female's or male's *TLR3* characteristics. We used the same dataset as in (Raj Pant et al. 2020), i.e., offspring produced by the dominant female in all the major seasons from 1997-2014, but only individuals from this dataset with both *TLR3* and MHC-I data were included. Dominant birds of each sex were analysed separately, i.e., 369 offspring from 130 dominant female mothers and 107 dominant males in 209 breed groups with offspring. *TLR3* genotype (*TLR3*^{AA}, *TLR3*^{AC} or *TLR3*^{CC}), Hs (continuous), MHC-I diversity (continuous), and *Ase-ua4* (present, absent) were included as fixed factors, as were other factors previously found to be important predictors of offspring EPP likelihood: offspring's natal group size (continuous) (Raj Pant et al. 2019), the age difference between the mother and her social male, and either the social male's or female's chronological age (linear term), and age^2 (quadratic term) (Raj Pant et al. 2020). All ages were measured in years. Breed year (cohort), territory, mother identity, and social male identity were initially included as random factors. To aid model convergence territory identity was removed from both models, and offspring breed year was removed from the male model, as they explained zero variance.

3.3.3.6. Genetic mate choice: Dominant male annual EGP and WGP success

Two GLMMs with a Poisson error structure were constructed to determine whether the dominant male's **A**) annual extra-group offspring (EGP success) or **B**) within-group offspring (WGP success), was associated with *TLR3* variation. The same dataset as in (Raj Pant et al. 2020) was used, i.e., annual records for all dominant males alive between 1997 and 2014 that were genetically assigned at least one offspring across the data period, except that only individuals with *TLR3* and MHC-I data were included (584 annual records from 96 dominant males). Only offspring which survived to at least 3 months of age were included in annual success measures. *TLR3* genotype ($TLR3^{AA}$, $TLR3^{AC}$ or $TLR3^{CC}$), Hs (continuous), MHC diversity (continuous), and *Ase-ua4* (present, absent) were included as fixed factors, along with the social males chronological age (linear term), and age^2 (quadratic term). As in (Raj Pant et al. 2020), either annual WGP or annual EGP were entered as a fixed factor in either the EGP, or WGP model respectively, as these factors may affect one another. All ages were measured in years. Breed year, territory, and social male identity were all included as random factors in the models

All GLMM, GLM and LMM models, were constructed using glmmTMB 0.2.3 (Brooks, Kristensen et al. 2017). Continuous factors were standardised (scaled and centred) using the package arm 1.10-1 (Gelman et al. 2018). In all models, $TLR3^{AA}$ genotype was set as the reference value. However, to gain pairwise comparisons of $TLR3^{AC}$ and $TLR3^{CC}$ individuals, we re-ran the models with $TLR3^{AC}$ genotype set as the reference value – but all other factors held constant. Collinearity among fixed effects was tested using variance inflation factors, with an upper limit of three. To confirm that there was no over- or under-dispersion, residual spatial or temporal autocorrelation in the models we used the package DHARMA 0.2.4 (Hartig 2017). Model averaging - an information-theoretic approach - was used to select plausible models using the dredge function in MUMIn 1.43.6 (Barton and Barton 2019). All biologically relevant interactions were initially included in models, and non-significant ($P > 0.05$) interactions were removed prior to model averaging to prevent over-parametrisation, and to allow assessment of first-order effects. All models within seven AICc of the top model were included in the averaged model, to get the final conditional (natural) model (Burnham et al. 2011).

3.3.3.6. Post-copulatory: Mendelian expectations

We tested whether observed offspring genotypes conformed to Mendelian expectations given their genetic parents using 272 offspring for which we had *TLR3* genotypes for the mother, genetic father, and offspring. Where both parents were homozygous, the offspring genotype cannot deviate from the expected genotype, so these cases were

excluded – i.e., $TLR3^{AA}$ fathers with $TLR3^{AA}$ mothers, and $TLR3^{CC}$ fathers with $TLR3^{CC}$ mothers (*n* = 99), leaving only offspring where one or more parents are heterozygous (*n* = 173). Second, we tested whether the father contributed a $TLR3^{A}$ or $TLR3^{C}$ allele when the mothers *TLR3* allele was fixed – i.e., $TLR3^{AC}$ fathers with $TLR3^{AA}$ or $TLR3^{CC}$ mothers (*n* = 52). For both questions, we tested whether observed genotypes were significantly different to expected genotype using χ^2 tests.

3.4. Results

3.4.1. Competitive ability: Physiological condition

$TLR3^{AC}$ males tended to be in better condition than $TLR3^{AA}$ males, but this result was not significant (*P* = 0.054, Table 3.2, Fig 3.1a), and there was no difference in condition between $TLR3^{AC}$ (Table 3.2) and $TLR3^{CC}$ males (*P* = 0.466). $TLR3^{AC}$ females were in better body condition than $TLR3^{CC}$ females (β ± SE = -0.68 ± 0.23, z = 2.93, *P* = 0.003, Fig 1b) and tended to be in better condition than $TLR3^{AA}$ females, but this result was not significant (*P* = 0.062). $TLR3^{AA}$ females also tended to be in better condition than TLR^{CC} females, but this result was also not significant (*P* = 0.073, Table 2, Fig 1b). *TLR3* genotype was not correlated with tarsus size in either sex (Table 3.2).

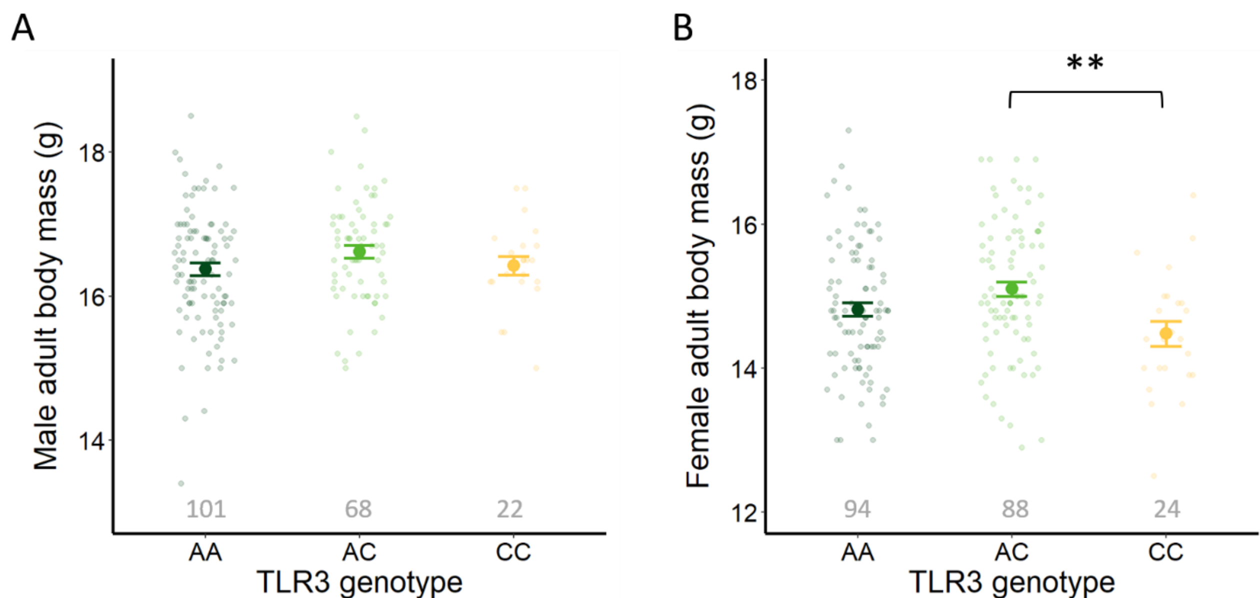


Fig 3.1: The effect of *TLR3* genotype on body condition (mass) **in A**) adult male (*n* = 97), and **B**) adult female (*n* = 120) Seychelles warbler. In bold are raw means ± SE, with raw data points shown in fade. *TLR3* genotype indicated by colour ($TLR3^{AA}$ = dark green, $TLR3^{AC}$ = light green, $TLR3^{CC}$ = yellow), with associated sample sizes in grey above the x axis. ** *P* < 0.01.

MHC diverse males were larger in size, whereas MHC diverse females were smaller in size (Table 3.2). Males with at least one copy of the *Ase-ua4* alleles were also larger in size, but presence of *Ase-ua4* did not predict female size (Table 3.2), and there was no effect of MHC diversity, or presence of *Ase-ua4* on condition in either sex (Table 3.2).

Table 3.2: Predictors of **A)** body mass and **B**) size for adult male and female Seychelles warblers. LMMs were used to generate conditional model-averaged estimates (β), their standard error (SE), z value, *P* value, and relative importance (ω) are shown for all predictors featuring in the top model set ($\Delta AIC_c \leq 7$). Bird identity and catch year were included as random factors. All continuous factors were standardised. Estimates are in reference to catch period = AM, and *TLR3* genotype = $TLR3^{AA}$. Significant terms are in bold and underlined. *** $P < 0.001$, * $P < 0.05$, $P < 0.1$.

Model	Factor	Male (97 males, 191 catches)						Female (120 females, 206 catches)					
		ω	β	SE	z	*P*		ω	β	SE	z	*P*	
A) Condition	(Intercept)		**16.32**	**0.13**	**128.64**	**<0.001**	***		**14.69**	**0.14**	**106.61**	**<0.001**	***
	Catch Period	**0.84**	**0.26**	**0.11**	**2.29**	**0.022**	*	**1.00**	**0.41**	**0.12**	**3.36**	**0.001**	***
	Tarsus	**0.72**	**0.24**	**0.12**	**1.97**	**0.049**	*	0.69	0.25	0.13	1.93	0.053	.
	Age	0.25	-0.06	0.11	0.48	0.632		0.26	0.07	0.14	0.48	0.630	
	Hs	0.24	0.05	0.14	0.34	0.737		0.24	0.03	0.14	0.20	0.838	
	MHC Diversity	0.26	-0.08	0.15	0.53	0.595		0.35	-0.14	0.14	0.98	0.326	
	Ase-ua4	0.25	0.11	0.23	0.45	0.650		0.30	0.14	0.17	0.79	0.432	
	TLR3: ^{AC}	0.40	0.29	0.15	1.93	0.054	.	0.95	0.27	0.14	1.88	0.060	.
	TLR3: ^{CC}		0.11	0.23	0.47	0.642			-0.40	0.23	1.76	0.078	.
B) Size	(Intercept)		**25.79**	**0.08**	**313.31**	**<0.001**	***		**24.17**	**0.06**	**386.20**	**<0.001**	***
	Catch Period	0.22	0.00	0.08	0.02	0.981		0.58	0.11	0.07	1.65	0.099	.
	Age	0.24	-0.03	0.10	0.26	0.794		0.25	-0.02	0.07	0.33	0.741	
	Hs	0.31	0.10	0.12	0.85	0.398		0.38	0.10	0.10	1.07	0.283	
	MHC Diversity	**0.80**	**0.27**	**0.13**	**2.14**	**0.032**	*	**0.82**	**-0.22**	**0.10**	**2.19**	**0.028**	*
	Ase-ua4	**0.83**	**0.42**	**0.19**	**2.23**	**0.026**	*	0.24	-0.03	0.13	0.21	0.831	
	TLR3: ^{AC}	0.18	-0.01	0.13	0.07	0.945		0.07	0.00	0.11	0.04	0.970	
	TLR3: ^{CC}		-0.26	0.20	1.28	0.201			0.07	0.17	0.42	0.674	

3.4.2. Social mate choice: Likelihood of dominance gain likelihood

Overall, the likelihood of gaining a dominance position before two years of age was not directly related to *TLR3* genotype in either males or females (Table 3.3). However, in males - but not females - natal group size was related to dominance gain, with individuals hatched from bigger groups more likely to gain dominance before two years of age (Table 3.3). Furthermore, there was an interaction between male *TLR3* genotype and natal group size; $TLR3^{AC}$ and $TLR3^{CC}$ males were more likely to gain dominance as natal group size decreased, whereas $TLR3^{AA}$ males were more likely to gain dominance with increasing natal group size (Table 3.3; Fig 3.2). There was no difference between group size interacted with $TLR3^{AC}$ and $TLR3^{CC}$ males ($P = 0.339$). There was no interaction between natal group size and *TLR3* genotype in females ($P > 0.1$).

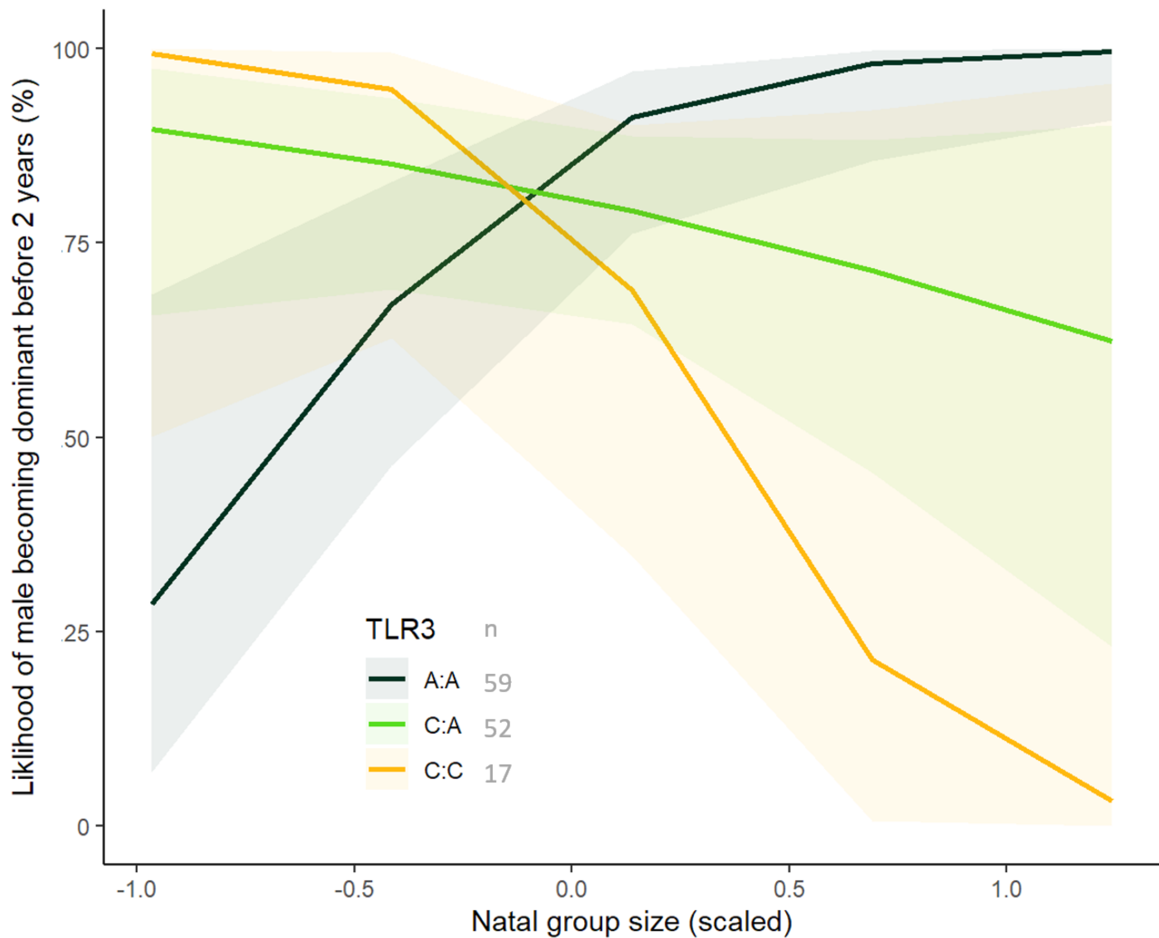


Fig 3.2: Effect of *TLR3* genotype and natal group size on likelihood of gaining dominance by 2 years of age for male Seychelles warblers. The fit lines for *TLR3* genotype (*TLR3^{AA}* = dark green, *TLR3^{AC}* = light green, *TLR3^{CC}* =yellow) are derived from conditional model averaged GLMMs (see Table 3), and the shaded areas are the 95% confidence intervals. Sample sizes (number of individuals) are indicated in grey in the key.

There was no significant effect of individual Hs, MHC diversity, presence of *Ase-ua4*, or natal TQ on dominance gain in males or females (Table 3.3). In females, season hatched was marginally related to dominance gain (Table 3.3). Females hatched in the major season were more likely to become dominant by 2 years of age than those hatched in the minor season, the same trend was observed for males, but was not significant (*P* = 0.065).

Table 3.3: Likelihood of gaining a dominant breeding position before two years of age for male and female Seychelles warblers in relation to *TLR3* genotype. A GLMM (females) and a GLM (males) were used to generate conditional model-averaged estimates (β), their standard error (SE), z value, *P* value, and relative importance (ω) are shown for all predictors featuring in the top model set ($\Delta AIC_c \leq 7$). Hatch year is included as a random factor in female models. All continuous factors were standardised. Estimates are in reference to: Season hatched = Minor, *Ase-ua4* = present, and *TLR3* genotype = $TLR3^{AA}$. Significant terms are in bold and underlined. *** $P < 0.001$, ** $P < 0.01$, * $P < 0.05$, . $P < 0.1$.

Factor	Male (n = 128)						Female (n = 142)					
	ω	β	SE	z	P		ω	β	SE	z	P	
(Intercept)		**1.79**	**0.49**	**3.59**	**<0.001**	***		**0.63**	**0.26**	**2.43**	**0.015**	*
Season hatched	0.65	-1.04	0.56	1.84	0.065	.	**0.78**	**-0.87**	**0.42**	**2.06**	**0.039**	*
Natal group size	**0.91**	**2.83**	**1.17**	**2.40**	**0.016**	*	0.35	-0.38	0.37	1.01	0.313	
Natal TQ	0.27	-0.37	0.53	0.69	0.492		0.49	0.62	0.49	1.27	0.204	
Hs	0.26	-0.29	0.45	0.63	0.532		0.70	-0.72	0.38	1.85	0.064	.
MHC Diversity	0.72	1.05	0.53	1.96	0.051	.	0.73	-0.75	0.39	1.93	0.054	.
Ase-ua4	-	-	-	-	-		0.28	0.32	0.47	0.67	0.504	
TLR3 : AC	0.87	-0.50	0.58	0.84	0.401		0.12	-0.18	0.40	0.45	0.653	
TLR3 : CC		-0.59	0.76	0.77	0.442			-0.60	0.58	1.02	0.307	
Natal group size * *TLR3* : AC	**0.87**	**-3.61**	**1.24**	**2.88**	**0.004**	**	-	-	-	-	-	
Natal group size * *TLR3* : CC		**-6.18**	**2.84**	**2.16**	**0.031**	*	-	-	-	-	-	

3.4.3. Social mate choice: Assortative mating and/or directional mate choice

There was some evidence of assortative mating relative to *TLR3* genotype. Observed social pairs shared the same *TLR3* alleles more often than simulated social pairs (mean ± SEM pair *TLR3* similarity for: observed pairs = 0.69 ± 0.02, *n* = 364, and simulated pairs = 0.66 ± 0.01, *n* = 7890; *P* = 0.026; Fig 3.3a).

Furthermore, $TLR3^{AA}$ females paired with males with a greater number of $TLR3^{A}$ alleles than expected by chance (mean ± SEM male $TLR3^{A}$ alleles for: observed pairs = 1.50 ± 0.05, *n* = 174, and simulated pairs = 1.39 ± < 0.01, *n* = 3720; *P* = 0.026; Fig 3.3b). In contrast, $TLR3^{AC}$ females paired with males with a reduced number of $TLR3^{A}$ alleles than expected (observed pairs = 1.29 ± 0.06, *n* = 146, and simulated pairs = 1.36 ± < 0.01, *n* = 3233; *P* = 0.043; Fig 3.3c). There was no difference between observed and simulated pairings for $TLR3^{CC}$ female (observed pairs = 1.36 ± 0.11, *n* = 44, and simulated pairs = 1.36 ± < 0.01, *n* = 937; *P* = 0.628; Fig 3.3d)

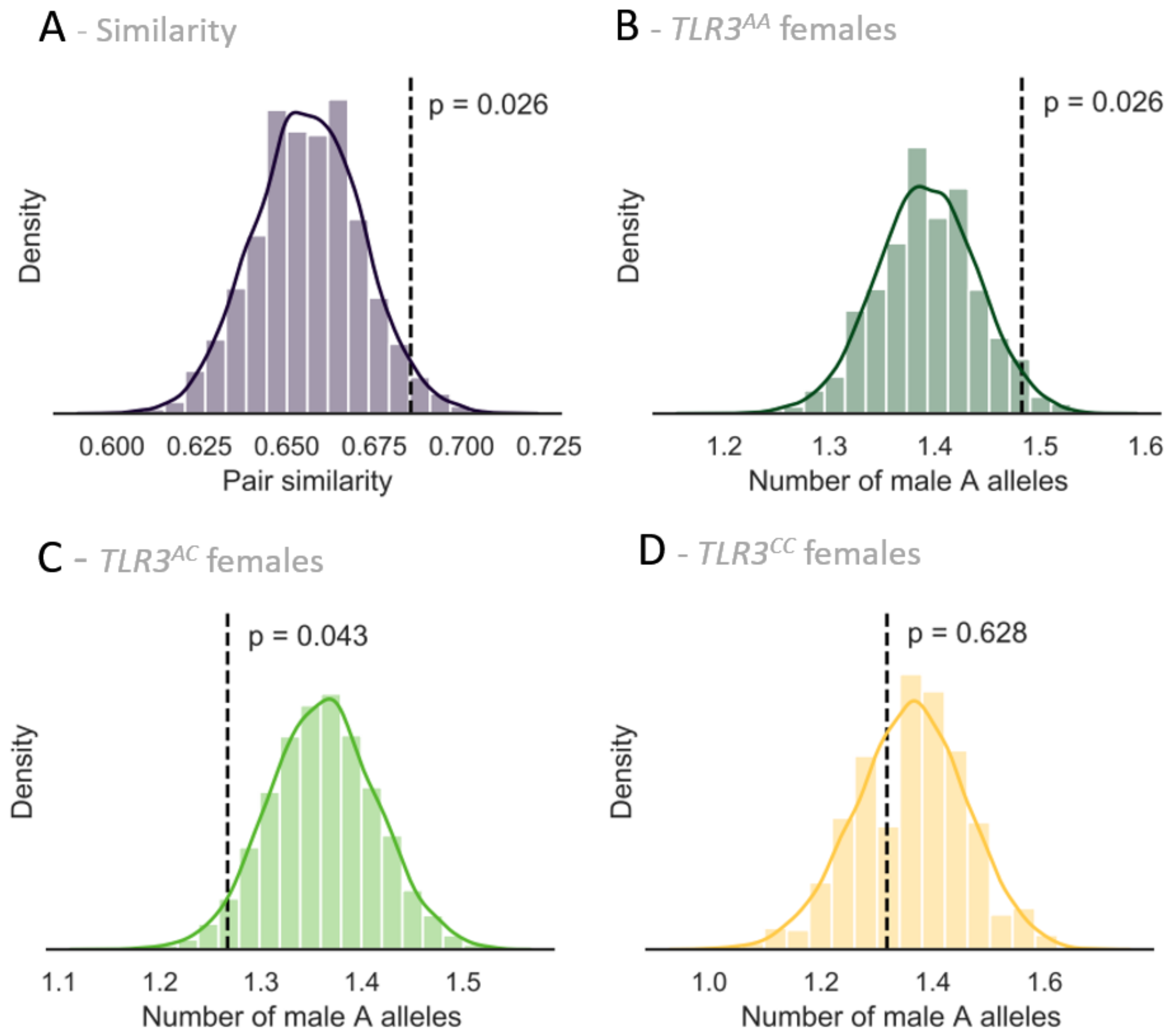


Fig 3.3: The association between *TLR3* genotype and social mate-choice in Seychelles warblers. **A**) Assortative social mating based on pair *TLR3* similarity (purple), or directional female social pair preference of **B**) $TLR3^{AA}$ females (dark green), **C**) $TLR3^{AC}$ females (light green), **D**) $TLR3^{CC}$ females (yellow), based on the number of $TLR3^{A}$ alleles in the social male. Frequency distributions of *TLR3* genotypes in simulated social pairings are generated by Monte Carlo simulations of 10,000 randomised male–female pairings for the Cousin population. Dotted line indicates the mean value of the observed social pairings. P value are calculated by comparing the proportion of times the observed mean was further away from the distribution mean, compared to the simulated means.

3.4.4. Genetic mate choice: Offspring EPP likelihood

Dominant male genetic characteristics (*TLR3* genotype, Hs, *Ase-ua4*, or MHC diversity) had no effect on the likelihood of them being cuckolded, neither did their age (linear or quadratic; Table 3.5; Fig 3.4a). As previously shown using this data set (Raj Pant, Hammers et al. 2020) dominant males were more likely to be cuckolded when the age

difference between them and the dominant female was greater, and they were in bigger groups (Table 3.5).

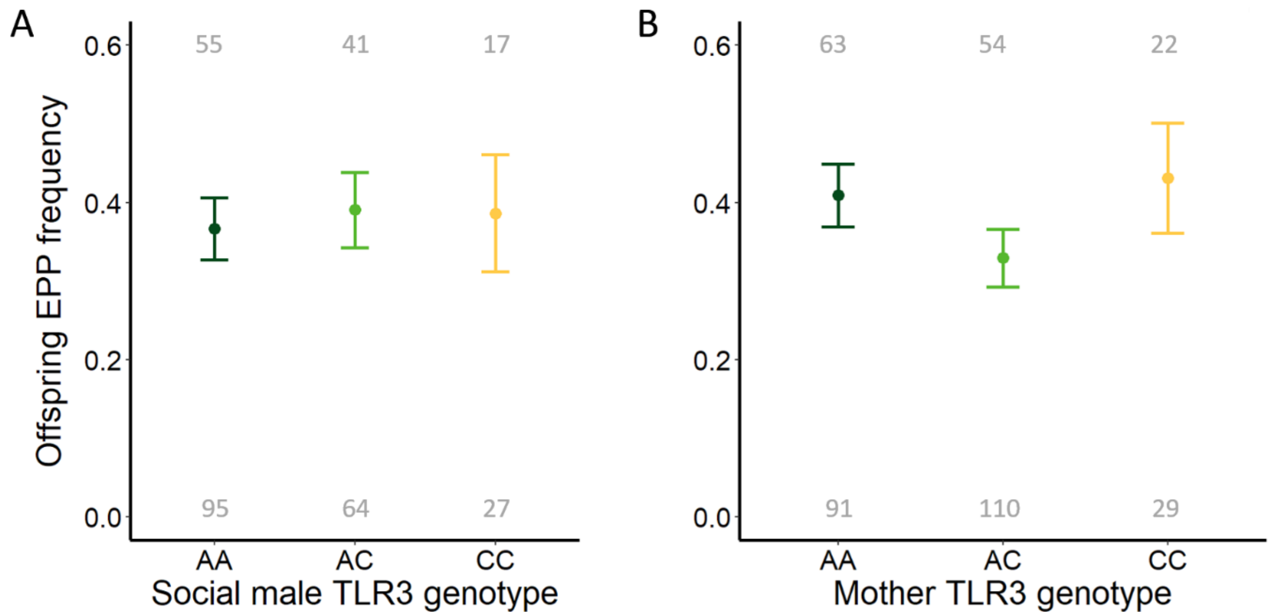


Fig 3.4: The likelihood of offspring being extra-pair in relation to the parent's *TLR3* genotype in the Seychelles warbler: **A**) social males (n = 107), **B**) mothers (n = 137). Data are raw means ± SE, with *TLR3* genotype indicated by colour ($TLR3^{AA}$ = dark green, $TLR3^{AC}$ = light green, $TLR3^{CC}$ =yellow), with associated sample sizes (offspring number) in grey just above the x axis for within-pair offspring, and at the top for extra-pair offspring. No genotypes were significantly different from another.

Table 3.5: Offspring extra-pair paternity likelihood based on the social male and mother characteristics in the Seychelles warbler in relation to *TLR3* genotype. GLMMs were used to generate conditional model-averaged estimates (β), their standard error (SE), z value, *P* value, and relative importance (ω) are shown for all predictors in the top model set ($\Delta AIC_c \leq 7$). All continuous factors were standardised. Mother and social male ID are included as random factors in both models, and breed year is additionally included in the mother model. Estimates are in reference to *Ase-ua4* = present, and *TLR3* genotype = $TLR3^{AA}$. Significant terms are in bold and underlined. ** $P < 0.01$, * $P < 0.05$, . $P < 0.1$.

Factor	Social male (107 social males, 299 offspring)						Mother (137 mothers, 369 offspring)					
	ω	β	SE	z	*P*		ω	β	SE	z	*P*	
(Intercept)		**-0.59**	**0.18**	**3.18**	**0.001**	**		**-0.55**	**0.25**	**2.23**	**0.026**	*
Age	0.32	-0.65	1.01	0.64	0.522		0.72	2.19	1.25	1.75	0.080	.
Age^2	0.28	0.55	1.02	0.54	0.592		0.80	-2.24	1.25	1.79	0.073	.
Social pair age difference	**0.91**	**0.89**	**0.39**	**2.29**	**0.022**	*	**0.80**	**0.75**	**0.36**	**2.04**	**0.041**	*
Group size	**0.77**	**0.61**	**0.31**	**1.97**	**0.049**	*	0.55	0.41	0.26	1.56	0.119	
Hs	0.28	-0.25	0.34	0.75	0.453		0.24	-0.04	0.28	0.13	0.899	
MHC Diversity	0.22	0.02	0.34	0.06	0.955		**0.84**	**0.64**	**0.29**	**2.16**	**0.030**	*
Ase-ua4	0.23	0.18	0.47	0.38	0.705		0.52	0.51	0.34	1.52	0.130	
TLR3 : ^{AC}	0.05	0.12	0.38	0.31	0.758		0.51	-0.58	0.31	1.88	0.061	.
TLR3 : ^{CC}		0.13	0.48	0.27	0.786			0.08	0.44	0.19	0.853	

TLR3^{AC} dominant mothers were marginally less likely to produce offspring sired by extra group males than *TLR3*^{AA} females, but not *TLR3*^{CC} females, but this result was not significant (*P* = 0.061 and *P* = 0.135 respectively), and there was no difference between *TLR3*^{AA} and *TLR3*^{CC} mothers (Table 3.5; Fig 3.4b). The proportion of offspring produced via EPP was positively correlated with the dominant mother's MHC diversity (Table 3.5). But the proportion of offspring produced via EPP was not affected by the dominant mothers' Hs or presence of *Ase-ua4*, and there was no effect of group size, or age (linear or quadratic; Table 3.5). Dominant mothers were more likely to gain EPP when there was a greater age difference between themselves and their social partner (Table 3.5).

3.4.5. Dominant male annual EGP and WGP success

There was no significant effect of dominant male *TLR3* genotype on annual EGP success (Table 3.6; Fig 3.5a), but *TLR3*^{CC} males had reduced annual WGP success compared to *TLR3*^{AA} (Table 3.6; Fig 3.5b), but not *TLR3*^{AC} males (*P* = 0.184), and there was no difference between *TLR3*^{AA} and *TLR3*^{AC} males. This corresponded to *TLR3*^{CC} males producing on average 33% fewer within-group offspring per year (mean ± SEM: 0.20 ± 0.04) than *TLR3*^{AA} (mean ± SEM: 0.32 ± 0.03), or *TLR3*^{AC} (mean ± SEM: 0.29 ± 0.04) males. There was no effect of *H*s, *Ase-ua4*, or MHC diversity on annual EGP or WGP success (Table 3.6). Additionally, annual EPP and WPP success were not influenced by male age (linear or quadratic; Table 3.6).

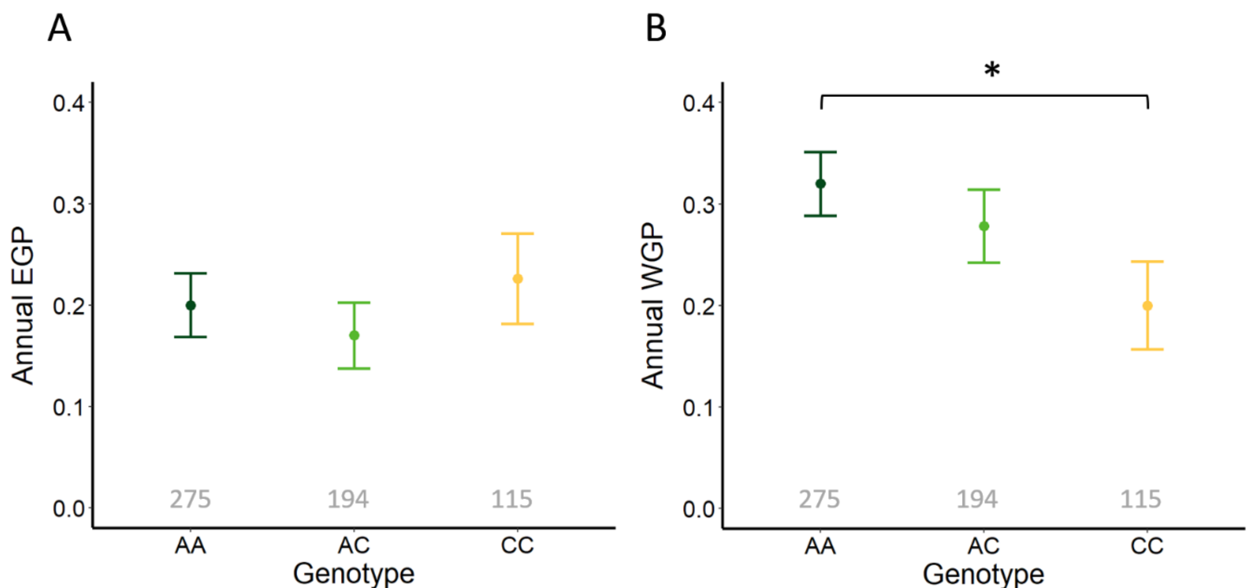


Fig 3.5: Annual **A**) extra-group paternity (EGP), and **B**) within-group paternity (WGP) gained by dominant male Seychelles warblers in relation to their TLR3 genotype (n= 96). Data are raw means and standard errors, with TLR3 genotype indicated by colour (*TLR3*^{AA} = dark green, *TLR3*^{AC} = light green, *TLR3*^{CC} = yellow), and associated sample sizes (years) in grey above the x axis. * *P* <0.05.

Table 3.6: Annual extra-group (EGP) and within-group (WGP) paternity gained by dominant male Seychelles warblers in relation to *TLR3* genotype. GLMMs were used to generate conditional model-averaged estimates (β), their standard error (SE), z value, *P* value, and relative importance (ω) are shown for all predictors featuring in the top model set ($\Delta AIC_c \leq 7$). Years refers to total number of years included for all males. All continuous factors were standardised. Breed year, territory and individual ID are included as random factors. Estimates are in reference to *Ase-ua4* = present, and *TLR3* genotype = $TLR3^{AA}$. Significant terms are in bold and underlined. *** $P < 0.001$, * $P < 0.05$.

Factors	Annual EGP (96 males, 584 years)						Annual WGP (96 males, 584 years)					
	ω	β	SE	z	*P*		ω	β	SE	z	*P*	
(Intercept)		**-1.90**	**0.18**	**10.40**	**<0.001**	***		**-1.21**	**0.12**	**9.69**	**<0.001**	***
Age	0.64	1.63	1.03	1.58	0.114		0.35	0.32	0.64	0.50	0.617	
Age^2	0.62	-1.62	1.03	1.57	0.116		0.42	-0.49	0.59	0.83	0.408	
Annual EGP	-	-	-	-	-		0.25	0.03	0.15	0.20	0.842	
Annual WGP	0.23	0.03	0.19	0.14	0.891		-					
Hs	0.24	0.01	0.24	0.06	0.955		0.25	-0.01	0.17	0.07	0.943	
MHC Diversity	0.40	0.26	0.24	1.12	0.261		0.25	-0.01	0.16	0.05	0.96	
Ase-ua4	0.25	0.14	0.32	0.44	0.663		0.25	0.01	0.22	0.03	0.979	
TLR3 : AC	0.15	-0.18	0.28	0.66	0.512		0.56	-0.14	0.18	0.79	0.428	
TLR3 : CC		0.16	0.30	0.54	0.589			**-0.48**	**0.24**	**2.03**	**0.042**	*

3.4.6. Post-copulatory: Mendelian expectations

There was no significant difference between observed and expected offspring genotype based on their combined parental genotypes ($n = 173$, df = 1, $\chi^2 = 1.19$, $P = 0.275$; Fig 3.6). However, when focusing on the inheritance of paternal TLR3 alleles there was a difference, with offspring from $TLR3^{AC}$ fathers less likely to inherit the $TLR3^{C}$ allele, and more likely to inherit the $TLR3^{A}$ allele than expected ($n = 52$, df = 1, $\chi^2 = 4.96$, $P = 0.026$; Fig 3.7).

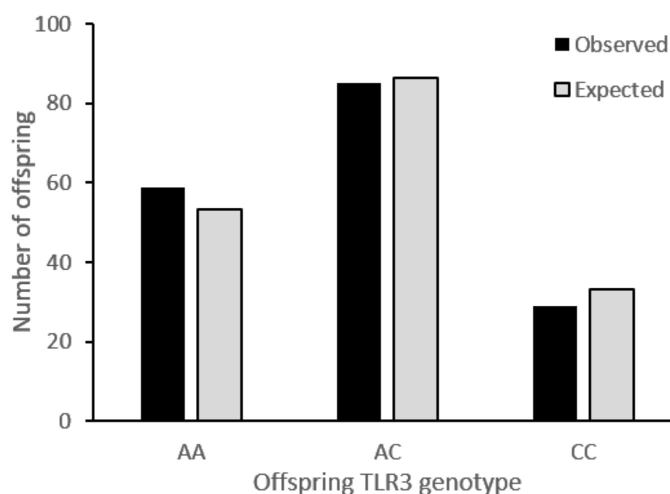


Fig 3.6: Observed and expected values of offspring's *TLR3* genotype based on parental *TLR3* genotypes, when at least one parent is *TLR3* heterozygous ($n = 173$) in the Seychelles warbler. Observed values are shown in black and expected values in grey.

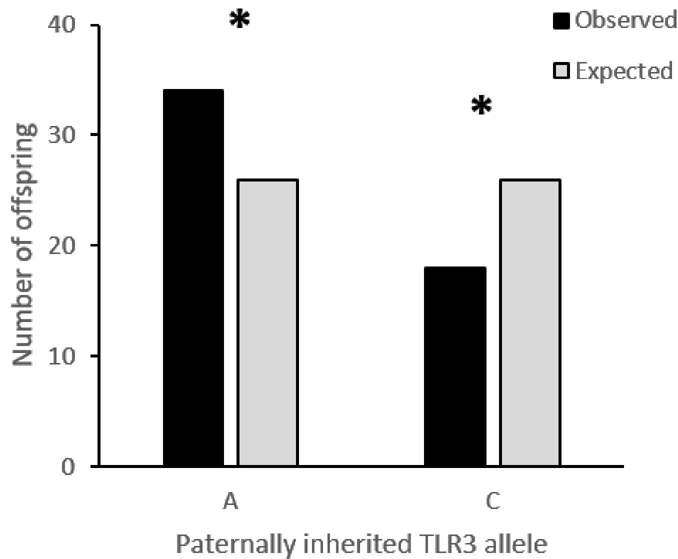


Fig 3.7: Observed and expected values of whether offspring paternally inherit a $TLR3^A$ or $TLR3^C$ allele when genetic father is heterozygous ($TLR3^{AC}$), and the mother is homozygous ($TLR3^{AA}$ or $TLR3^{CC}$) (*n* = 52) in the Seychelles warbler. Observed values are shown in black and expected values in grey. * *P* <0.05.

3.5. Discussion

Our result show that *TLR3* genotype was associated with a difference in body condition (but not in skeletal size); $TLR3^{AC}$ adult females were in better condition. Though there was no difference between *TLR3* genotypes in the likelihood of gaining a dominant breeding position early in life, we did detect evidence of a link between *TLR3* genotype and social mate choice; $TLR3^{AA}$ females were more likely to pair with $TLR3^{AA}$ males, and $TLR3^{AC}$ females were less likely to pair with $TLR3^{AA}$ males. However, there was no evidence of extra-pair mate choice based on male *TLR3* genotypes, although $TLR3^{CC}$ males had reduced within-group paternity. Lastly, while offspring genotypes overall conformed to Mendelian expectations, we did find that when the genetic father was heterozygous for the *TLR3* SNP, offspring were more likely to inherit the $TLR3^A$ allele from the father than the $TLR3^C$ allele.

3.5.1. TLR3 *associated condition dependent cues and dominance gain*

There were no skeletal size differences between Seychelles warblers with the different *TLR3* genotypes. In contrast, $TLR3^{AC}$ heterozygous individuals had a non-significant tendency to have greater body mass (controlling for skeletal size) than did *TLR3* homozygous individuals. We predicted that $TLR3^{CC}$ individuals of both sexes would be in worse condition than their counterparts (based on the logic of this genotype having greater mortality), and likewise that $TLR3^{AA}$ males – which had increased lifetime reproductive success than other genotypes - would be in better condition. As expected,

TLR3^{CC} females were lighter than other *TLR3* genotypes, suggesting that there could be condition-dependent consequences at this *TLR3* locus. However, there was no difference in the condition of *TLR3*^{CC} males with any of the different *TLR3* genotypes. It must be noted that our power to resolve such differences was limited because the sample size of *TLR3*^{CC} individuals was low (22 measurements, from 12 individuals) due to the lower frequency of the *TLR3*^{C} allele and greater early mortality of *TLR3*^{CC} individuals in this population (see Chapter 2). Furthermore, relative mass is only one aspect of condition, and it could be that *TLR3* genotype is linked to other differences in condition or vigour. Future studies could expand the data set and incorporate status and reproductive state into the analysis, assess different metrics of condition (e.g., telomere length (Bebbington et al. 2016), and oxidative stress (van de Crommenacker et al. 2017), or include potentially secondary sexual characteristics and signals such as song (Saino et al. 1997).

We found no difference in the likelihood of gaining dominance before two years of age relative to *TLR3* genotype for either sex, but there was an interaction between *TLR3* genotype and natal group size in males. *TLR3*^{AA} males were more likely to gain a dominant position when from larger natal groups, whilst the opposite was true for *TLR3*^{CC} males - with heterozygous *TLR3*^{AC} males showing an intermediate pattern. While this result could partly be due to few individuals originating from very big or small natal groups, it raises some intriguing possibilities. There are differences in strategies and costs of dispersal between the sexes (Eikenaar et al. 2008b, Kingma et al. 2016a, Groenewoud et al. 2018), and intra-sexual competition for territories may be stronger in males (Eikenaar et al. 2009). Male Seychelles warblers suffer greater mortality from dispersal than do females (Kingma et al. 2017). Likewise, individuals are more likely to become floaters when natal group sizes are larger, and floaters suffer from increased mortality (Kingma et al. 2016a). And so, condition could be an important determinant of male survival and successful dominance gain post dispersal. As *TLR3*^{AA} males exhibited greatest fitness (Chapter 2), they may be better able to survive and successfully gain dominance after dispersing from the natal territory.

3.5.2. TLR3 *mediated social mate choice through assortative mating*

Our result provided evidence of assortative social pairing in respect to the *TLR3* locus, which could potentially explain the previous finding of a deficiency of heterozygotes in young individuals (Chapter 2). This increased social pair similarity could be driven by directional social pairing, rather than assortative mating *per se*. Indeed, we found that *TLR3*^{AA} females were more likely to pair with *TLR3*^{AA} males, while *TLR3*^{AC} females were

less likely to pair with *TLR3*^{AA} males – and presumably paired more often with *TLR3*^{AC} males than *TLR3*^{CC} males due to the greater number of *TLR3*^{AC} individuals in the population. There was no detected pattern of pairing relative to *TLR3*^{CC} females, though this could be a power issue, due to the small number of adults of this genotype in the population - and thus our sample. In socially monogamous species, particularly those with long partnerships, we would expect both sexes to be selective about their partner and be mutually choosy (Hooper and Miller 2008, Baldauf et al. 2009). Individuals of both sexes that are in better condition should be more discriminatory and pair together - as they would both be better at defending and holding high quality territories. While individuals in worse condition would be less selective. This may explain the lack of directional preference in *TLR3*^{CC} females (Fawcett and Johnstone 2003). Although mutual mate choice, whether it be based on direct or indirect mechanisms, is an obvious possible cause of the observed bias in *TLR3*^{AA} pairings, it is possible that intra-sexual competition could be causing these mating patterns. In studies on free living organisms, it can be very difficult to separate competition from mate choice, with the same signals or traits driving both mechanisms and feeding back on each other, i.e., bigger males may be more likely to win against smaller males, and also be more likely to be preferred as mates (Wong and Candolin 2005).

Evidence for a genetic basis to social or genetic assortative mating patterns is relatively uncommon in vertebrates, with the examples that do exist mainly based on measures of heterozygosity (Garcia-Navas et al. 2009, Ortego et al. 2009), or MHC variation (Bonneaud et al. 2006, Sin et al. 2015, Santos et al. 2016, Santos et al. 2017). Typically, disassortative rather than assortative mating based on genetic characteristics is predicted, as this would reduce inbreeding likelihood and/or increase offspring heterozygosity, thus increasing offspring fitness (reviewed in Kamiya et al. 2014)). However, many studies have found no genetic effects either way (Westerdahl 2004, Sepil et al. 2015, Indykiewicz et al. 2017). Assortative mating based on phenotypic traits has been well studied (Jiang et al. 2013, Wang et al. 2019), although whether this evolves due to selection or constraint is debated. Whilst there is evidence for MHC-mediated olfactory (Grieves et al. 2019) and visual cues (Dunn et al. 2013), acting as sexually selected traits in passerines, *TLR3*-mediated cues are more likely to be condition dependent - due to *TLR3*'s role in resistance to viral infection (Barton 2007). While we found little difference in mass or size in relation to *TLR3* genotype in this study, there are many different potential cues which we did not measure (see Indykiewicz et al. 2017).

Assortative mating patterns could potentially be explained by passive mechanisms such as active inbreeding or reduced natal dispersal (as discussed in Slade et al. 2019) but these mechanisms are unlikely to be driving the *TLR3* dependent pairing pattern we find in the present study. While inbreeding does occur in the Seychelles warbler (Eikenaar et al. 2008a), there is no evidence for active inbreeding in social pair formation (Richardson et al. 2004, Wright et al. 2016). Likewise, there is no evidence for assortative social pair mate choice based on MHC variation (Richardson et al. 2005), which we would expect if our result is due to inbreeding. Unlike the more common $TLR3^{AA}$ and $TLR3^{AC}$ genotypes, $TLR3^{CC}$ individuals exhibited random mating. We also found no effect of *TLR3* genotype on the overall likelihood of gaining dominance, and that available frequency of *TLR3* genotypes in the paired population was similar to that of the larger adult population (including subordinate and floater birds), suggesting inclusion into a pair *per se* was not biased towards certain *TLR3* genotypes, but rather who you paired with.

3.5.3. TLR3 dependent genetic mate choice

We found no evidence of dominant males being more likely to be cuckolded, or of gaining more EPP in relation to their *TLR3* genotype. The lack of association between male *TLR3* genotype and EPP suggests that genetic mate choice is not driving the differential reproductive success found previously (Chapter 2). Although we found more MHC diverse females to be more likely to produce EPP offspring, we didn't find any effect of male MHC diversity on offspring EPP, despite previous evidence that male based MHC-dependent EPP takes place (Richardson, Komdeur et al. 2005), though due to sample size limitations we were not able to directly compare the genotype of the social male being cuckolded and the chosen genetic male. There was a marginal (but non-significant) result suggesting that $TLR3^{AC}$ females could be less likely to produce EPP offspring than other genotypes. Based on the assortative pairing results, we would expect that to result in $TLR3^{AC}$ dominant males being less likely to lose within pair-paternity, however this was not the case. As these females were in better condition (greater body mass) than $TLR3^{AA}$ or $TLR3^{CC}$ females, this seems odd. Females in better physical condition are usually more attractive to other males, or better able to avoid mate-guarding, both of which would result in increased EPP offspring (Gowaty 1996, Hoi-Leitner et al. 1999).

In contrast to levels of EPP, we did find that $TLR3^{CC}$ males had reduced annual WGP compared to $TLR3^{AA}$ or $TLR3^{AC}$ males. While this can explain the reduced LRS of $TLR3^{CC}$ males, we also previously found that $TLR3^{AC}$ males had reduced reproductive success (Chapter 2). Given there appears to be no bias against gaining a social or genetic mate

for *TLR3*^{CC} males, and they do not appear more likely to lose paternity, reduced WGP is likely due to post-copulatory rather than pre-copulatory mechanisms. This could be via poorer parental care, resource allocation, or through increased early life mortality of offspring (which we know selects against *TLR3*^{CC} offspring). Incorporating measures of territory quality, male nest guarding, and provisioning rate with *TLR3* genotype information could provide further resolution on if, and why, *TLR3* dependent post-copulatory selective mortality occurs and explain the lower reproductive rate of *TLR3*^{AC} males. If *TLR3*^{CC} offspring mortality is contributing to reduced WGP, this would, to a lesser extent, also affect *TLR3*^{AC} males that may also produce such offspring depending on their partner's *TLR3* genotype. Indeed, although not significant, our results suggest that *TLR3*^{AC} males may tend to do slightly worse in gaining dominance, EGP, and WGP, and are cuckolded more often, than *TLR3*^{AA} males. Although separately, there was not sufficient power to detect *TLR3* bias, perhaps together the combined accumulation of slightly deleterious effects could cause the overall reduction in male *TLR3*^{AC} individuals' reproductive success.

3.5.4. Mendelian inheritance and post-copulatory selection

Offspring *TLR3* genotype conformed to Mendelian expectations when considering all parental genotypes. However, we then focused on offspring where the genetic father was *TLR3*^{AC} (heterozygous) and the mother was homozygous. In these cases, the offspring could inherit either allele from the male, but the maternal allele was fixed, so we could resolve which allele was paternally inherited. We found offspring were significantly more likely to inherit the *TLR3*^{A} allele rather than the *TLR3*^{C} allele from the father. *TLR3*^{AA} females are more common in the population than *TLR3*^{CC} females, consequently, this bias in the inheritance of paternal *TLR3*^{A} alleles could contribute to the heterozygote deficiency in offspring that we previously found in this population (Chapter 2). Repeating this test for heterozygous females showed no bias in the inheritance of maternal allele, which would be expected if post-fertilization processes were driving this skew. This result, and that the overall offspring genotypes conform to Mendelian expectations, indicates that bias in the inheritance of male alleles is not due to sperm-egg incompatibility, or early embryo/offspring mortality. Instead, the bias in paternal allele inheritance we observe is likely due to pre-, rather than post-fertilization processes, either through cryptic mate-choice (Eberhard 1996) or sperm competition (Birkhead and Møller 1998).

There is increasing appreciation of the role within-male sperm competition may play in driving evolutionary processes (reviewed in Sutter and Immler 2020). Within-male

variation in sperm can be expressed via differential longevity, performance, and morphology (Holt and Look 2004) and have consequences for offspring development and survival (Immler et al. 2014, Alavioon et al. 2017). There is some evidence that sperm variation can be correlated with genetic variability (Borowsky et al. 2019), and that this can result in balancing selection at the haploid level (Joseph and Kirkpatrick 2004). There is also mounting evidence that the MHC can play a role in gamete-mediated mate choice (Reusch et al. 2001, Yeates et al. 2009, Ziegler et al. 2010, Løvlie et al. 2013, Kekäläinen and Evans 2018). However, why there would be differential inheritance of the *TLR3* molecule – which recognises viral dsRNA intracellulary (Hutchens et al. 2008, Brownlie and Allan 2011 (Chen et al. 2013)) – is unclear. Though in mice *(Mus musculus*) *TLR3* expression has been linked to infertility through severe inflammation in testicular cells (reviewed in Dutta, Sengupta et al. 2020). In the Seychelles warbler, testing for differences in sperm quality between *TLR3* genotypes could be a valuable avenue for future research, with various methods to collect sperm from passerines existing (Wolfson 1952, Immler and Birkhead 2005). Likewise, the reproductive microbiome can be an important driver of sexual selection (reviewed in Rowe et al. 2020). Specifically, sexually transmitted infections - including viruses - can damage sperm, resulting in differences in success (Wigby et al. 2019), and are widespread in avian species (Sheldon 1993). Although difficult (reviewed in Garmaeva et al. 2019), conducting virome analysis of the reproductive tract could provide insight on what is driving this difference in paternal *TLR3* allele inheritance.

3.5.5. Conclusion

Our results provide evidence that both pre- and post-copulatory sexual selection are associated with *TLR3* variation in the Seychelles warblers, a population where we previously found evidence of strong positive selection on variation at this locus (Chapter 2). Unidirectional assortative mating appears to be occurring in terms of social pairs (precopulatory mechanism), while a bias in the inheritance of a specific *TLR3* haplotype from the heterozygous genetic father was also observed (post copulatory mechanism). These mechanisms could explain the offspring heterozygote deficiency previously observed in this population. We also found $TLR3^{CC}$ males had reduced WGP, this combined with their reduced survival can explain their lower LRS compared to other genotypes. These results outline that multiple mechanisms of selection, including both pre- and post-copulatory sexual selection, as well as natural selection may act to affect contemporary evolution within a natural population.

3.6. References

Acevedo-Whitehouse, K. and Cunningham, A. A. (2006). Is MHC enough for understanding wildlife immunogenetics? *Trends in Ecology & Evolution.* **21**(8): 433-438.

Akira, S., Uematsu, S. and Takeuchi, O. (2006). Pathogen Recognition and Innate Immunity. *Cell.* **124**(4): 783-801.

Alavioon, G., Hotzy, C., Nakhro, K., Rudolf, S., Scofield, D. G., Zajitschek, S., . . . Immler, S. (2017). Haploid selection within a single ejaculate increases offspring fitness. *Proceedings of the National Academy of Sciences.* **114**(30): 8053-8058.

Alcaide, M. and Edwards, S. V. (2011). Molecular evolution of the toll-like receptor multigene family in birds. *Molecular Biology and Evolution.* **28**(5): 1703-1715.

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

Antonides, J., Mathur, S., Sundaram, M., Ricklefs, R. and DeWoody, A. J. (2019). Immunogenetic response of the bananaquit in the face of malarial parasites. *BMC Evolutionary Biology.* **19**(1): 107.

Areal, H., Abrantes, J. and Esteves, P. J. (2011). Signatures of positive selection in Toll-like receptor (TLR) genes in mammals. *BMC Evolutionary Biology.* **11**(1): 368.

Baldauf, S. A., Kullmann, H., Schroth, S. H., Thünken, T. and Bakker, T. C. M. (2009). You can't always get what you want: size assortative mating by mutual mate choice as a resolution of sexual conflict. *BMC Evolutionary Biology.* **9**(1): 129.

Barton, G. M. (2007). Viral recognition by Toll-like receptors. *Seminars in Immunology.* **19**(1): 33-40.

Barton, K. and Barton, M. K. (2019). Package ‘MuMIn’.

Bateson, Z. W., Hammerly, S. C., Johnson, J. A., Morrow, M. E., Whittingham, L. A. and Dunn, P. O. (2016). Specific alleles at immune genes, rather than genome-wide heterozygosity, are related to immunity and survival in the critically endangered Attwater's prairie-chicken. *Molecular Ecology.* **25**(19): 4730-4744.

Bebbington, K., Spurgin, L. G., Fairfield, E. A., Dugdale, H. L., Komdeur, J., Burke, T. and Richardson, D. S. (2016). Telomere length reveals cumulative individual and transgenerational inbreeding effects in a passerine bird. *Molecular Ecology.* **25**(12): 2949-2960.

Bernatchez, L. and Landry, C. (2003). MHC studies in nonmodel vertebrates: what have we learned about natural selection in 15 years? *Journal of Evolutionary Biology.* **16**(3): 363-377.

Birkhead, T. R. and Møller, A. (1998). Sperm competition. *New York: Academic.*

Birkhead, T. R. and Pizzari, T. (2002). Postcopulatory sexual selection. *Nature Reviews Genetics.* **3**(4): 262-273.

Bollmer, J. L., Dunn, P. O., Whittingham, L. A. and Wimpee, C. (2010). Extensive MHC Class II B Gene Duplication in a Passerine, the Common Yellowthroat (Geothlypis trichas). *Journal of Heredity.* **101**(4): 448-460.

Bonneaud, C., Chastel, O., Federici, P., Westerdahl, H. and Sorci, G. (2006). Complex Mhc-based mate choice in a wild passerine. *Proceedings of the Royal Society B: Biological Sciences.* **273**(1590): 1111-1116.

Borowsky, R., Luk, A. and Kim, R. S. (2019). Sperm swimming behaviors are correlated with sperm haploid genetic variability in the Mexican tetra, Astyanax mexicanus. *PLoS ONE.* **14**(6).

Brandies, P. A., Grueber, C. E., Ivy, J. A., Hogg, C. J. and Belov, K. (2018). Disentangling the mechanisms of mate choice in a captive koala population. *PeerJ.* **6**: e5438.

Brooks, M. E., Kristensen, K., van Benthem, K. J., Magnusson, A., Berg, C. W., Nielsen, A., . . . Bolker, B. M. (2017). glmmTMB balances speed and flexibility among packages for zero-inflated generalized linear mixed modeling. *The R journal.* **9**(2): 378-400.

Brouwer, L., Barr, I., Van De Pol, M., Burke, T., Komdeur, J. and Richardson, D. S. (2010). MHC-dependent survival in a wild population: evidence for hidden genetic benefits gained through extra-pair fertilizations. *Molecular Ecology.* **19**(16): 3444-3455.

Brown, T. J., Dugdale, H., Spurgin, L., Komdeur, J., Burke, T. and Richardson, D. (2021). Causes and Consequences of Telomere Lengthening in a Wild Vertebrate Population. *Authorea Preprints.*

Brownlie, R. and Allan, B. (2011). Avian toll-like receptors. *Cell and Tissue Research.* **343**(1): 121-130.

Burnham, K. P., Anderson, D. R. and Huyvaert, K. P. (2011). AIC model selection and multimodel inference in behavioral ecology: some background, observations, and comparisons. *Behavioral Ecology and Sociobiology.* **65**(1): 23-35.

Chen, S., Cheng, A. and Wang, M. (2013). Innate sensing of viruses by pattern recognition receptors in birds. *Veterinary Research.* **44**(1): 82.

Clutton-Brock, T. (2007). Sexual Selection in Males and Females. *Science.* **318**(5858): 1882-1885.

Coulon, A. (2010). genhet: an easy-to-use R function to estimate individual heterozygosity. *Molecular ecology resources.* **10**(1): 167-169.

Downing, T., Lloyd, A. T., O’Farrelly, C. and Bradley, D. G. (2010). The Differential Evolutionary Dynamics of Avian Cytokine and TLR Gene Classes. *The Journal of Immunology.* **184**(12): 6993-7000.

Dunn, P. O., Bollmer, J. L., Freeman-Gallant, C. R. and Whittingham, L. A. (2013). MHC variation is related to a sexually selected ornament, survival, and parasite resistance in common yellowthroats. *Evolution.* **67**(3): 679-687.

Dutta, S., Sengupta, P., Hassan, M. F. and Biswas, A. (2020). Role of toll-like receptors in the reproductive tract inflammation and male infertility. *Chemical Biology Letters.* **7**(2): 113-123.

Eberhard, W. (1996). *Female control: sexual selection by cryptic female choice.* Princeton University Press.

Edward, D. A. (2014). The description of mate choice. *Behavioral Ecology.* **26**(2): 301-310.

Edwards, S. V. and Hedrick, P. W. (1998). Evolution and ecology of MHC molecules: from genomics to sexual selection. *Trends in Ecology & Evolution.* **13**(8): 305-311.

Eikenaar, C., Komdeur, J. and Richardson, D. (2008a). Natal dispersal patterns are not associated with inbreeding avoidance in the Seychelles warbler. *Journal of Evolutionary Biology.* **21**(4): 1106-1116.

Eikenaar, C., Richardson, D. S., Brouwer, L., Bristol, R. and Komdeur, J. (2009). Experimental evaluation of sex differences in territory acquisition in a cooperatively breeding bird. *Behavioral Ecology.* **20**(1): 207-214.

Eikenaar, C., Richardson, D. S., Brouwer, L. and Komdeur, J. (2007). Parent presence, delayed dispersal, and territory acquisition in the Seychelles warbler. *Behavioral Ecology.* **18**(5): 874-879.

Eikenaar, C., Richardson, D. S., Brouwer, L. and Komdeur, J. (2008b). Sex biased natal dispersal in a closed, saturated population of Seychelles warblers Acrocephalus sechellensis. *Journal of Avian Biology.* **39**(1): 73-80.

Eizaguirre, C., Lenz, T. L., Kalbe, M. and Milinski, M. (2012). Rapid and adaptive evolution of MHC genes under parasite selection in experimental vertebrate populations. *Nature Communications.* **3**: 621.

Ejsmond, M. J., Radwan, J. and Wilson, A. B. (2014). Sexual selection and the evolutionary dynamics of the major histocompatibility complex. *Proceedings of the Royal Society of London B: Biological Sciences.* **281**(1796): 20141662.

Emlen, D. J., Warren, I. A., Johns, A., Dworkin, I. and Lavine, L. C. (2012). A Mechanism of Extreme Growth and Reliable Signaling in Sexually Selected Ornaments and Weapons. *Science.* **337**(6096): 860-864.

Fawcett, T. W. and Johnstone, R. A. (2003). Mate choice in the face of costly competition. *Behavioral Ecology.* **14**(6): 771-779.

Ferrer-Admetlla, A., Bosch, E., Sikora, M., Marquès-Bonet, T., Ramírez-Soriano, A., Muntasell, A., . . . Casals, F. (2008). Balancing Selection Is the Main Force

Shaping the Evolution of Innate Immunity Genes. *The Journal of Immunology.* **181**(2): 1315-1322.

Garcia-Navas, V., Ortego, J. and Sanz, J. J. (2009). Heterozygosity-based assortative mating in blue tits (Cyanistes caeruleus): implications for the evolution of mate choice. *Proceedings of the Royal Society B: Biological Sciences.* **276**(1669): 2931-2940.

Garmaeva, S., Sinha, T., Kurilshikov, A., Fu, J., Wijmenga, C. and Zhernakova, A. (2019). Studying the gut virome in the metagenomic era: challenges and perspectives. *BMC Biology.* **17**(1): 84.

Gelman, A., Su, Y., Masanao, Y., Zheng, T. and Dorie, V. (2018). arm: Data Analysis Using Regression and Multilevel/Hierarchical Models, version 1.10-1.

Gilroy, D., van Oosterhout, C., Komdeur, J. and Richardson, D. S. (2017). Toll-like receptor variation in the bottlenecked population of the endangered Seychelles warbler. *Animal Conservation.* **20**(3): 235-250.

Gowaty, P. A. (1996). Battles of the sexes and origins of monogamy. *Partnerships in birds.* **71**: 21-52.

Grieves, L. A., Gloor, G. B., Bernards, M. A. and MacDougall-Shackleton, E. A. (2019). Songbirds show odour-based discrimination of similarity and diversity at the major histocompatibility complex. *Animal Behaviour.* **158**: 131-138.

Griffiths, R., Double, M. C., Orr, K. and Dawson, R. J. G. (1998). A DNA test to sex most birds. *Molecular Ecology.* **7**(8): 1071-1075.

Groenewoud, F., Kingma, S. A., Hammers, M., Dugdale, H. L., Burke, T., Richardson, D. S. and Komdeur, J. (2018). Subordinate females in the cooperatively breeding Seychelles warbler obtain direct benefits by joining unrelated groups. *Journal of Animal Ecology.* **87**(5): 1251-1263.

Grueber, C. E., Wallis, G. P. and Jamieson, I. G. (2013). Genetic drift outweighs natural selection at toll-like receptor (TLR) immunity loci in a re-introduced population of a threatened species. *Molecular Ecology.* **22**(17): 4470-4482.

Grueber, C. E., Wallis, G. P. and Jamieson, I. G. (2014). Episodic Positive Selection in the Evolution of Avian Toll-Like Receptor Innate Immunity Genes. *PLoS ONE.* **9**(3): e89632.

Hadfield, J., Richardson, D. and Burke, T. (2006). Towards unbiased parentage assignment: combining genetic, behavioural and spatial data in a Bayesian framework. *Molecular Ecology.* **15**(12): 3715-3730.

Hamilton, W. D. and Zuk, M. (1982). Heritable true fitness and bright birds: a role for parasites? *Science.* **218**(4570): 384-387.

Hammers, M. and Brouwer, L. (2017). Rescue behaviour in a social bird: removal of sticky 'bird-catcher tree'seeds by group members. *Behaviour.* **154**(4): 403-411.

Hammers, M., Kingma, S. A., Bebbington, K., van de Crommenacker, J., Spurgin, L. G., Richardson, D. S., . . . Komdeur, J. (2015). Senescence in the wild: Insights from a long-term study on Seychelles warblers. *Experimental Gerontology.* **71**: 69-79.

Hammers, M., Kingma, S. A., Spurgin, L. G., Bebbington, K., Dugdale, H. L., Burke, T., . . . Richardson, D. S. (2019). Breeders that receive help age more slowly in a cooperatively breeding bird. *Nature Communications.* **10**(1): 1301.

Hammers, M., Richardson, D. S., Burke, T. and Komdeur, J. (2013). The impact of reproductive investment and early-life environmental conditions on senescence: support for the disposable soma hypothesis. *Journal of Evolutionary Biology.* **26**(9): 1999-2007.

Hansson, B. and Richardson, D. S. (2005). Genetic variation in two endangered Acrocephalus species compared to a widespread congener: estimates based on functional and random loci. Animal Conservation forum, Cambridge University Press.

Hartig, F. (2017). DHARMa: residual diagnostics for hierarchical (multi-level/mixed) regression models. *R package version 0.2.4.* **5**.

Hedrick, P. W. (1994). Evolutionary genetics of the major histocompatibility complex. *The American Naturalist.* **143**(6): 945-964.

Hill, A. V. S., Allsopp, C. E. M., Kwiatkowski, D., Anstey, N. M., Twumasi, P., Rowe, P. A., . . . Greenwood, B. M. (1991). Common West African HLA antigens are associated with protection from severe malaria. *Nature.* **352**(6336): 595-600.

Hoi-Leitner, M., Hoi, H., Romero-Pujante, M. and Valera, F. (1999). Female extra–pair behaviour and environmental quality in the Serin (Serinus serinus): a test of the 'constrained female hypothesis'. *Proceedings of the Royal Society of London. Series B: Biological Sciences.* **266**(1423): 1021-1026.

Holt, W. V. and Look, K. J. W. V. (2004). Concepts in sperm heterogeneity, sperm selection and sperm competition as biological foundations for laboratory tests of semen quality. *Reproduction.* **127**(5): 527.

Hooper, P. L. and Miller, G. F. (2008). Mutual mate choice can drive costly signaling even under perfect monogamy. *Adaptive Behavior.* **16**(1): 53-70.

Hutchens, M., Luker, K. E., Sottile, P., Sonstein, J., Lukacs, N. W., Núñez, G., . . . Luker, G. D. (2008). TLR3 increases disease morbidity and mortality from vaccinia infection. *Journal of immunology (Baltimore, Md. : 1950).* **180**(1): 483-491.

Immler, S. and Birkhead, T. R. (2005). A non-invasive method for obtaining spermatozoa from birds. *Ibis.* **147**(4): 827-830.

Immler, S., Hotzy, C., Alavioon, G., Petersson, E. and Arnqvist, G. (2014). Sperm variation within a single ejaculate affects offspring development in Atlantic salmon. *Biol Lett.* **10**(2): 20131040.

Indykiewicz, P., Podlaszczuk, P., Surmacki, A., Kudelska, K., Kosicki, J., Kamiński, M. and Minias, P. (2017). Scale-of-choice effect in the assortative mating by multiple ornamental and non-ornamental characters in the black-headed gull. *Behavioral Ecology and Sociobiology.* **71**(12): 183.

Jiang, Y., Bolnick, D. I. and Kirkpatrick, M. (2013). Assortative Mating in Animals. *The American Naturalist.* **181**(6): E125-E138.

Johnston, S. E., Gratten, J., Berenos, C., Pilkington, J. G., Clutton-Brock, T. H., Pemberton, J. M. and Slate, J. (2013). Life history trade-offs at a single locus maintain sexually selected genetic variation. *Nature.* **502**(7469): 93-95.

Joseph, S. B. and Kirkpatrick, M. (2004). Haploid selection in animals. *Trends in Ecology & Evolution.* **19**(11): 592-597.

Kamiya, T., O'Dwyer, K., Westerdahl, H., Senior, A. and Nakagawa, S. (2014). A quantitative review of MHC-based mating preference: the role of diversity and dissimilarity. *Molecular Ecology.* **23**(21): 5151-5163.

Kekäläinen, J. and Evans, J. P. (2018). Gamete-mediated mate choice: towards a more inclusive view of sexual selection. *Proceedings of the Royal Society B: Biological Sciences.* **285**(1883): 20180836.

Khan, I., Maldonado, E., Silva, L., Almeida, D., Johnson, W. E., O'Brien, S. J., . . . Antunes, A. (2019). The vertebrate TLR supergene family evolved dynamically by gene gain/loss and positive selection revealing a host–pathogen arms race in birds. *Diversity.* **11**(8): 131.

Khila, A., Abouheif, E. and Rowe, L. (2012). Function, Developmental Genetics, and Fitness Consequences of a Sexually Antagonistic Trait. *Science.* **336**(6081): 585-589.

Kingma, S. A., Bebbington, K., Hammers, M., Richardson, D. S. and Komdeur, J. (2016a). Delayed dispersal and the costs and benefits of different routes to independent breeding in a cooperatively breeding bird. *Evolution.* **70**(11): 2595-2610.

Kingma, S. A., Komdeur, J., Burke, T. and Richardson, D. S. (2017). Differential dispersal costs and sex-biased dispersal distance in a cooperatively breeding bird. *Behavioral Ecology.* **28**(4): 1113-1121.

Kingma, S. A., Komdeur, J., Hammers, M. and Richardson, D. S. (2016b). The cost of prospecting for dispersal opportunities in a social bird. *Biol Lett.* **12**(6): 20160316.

Kloch, A., Wenzel, M. A., Laetsch, D. R., Michalski, O., Bajer, A., Behnke, J. M., . . . Piertney, S. B. (2018). Signatures of balancing selection in toll-like receptor (TLRs) genes – novel insights from a free-living rodent. *Scientific Reports.* **8**(1): 8361.

Komdeur, J. (1992). Importance of habitat saturation and territory quality for evolution of cooperative breeding in the Seychelles warbler. *Nature.* **358**(6386): 493-495.

Komdeur, J. (1994a). Conserving the Seychelles warbler Acrocephalus sechellensis by translocation from Cousin Island to the islands of Aride and Cousine. *Biological Conservation.* **67**(2): 143-152.

Komdeur, J. (1994b). Experimental evidence for helping and hindering by previous offspring in the cooperative-breeding Seychelles warbler Acrocephalus sechellensis. *Behavioral Ecology and Sociobiology.* **34**(3): 175-186.

Komdeur, J. (1997). Inter-island transfers and population dynamics of Seychelles warblers Acrocephalus sechellensis. *Bird Conservation International.* **7**(1): 7-26.

Komdeur, J., Piersma, T., Kraaijeveld, K., Kraaijeveld-Smit, F. and Richardson, D. S. (2004). Why Seychelles Warblers fail to recolonize nearby islands: unwilling or unable to fly there? *Ibis.* **146**(2): 298-302.

Liu, G., Zhang, H., Zhao, C. and Zhang, H. (2019). Evolutionary History of the Toll-Like Receptor Gene Family across Vertebrates. *Genome Biology and Evolution.* **12**(1): 3615-3634.

Løvlie, H., Gillingham, M. A., Worley, K., Pizzari, T. and Richardson, D. S. (2013). Cryptic female choice favours sperm from major histocompatibility complex-dissimilar males. *Proceedings of the Royal Society of London B: Biological Sciences.* **280**(1769): 20131296.

Medzhitov, R. (2001). Toll-like receptors and innate immunity. *Nature Reviews Immunology.* **1**(2): 135-145.

Milinski, M. (2006). The Major Histocompatibility Complex, Sexual Selection, and Mate Choice. *Annual Review of Ecology, Evolution, and Systematics.* **37**(1): 159-186.

Miller, H. C. and Lambert, D. M. (2004). Gene duplication and gene conversion in class II MHC genes of New Zealand robins (Petroicidae). *Immunogenetics.* **56**(3): 178-191.

Moller, A. P. and Thornhill, R. (1998). Male parental care, differential parental investment by females and sexual selection. *Animal Behaviour.* **55**(6): 1507-1515.

Ortego, J., Calabuig, G., Bonal, R., Muñoz, A., Aparicio, J. M. and Cordero, P. J. (2009). Temporal variation of heterozygosity-based assortative mating and related benefits in a lesser kestrel population. *Journal of Evolutionary Biology.* **22**(12): 2488-2495.

Penn, D. J. and Potts, W. K. (1999). The evolution of mating preferences and major histocompatibility complex genes. *The American Naturalist.* **153**(2): 145-164.

Piertney, S. B. and Oliver, M. K. (2005). The evolutionary ecology of the major histocompatibility complex. *Heredity.* **96**(1): 7-21.

Poltorak, A., Smirnova, I., He, X., Liu, M.-Y., Van Huffel, C., Birdwell, D., . . . Thompson, P. (1998). Genetic and physical mapping of theLpsLocus: identification of the Toll-4 receptor as a candidate gene in the critical region. *Blood Cells, Molecules, and Diseases.* **24**(3): 340-355.

Raj Pant, S., Hammers, M., Komdeur, J., Burke, T., Dugdale, H. L. and Richardson, D. S. (2020). Age-dependent changes in infidelity in Seychelles warblers. *Molecular Ecology.* **29**(19): 3731-3746.

Raj Pant, S., Komdeur, J., Burke, T. A., Dugdale, H. L. and Richardson, D. S. (2019). Socio-ecological conditions and female infidelity in the Seychelles warbler. *Behavioral Ecology.* **30**(5): 1254-1264.

Reusch, T. B. H., Haberli, M. A., Aeschlimann, P. B. and Milinski, M. (2001). Female sticklebacks count alleles in a strategy of sexual selection explaining MHC polymorphism. *Nature.* **414**(6861): 300-302.

Richardson, D., Jury, F., Blaakmeer, K., Komdeur, J. and Burke, T. (2001). Parentage assignment and extra-group paternity in a cooperative breeder: The Seychelles warbler (Acrocephalus sechellensis). *Molecular Ecology.* **10**(9): 2263-2273.

Richardson, D. and Westerdahl, H. (2003). MHC diversity in two Acrocephalus species: the outbred Great reed warbler and the inbred Seychelles warbler. *Molecular Ecology.* **12**(12): 3523-3529.

Richardson, D. S., Burke, T. and Komdeur, J. (2007). Grandparent helpers: the adaptive significance of older, postdominant helpers in the Seychelles warbler. *Evolution.* **61**(12): 2790-2800.

Richardson, D. S., Burke, T., Komdeur, J. and Dunn, P. (2002). Direct benefits and the evolution of female-biased cooperative breeding in Seychelles warblers. *Evolution.* **56**(11): 2313-2321.

Richardson, D. S., Komdeur, J. and Burke, T. (2004). Inbreeding in the Seychelles warbler: environment-dependent maternal effects. *Evolution.* **58**(9): 2037-2048.

Richardson, D. S., Komdeur, J., Burke, T. and von Schantz, T. (2005). MHC-based patterns of social and extra-pair mate choice in the Seychelles warbler. *Proceedings of the Royal Society B: Biological Sciences.* **272**(1564): 759-767.

Roach, J. C., Glusman, G., Rowen, L., Kaur, A., Purcell, M. K., Smith, K. D., . . . Aderem, A. (2005). The evolution of vertebrate Toll-like receptors. *Proceedings of the National Academy of Sciences of the United States of America.* **102**(27): 9577-9582.

Rosvall, K. A. (2011). Intrasexual competition in females: evidence for sexual selection? *Behavioral Ecology.* **22**(6): 1131-1140.

Rowe, M., Veerus, L., Trosvik, P., Buckling, A. and Pizzari, T. (2020). The Reproductive Microbiome: An Emerging Driver of Sexual Selection, Sexual Conflict, Mating

Systems, and Reproductive Isolation. *Trends in Ecology & Evolution.* **35**(3): 220-234.

Saino, N., Galeotti, P., Sacchi, R. and Møller, A. P. (1997). Song and immunological condition in male barn swallows (Hirundo rustica). *Behavioral Ecology.* **8**(4): 364-371.

Santos, P. S. C., Courtiol, A., Heidel, A. J., Höner, O. P., Heckmann, I., Nagy, M., . . . Sommer, S. (2016). MHC-dependent mate choice is linked to a trace-amine-associated receptor gene in a mammal. *Scientific Reports.* **6**(1): 38490.

Santos, P. S. C., Michler, F.-U. and Sommer, S. (2017). Can MHC-assortative partner choice promote offspring diversity? A new combination of MHC-dependent behaviours among sexes in a highly successful invasive mammal. *Molecular Ecology.* **26**(8): 2392-2404.

Schröder, N. W. and Schumann, R. R. (2005). Single nucleotide polymorphisms of Toll-like receptors and susceptibility to infectious disease. *The Lancet infectious diseases.* **5**(3): 156-164.

Sepil, I., Radersma, R., Santure, A. W., De Cauwer, I., Slate, J. and Sheldon, B. C. (2015). No evidence for MHC class I-based disassortative mating in a wild population of great tits. *Journal of Evolutionary Biology.* **28**(3): 642-654.

Setchell, J. M., Charpentier, M. J. E., Abbott, K. M., Wickings, E. J. and Knapp, L. A. (2010). Opposites attract: MHC-associated mate choice in a polygynous primate. *Journal of Evolutionary Biology.* **23**(1): 136-148.

Sheldon, B. C. (1993). Sexually transmitted disease in birds: occurrence and evolutionary significance. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences.* **339**(1290): 491-497.

Sin, Y. W., Annavi, G., Dugdale, H. L., Newman, C., Burke, T. and MacDonald, D. W. (2014). Pathogen burden, co-infection and major histocompatibility complex variability in the European badger (Meles meles). *Molecular Ecology.* **23**(20): 5072-5088.

Sin, Y. W., Annavi, G., Newman, C., Buesching, C., Burke, T., Macdonald, D. W. and Dugdale, H. L. (2015). MHC class II-assortative mate choice in European badgers (Meles meles). *Molecular Ecology.* **24**(12): 3138-3150.

Slade, J. W. G., Watson, M. J. and MacDougall-Shackleton, E. A. (2019). Balancing balancing selection? Assortative mating at the major histocompatibility complex despite molecular signatures of balancing selection. *Ecology and evolution.* **9**(9): 5146-5157.

Sparks, A. M., Spurgin, L. G., van der Velde, M., Fairfield, E. A., Komdeur, J., Burke, T., . . . Dugdale, H. (2020). Telomere heritability and parental age at conception effects in a wild avian population.

Spurgin, L. G., Bebbington, K., Fairfield, E. A., Hammers, M., Komdeur, J., Burke, T., . . . Richardson, D. S. (2018). Spatio-temporal variation in lifelong telomere dynamics in a long-term ecological study. *Journal of Animal Ecology.* **87**(1): 187-198.

Spurgin, L. G. and Richardson, D. S. (2010). How pathogens drive genetic diversity: MHC, mechanisms and misunderstandings. *Proceedings of the Royal Society B: Biological Sciences.* **277**(1684): 979-988.

Spurgin, L. G., Wright, D. J., van der Velde, M., Collar, N. J., Komdeur, J., Burke, T. and Richardson, D. S. (2014). Museum DNA reveals the demographic history of the endangered Seychelles warbler. *Evolutionary Applications.* **7**(9): 1134-1143.

Sutter, A. and Immler, S. (2020). Within-ejaculate sperm competition. *Philosophical Transactions of the Royal Society B: Biological Sciences.* **375**(1813): 20200066.

van de Crommenacker, J., Hammers, M., van der Woude, J., Louter, M., Santema, P., Richardson, D. S. and Komdeur, J. (2017). Oxidative status and fitness components in the Seychelles warbler. *Functional Ecology.* **31**(6): 1210-1219.

van de Crommenacker, J., Komdeur, J., Burke, T. and Richardson, D. S. (2011a). Spatio-temporal variation in territory quality and oxidative status: a natural experiment in the Seychelles warbler (Acrocephalus sechellensis). *Journal of Animal Ecology.* **80**(3): 668-680.

van de Crommenacker, J., Komdeur, J. and Richardson, D. S. (2011b). Assessing the Cost of Helping: The Roles of Body Condition and Oxidative Balance in the Seychelles Warbler (Acrocephalus sechellensis). *PLoS ONE.* **6**(10): e26423.

Velová, H., Gutowska-Ding, M. W., Burt, D. W. and Vinkler, M. (2018). Toll-like receptor evolution in birds: gene duplication, pseudogenisation and diversifying selection. *Molecular Biology and Evolution.* **35**(9): 2170-2184.

Wang, D., Forstmeier, W., Valcu, M., Dingemanse, N. J., Bulla, M., Both, C., . . . Kempenaers, B. (2019). Scrutinizing assortative mating in birds. *PLOS Biology.* **17**(2): e3000156.

Westerdahl, H. (2004). No evidence of an MHC-based female mating preference in great reed warblers. *Molecular Ecology.* **13**(8): 2465-2470.

Westerdahl, H., Stjernman, M., Råberg, L., Lannefors, M. and Nilsson, J.-Å. (2013). MHC-I Affects Infection Intensity but Not Infection Status with a Frequent Avian Malaria Parasite in Blue Tits. *PLoS ONE.* **8**(8): e72647.

Whitcomb, A. C., Banks, M. A. and O'Malley, K. G. (2014). Influence of immune-relevant genes on mate choice and reproductive success in wild-spawning hatchery-reared and wild-born coho salmon (Oncorhynchus kisutch). *Canadian journal of fisheries and aquatic sciences.* **71**(7): 1000-1009.

Wigby, S., Suarez, S. S., Lazzaro, B. P., Pizzari, T. and Wolfner, M. F. (2019). Sperm success and immunity. *Current topics in developmental biology.* **135**: 287-313.

Wilkinson, G. S., Breden, F., Mank, J. E., Ritchie, M. G., Higginson, A. D., Radwan, J., . . . Rowe, L. (2015). The locus of sexual selection: moving sexual selection studies into the post-genomics era. *Journal of Evolutionary Biology.* **28**(4): 739-755.

Wolfson, A. (1952). The Cloacal Protuberance: A Means for Determining Breeding Condition in Live Male Passerines. *Bird-Banding.* **23**(4): 159-165.

Wong, B. B. M. and Candolin, U. (2005). How is female mate choice affected by male competition? *Biological Reviews.* **80**(4): 559-571.

Worley, K., Collet, J., Spurgin, L. G., Cornwallis, C., Pizzari, T. and Richardson, D. S. (2010). MHC heterozygosity and survival in red junglefowl. *Molecular Ecology.* **19**(15): 3064-3075.

Wright, D. J. (2014a). *Evolutionary and conservation genetics of the Seychelles warbler (Acrocephalus sechellensis)*. University of East Anglia.

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

Wright, D. J., Spurgin, L. G., Collar, N. J., Komdeur, J., Burke, T. and Richardson, D. S. (2014b). The impact of translocations on neutral and functional genetic diversity within and among populations of the Seychelles warbler. *Molecular Ecology.* **23**(9): 2165-2177.

Yeates, S. E., Einum, S., Fleming, I. A., Megens, H.-J., Stet, R. J. M., Hindar, K., . . . Gage, M. J. G. (2009). Atlantic salmon eggs favour sperm in competition that have similar major histocompatibility alleles. *Proceedings of the Royal Society B: Biological Sciences.* **276**(1656): 559-566.

Ziegler, A., Santos, P. S. C., Kellermann, T. and Uchanska-Ziegler, B. (2010). Self/nonself perception, reproduction, and the extended MHC. *Self/Nonself.* **1**(3): 176-191.

Chapter 4

Immunogenetic variation shapes the gut microbiome in a wild vertebrate population

Living in a bacterial world - RP 2021

4.1 Abstract:

The gut microbiome (GM) can influence many biological processes in the host, impacting its health and survival, but the GM can also be influenced by the host's traits. In vertebrates, Major Histocompatibility Complex (MHC) genes play a pivotal role in combating pathogens and are thought to shape the host's GM. However, despite this - and the importance of both MHC and GM variation for individual fitness - few studies have investigated the associations between the MHC and GM in the wild. Here, we characterised MHC class I (MHC-I), MHC class II (MHC-II), and GM variation across individuals in a wild population of the Seychelles warbler, *Acrocephalus sechellensis*. We determined how the diversity and composition of the GM varied with MHC characteristics, in addition to other environmental and host intrinsic traits. Our results show that the presence of specific MHC alleles, but not MHC diversity, influences both the diversity and composition of the GM in this population. We found that MHC-I alleles rather than MHC-II alleles had the greatest impact on the GM. The presence of three MHC-I alleles, and one MHC-II allele (*Ase-dab4*) were negatively associated with GM diversity, while the presence of four MHC-I alleles were associated with changes in GM composition. Genome-wide heterozygosity was positively correlated with GM diversity, probably due to decreased inbreeding. These results indicate that components of the host's immune system may play a role in shaping the GM of wild animals, either directly, or indirectly via differential fitness.

4.2 Introduction

Most animals harbour a complex microbial community including bacteria, archaea, viruses, and microbial eukaryotes, collectively known as the host's microbiome. Members of this diverse community have coevolved with their hosts over evolutionary time and, as a result, play a fundamental role in their host's biology and function (Hird 2017). Consequently, it is important to understand both the ecological and evolutionary processes that shape host-associated microbial communities (McFall-Ngai et al. 2013). This is particularly true of the vertebrate gut, where the gut microbiome (GM) contributes to many key biological processes in the host, from enabling nutrient uptake (Hooper et al. 2002), to pathogen defence (Pickard et al. 2017). As such, studies on domestic or laboratory animals have demonstrated that disrupting the GM can have significant consequences for host health and survival (Round and Mazmanian 2009). However, the factors that govern the diversity and composition of the GM often remain unclear, particularly in wild populations where organisms are exposed to far greater levels of environmental and microbial complexity (Amato 2013).

Studies on wild vertebrates are now attempting to unravel the importance of extrinsic factors, such as the environment, season, and diet in contributing to GM variation between species, populations, and individuals (e.g., Hird et al. 2015, Maurice et al. 2015, Greene et al. 2020). Within a population, individual variation in the GM has also been linked to a multitude of host factors including sex (Stoffel et al. 2020), age (Videvall et al. 2018), body condition (Bolnick et al. 2014), cognition (Davidson et al. 2018), sociality (Tung et al. 2015), and hormones (Noguera et al. 2018, Mallott et al. 2020). Host genetics may also play an important role in determining individual GM variation within a population (Spor et al. 2011). For example, genetic relatedness predicts GM dissimilarity (Perofsky et al. 2017, Stoffel et al. 2020) and composition (Suzuki et al. 2019, Couch et al. 2020).

Immunogenetic variation may play a particularly important role in driving differences in the GM across individuals (reviewed in Marietta et al. 2015). Microbial complexity presents a unique challenge to the host immune system, which has evolved to prevent pathogenic microbes from proliferating, whilst still allowing beneficial microbes (i.e., those that form mutualistic and commensal interactions) to remain (Hooper et al. 2012). In humans, inappropriate immune responses can lead to detrimental changes to the GM via a loss of diversity and stability, as well as the proliferation of pathogenic bacteria (Claesson et al. 2012). In turn, such changes in GM composition have been linked to serious health concerns (Round and Mazmanian 2009, Cho and Blaser 2012), so accurate detection of pathogenic microbes by the host is essential. Variation in immune

receptor genes (which detect pathogens, inducing an immune response), have been associated with variation in GM composition (Kurilshikov et al. 2017), making them key candidate genes through which to investigate the relationship between host genotype and the GM. In particular, pattern recognition receptor genes involved in the innate immune response, including toll-like receptors (TLR), nucleotide-binding oligomerization-like receptors, RIG-I-like receptors, and C-type lectin receptors, play an important role in recognising and regulating the GM in humans and captive animals (reviewed in Thaiss et al. 2016, Kurilshikov et al. 2017).

In vertebrates, Major Histocompatibility Complex (MHC) genes, which play a key role in pathogen detection for the adaptive immune response (Piertney and Oliver 2005), may also shape GM variation. MHC genes code for receptor molecules which bind to specific non-self antigens, initiating an adaptive immune response. MHC class I (MHC-I) molecules, expressed on the surface of virtually all somatic cells, mainly recognise intracellular pathogens, while MHC class II (MHC-II) molecules are found on antigen-presenting cells and mainly recognise extracellular pathogens, including bacterial GM species (Hughes and Yeager 1998). Functional variants of MHC genes can confer differential pathogen recognition and directly affect fitness (Hill et al. 1991, Loiseau et al. 2008), with pathogen-mediated selection thought to drive the extraordinary within- and among-population variation observed at the MHC (Hedrick 1994, Spurgin and Richardson 2010). Individuals with different MHC genotypes are, therefore, likely to recognise different microbial species, which could contribute to the inter-individual GM variation seen in wild vertebrate systems. Likewise, comparing functional immunogenetic variation with the diversity and composition of the GM could provide further understanding of the selective pressures acting on the maintenance of host genetic variation, with the GM being one factor driving pathogen-mediated selection at MHC genes.

Various studies on captive animals have found links between individual MHC variation and GM composition (Toivanen et al. 2001, Lin et al. 2014, Khan et al. 2019). However, captivity can profoundly affect an organism’s microbiome (Clayton et al. 2016) and the complexities associated with natural populations cannot be captured in such studies (Amato 2013). Few studies have investigated associations between individual MHC variation and the microbiome in wild animals. Those that have, all focus on MHC-II variation as this is central to humoral immunity against extracellular microbes (Jensen 2007). For example, in a population of threespine sticklebacks (*Gasterosteus aculeatus*), MHC-II and GM diversity were negatively correlated, and the presence of specific MHC-II alleles was associated with changes to GM composition and diversity (Bolnick et al.

2014). In contrast, in two species of giant salamander, while the presence of MHC-II alleles was associated with the composition of the cutaneous microbiome, individual MHC-II divergence was positively correlated with GM diversity and composition (Hernández-Gómez et al. 2018). Lastly, two studies on seabirds found that MHC-II variation is associated with variation in the body surface (feather and glandular) bacterial communities that are responsible for the production of volatile cues involved in mate choice (Pearce et al. 2017, Leclaire et al. 2019). To our knowledge no studies have tested the association between MHC-I genes (or any other host immune genes), and the microbiome in wild systems. This is despite the fact that MHC-I variation could impact the GM indirectly, for example via differences in individual susceptibility to intracellular infections (such as viruses), that would then impact the health and/or GM of the host (Li et al. 2019).

The natural isolated population of Seychelles warbler (*Acrocephalus sechellensis)* on Cousin Island has been extensively monitored since 1997 (Komdeur 1992, Hammers et al. 2019), and DNA samples, as well as the age, sex, status and territory, of nearly all individuals are available from this time. This population harbours reduced genetic variation as a result of a past genetic bottleneck (Spurgin et al. 2014). However, variation still exists at neutral loci across the genome (Richardson et al. 2000) and also, importantly, at both MHC-I (Richardson and Westerdahl 2003) and MHC-II (Hutchings 2009) loci, as well as at some TLR genes (Gilroy et al. 2017). This genetic variation has been linked to differences in individual fitness. For example, genome-wide homozygosity has been negatively linked to individual condition (Bebbington et al. 2016) and reproductive success (Brouwer et al. 2007). Likewise, variation in the viral-sensing TLR3 gene has been linked to differential survival and reproductive success (see chapter 2). Lastly, both a specific MHC-I allele, *ase-ua4,* and MHC-I diversity overall is positively associated with survival (Brouwer et al. 2010) and extra pair paternity (Richardson et al. 2005). Whether there is functional variation at MHC-II in the Seychelles warbler, and if this is linked to variation in fitness remains unresolved.

Here we aim to test whether immunogenetic variation is associated with individual GM variation in the Seychelles warbler. Specifically, we test if bacterial diversity and community composition are associated with individual MHC-I and MHC-II gene diversity, or the presence of specific alleles at MHC-I, MHC-II genes or *TLR3*. It is difficult to make clear predictions about how such associations may be structured. Individual GM diversity could be negatively associated with immunogenetic variation (e.g., Bolnick et al. 2014), if it enables hosts to recognise and eliminate more bacterial species overall, meaning that a reduced diversity of mutualistic strains, or a few highly competitive strains, can

proliferate in the gut. Alternatively, greater, or optimal, immunogenetic diversity may be positively associated with GM diversity. This could occur via two different pathways. First, directly, with greater immunogenetic diversity helping eliminate a greater number of highly competitive (potentially pathogenic) bacterial taxa, while still allowing tolerance of a network of commensal and mutualistic bacterial species. Second, indirectly with greater, or optimal, immunogenetic diversity conferring increased fitness or host health, which has often been associated with greater GM diversity in captive animals (Claesson et al. 2012). We also predict that GM composition and diversity would differ with the presence/absence of specific immune alleles, via differences in immunity and tolerance. This is expected to be most marked for MHC-II alleles as these are expressed extracellularly on antigen-presenting cells that can extend into the gut lumen and so are important in the recognition of gut microbes (Hooper et al. 2012).

4.3 Methods

4.3.1. Study species and sample collection

The Seychelles warbler is a small insectivorous passerine, endemic to the Seychelles archipelago. They are facultative cooperative breeders, defending strict year-round territories (Komdeur 1992). The population on Cousin Island (4°20'S,55°40'E; 0.29km^2) has been extensively monitored since 1985, increasing in intensity from 1997 onwards (Komdeur 1992, Hammers et al. 2019). The Cousin Island population is small, with a carrying capacity of around 320 adult individuals, existing in 115 territories (Komdeur 1992, Brouwer et al. 2009, Sparks et al. 2020) and there is virtually no migration to or from Cousin (Komdeur et al. 2004). Individuals can live for a maximum of 19 years, with a median post-fledgling lifespan of 5.5 years (Hammers and Brouwer 2017). A comprehensive population census is conducted bi-yearly during the major breeding season (June-September) in the south-east monsoon, and the minor breeding season (January-March) in the north-west monsoon (Komdeur and Daan 2005). Territory quality varies quantifiably within and between years (Komdeur et al. 2016); thus, it is possible to separate the influence of environmental factors from host-intrinsic factors when investigating GM variation at the individual level.

Individuals are either caught as chicks in the nest, or by mist net. At first catch, each bird is given a metal British Trust for Ornithology (BTO) ring and a unique combination of three colour rings, allowing them to be individually identifiable. Birds are aged based on hatch date, behaviour or eye colour; grey eyes indicate (< 5 months), light brown eyes

are sub-adults (5–12 months), and those with red-brown eyes are adults (>12 months) (Komdeur 1992, Wright 2014). Blood samples (25 μl) are collected by brachial venipuncture and stored in 0.8 ml of absolute ethanol at either room temperature or 4°C.

Caught birds were placed into a bag immediately following capture. In the first major breeding season (2017) this was a cleaned cotton bag, however, for all following seasons, birds were placed into a single use plastic-lined, flat-bottomed paper bag containing a plastic tray covered by a metal grate according to an established protocol (Knutie and Gotanda 2018). The metal grate and tray were sterilised with a 10% bleach solution between use. Once individuals had defecated, or after a maximum of 30 minutes, individuals were removed from the bag. A sterile flocked swab was used to transfer faecal material into a sterile microcentrifuge tube with 1 ml of absolute ethanol. If the bird defecated outside of the bag/tray this was also collected providing it was possible to do so in as sterile a manner as possible. Control samples from possible contaminants such as the bag, grate, tray, and fieldworkers' hands were taken throughout the sampling seasons using the sterile flocked swabs. Faecal samples were stored in the field at 4°C, before being transported to the lab where they were stored at -80°C prior to extraction.

4.3.2 Molecular methods

Genomic DNA was extracted from blood using the DNeasy blood and tissue kit (Qiagen) Individuals were genotyped at 30 polymorphic microsatellite loci (Richardson et al. 2001, Sparks et al. 2020) and standardized individual microsatellite heterozygosity (Hs) was calculated using the R 3.6.1 function Genhet 3.1 (Coulon 2010). Sex was determined via PCR (Griffiths et al. 1998, Sparks et al. 2020). Variation at one non-synonymous SNP within the leucine-rich repeat domain of TLR3 exon 4 was determined following chapter 2.

4.3.2.1 MHC sequencing and bioinformatics

In total 314 samples were MHC sequenced, including 229 samples from individuals which had microbiome data and 31 samples from individuals previously MHC screened at either MHC-I (Wright et al. 2014) or MHC-II (Hutchings 2009) using older techniques. The remainder included 30 repeat samples, 23 negative controls (making up at least 5% of each plate), and four samples (one per plate) from one great reed warbler (*Acrocephalus arundinaceus*) individual to serve as a positive control.

MHC-I exon 3 and MHC-II exon 2 were amplified using previously validated primer sets (Richardson and Westerdahl 2003, Hutchings 2009) (Table 4.1), with the addition of Illumina index sites. Additionally, six random hexamers (N) were added to the first round PCR (PCR1) primers to increase diversity and improve cluster separation (Miya et al. 2015). The two primer pairs amplifying MHC-I each have a motif-specific primer situated within exon 3, and one general primer situated in intron 3 and so amplify a 262 bp of the full exon (274 bp). These primers had been designed to preferentially amplify functional variants while avoiding known pseudogenes (Richardson and Westerdahl 2003, Westerdahl et al. 2004). The primers for MHC-II, situated within the flanking introns 1 and 2 of exon 2, amplify a 291 bp fragment. These sequences were then edited to the 270 bp MHC-II exon 2 (Hutchings 2009). The term 'allele' is used to describe the different variants amplified for each class of MHC, consistent with other publications investigating MHC diversity, however alleles cannot be assigned to specific (duplicated loci) loci within each MHC class. Previously work suggests that in the Seychelles warbler there is a minimum of four duplicated MHC-I loci and six MHC-II loci (Richardson and Westerdahl 2003, Hutchings 2009).

MHC class	Primer name	5' to 3' sequence	Direction
MHC-II exon 2	cO33	CACCNCCTGACCTGTGTCC	F
MHC-II exon 2	cO43	CGAGGGGACAYGCTCTGCC	R
MHC-I exon 3	HN36	TCCCCACAGGTCTCCACACAGT	F
MHC-I exon 3	HN38	TCCCCACAGGTCTCCACACACG	F
MHC-I exon 3	HN46	ATCCCAAATTCCCACCCACCTT	R

Table 4.1: Primer sequences used for MHC sequencing. Degenerate bases are shown according to IUPAC codes: Y = C/T, N = any base.

All gDNA was normalised to 10 ng/ul. PCR1 amplification occurred in a total volume of 20 µl and used 2 µl gDNA (20 ng), 2 µl of forward and reverse MHC specific primers (5 µM, Table 4.1), 10 µl of 2x Qiagen Multiplex PCR Master Mix (Qiagen) and 4 µl of sterile double-distilled water. PCR1 conditions were as follows: 95°C for 15 minutes followed by cycling (35x) conditions of 94°C for 30 seconds, 61°C for 90 seconds and 72°C for 60 seconds, and a final elongation step at 72°C for 10 minutes. Products were verified on a 1% agarose gel and cleaned using AMPure XP beads at a 1:1 ratio following the Agencourt AMPure XP PCR Purification Kit protocol (Beckman Coulter Genomics, Indianapolis) to remove primer dimers and nontarget DNA fragments. Purified products were resuspended in 15 µl of low TE. Following this, cleaned PCR1 products from the two MHC-I primer combinations (forward primers: HN36 and HN38, Table 4.1) were

combined. Second round PCR (PCR2) amplification to add Illumina indexes occurred in a total volume of 20 µl and used 8 µl of template (separately for pooled products from PCR1 for MHC-I, and PCR1 products for MHC-II), 1 µl of Fi5 and Ri7 illumina indexes (2 µM), and 10 µl of 2x Qiagen Multiplex PCR Master Mix (Qiagen). PCR2 conditions were as follows: 95°C for 15 minutes followed by cycling (12x) conditions of 98°C for 10 seconds, 65°C for 30 seconds and 72°C for 30 seconds, and a final elongation step at 72°C for 10 minutes.

Peak distribution from products before and after PCR2 were checked using an Agilent Tapestation 4200 (Agilent Technologies) with a subset of 8 samples to confirm barcodes had been attached. Concentrations of all PCR2 products were then quantified using QuantiFluor dsDNA kit (Promega) assay. Subsequently, MHC-I and MHC-II PCR2 products were separately pooled on an equimolar basis into groups of 24, resulting in a total of 26 pools, each pool equalised to 80 ng in 20 µl. Pooled amplicons were cleaned first with a 0.5 amplicon :1 bead ratio, and then again with a 1:1 ratio, of AmPure XP beads and eluted with 20 µl of low TE. Each pool was quantified via qPCR (Applied Biosystems, California) using the KAPPA library quantification kit (Kappa Biosystem) and checked using a Qubit™ 4.0 Fluorometer (Invitrogen, Carlsbad) with a Qubit dsDNA BR assay kit (Invitrogen, Carlsbad) before being pooled to create two pools (MHC-I and MHC-II) both at 4 nM. Each final amplicon pool was purified, and size selected by Pippin prep. Final concentrations were quantified with qPCR, and correct peak distribution checked with the Agilent Tapestation 4200 (Agilent Technologies) before the MHC class I and II pools were combined at a 4 nM concentration for subsequent sequencing using 2x250 bp paired-end Illumina MiSeq (Illumina, San Diego).

Processing and MHC genotyping of raw Illumina MiSeq data was conducted using the Amplicon Sequencing Assignment (AmpliSAS) tool (Sebastian et al. 2016). First, FastQC was used to check read quality, before merging pair-ended reads together using AmpliMERGE (10,257,832 merged sequences). AmpliCLEAN was then used to remove low quality reads (Phred score of <30) and those that were missing either primers or barcodes (e.g., from residual PhiX). MHC-I and MHC-II sequences were separated at this stage, resulting in 3,044,897 raw MHC-I sequences, and 6,144,575 raw MHC-II sequences. Following this step, all downstream bioinformatics and analyses were conducted separately for MHC-I and MHC-II. Cutadapt version 1.2.1 (Martin 2011) was used to remove MHC specific primers, the six random hexamers and short reads (<100bp). For MHC-II sequences, remaining intron regions were also removed, leaving a 270bp fragment spanning the full exon. AmpliCHECK was used for preliminary data exploration, using illumina based default settings. Lastly, AmpliSAS was used for

demultiplexing, clustering, and filtering reads. First, a subset of 30 duplicated samples were used to optimise parameters for MHC-I and MHC-II, testing both minimum dominant frequency settings for the clustering step, and minimum amplicon depth for the filtering step, as recommended in (Sebastian et al. 2016). Based on these results (Table S4.1, S4.2) Illumina default clustering settings were used (1% substitution errors, 0.001% indel errors, 25% minimum dominant frequency) for both MHC classes. For the filtering step, chimeras and sequences that only appeared in one sample were removed, and the minimum amplicon frequency was set as 1.6% for MHC-I, and 1.8% for MHC-II. This resulted in 1,267,410 raw MHC-I sequences, and 1,385,049 raw MHC-II sequences. Due to computational limitations the MHC-II dataset was split into two halves and analysed using AmpliSAS separately, before being combined using AmpliCOMBINE in the web version of AmpliSAS.

For MHC-I, the majority of putative alleles were 262 bp, although three sequences <262 bp - that were not homologous to any known MHC gene when checked using blastn - were present in >80% individuals. These shorter fragments were removed from downstream analysis. The majority of MHC-II putative alleles were the full 270 bp length, although there were also sequences between 267-269 bp which were similar to MHC genes (see results). All MHC-II sequences <267 bp in length were not similar to any known MHC genes and so were removed from downstream analysis as putative sequencing artifacts. MHC-I and MHC-II putative alleles were first checked against all known Seychelles warbler alleles. Any unknown putative alleles were then checked against the GenBank (NCBI) nucleotide database (accessed on 25th June 2020) to assess similarity to known MHC alleles from other related species. Additionally, samples of insufficient read depth based on rarefaction curves were removed, which equated to a minimum read depth of 150 per amplicon for MHC-I, and 100 per amplicon for MHC-II. For 30 individuals sequenced twice for repeatability estimates, the sample with the greatest read depth was retained. After processing, the total number of reads assigned to an allele was 1,071,525 for MHC-I (mean ± SEM = 4391.5 ± 149.3 per sample) and 1,123,211 for MHC-II (mean ± SEM = 4603.3 ± 888.2 per sample) in the Seychelles warbler.

4.3.2.2. Microbial extraction, sequencing, and bioinformatics

In total 400 samples were sequenced across three sequencing runs (two plates per sequencing run). This included 343 unique faecal samples (from 235 individuals), 14 control samples, including six extraction negative controls, four positive controls (using a microbial community standard), and four sampling controls. Additionally, 43 faecal

samples were sequenced twice (20 within the same run and 23 across different runs) to determine sequencing repeatability.

Faecal samples were centrifuged for 10 minutes at 10,000 rpm and supernatant was removed. To remove ethanol residue the resulting pellet was washed with 100 µl RNase/DNA free molecular grade water by centrifuging at 10,000 rpm for 10 minutes. Supernatant was then removed, and the washing step repeated a further two times. Microbial DNA was then extracted from 0.05–0.1 g of each sample using the DNeasy PowerSoil Kit (Qiagen), according to an optimised version of the manufacturer's instructions. Modifications included a heat block step (65°C for 10 minutes) prior to bead-beating and elution of DNA in a final volume of 60 µl elution buffer. A ZymoBIOMICS microbial community standard (D6300) was extracted as a positive control using a ZymoBIOMICS DNA miniprep kit (Zymo Research), according to the manufacturer's instructions.

Extracted gDNA was quantified using a Qubit™ 4.0 Fluorometer (Invitrogen) with a Qubit dsDNA HS assay kit (Invitrogen). Aliquots of gDNA were shipped on dry ice to the Centre for Genomic Research, University of Liverpool for library preparation, pooling and sequencing. Bacterial barcoding was performed with a 2-step amplification process using the primers 515F (5'TGCCAGCMGCCGCGGTAA3') and 806R (5'GGACTACHVGGGTWTCTAAT3') (Caporaso et al. 2011) which amplify the V4 region of the 16S rRNA gene. In brief, PCR1 amplification occurred in a total volume of 20 µl and used 5 ng gDNA, 0.5 µl of forward and reverse 16S V4 specific primers (10 µM), and 10 µl of 2x Kapa Hi Fi amplification mix. PCR1 conditions were as follows: 95°C for 2 minutes (hot start) followed by cycling (15x) conditions of 98°C for 20 seconds, 65°C for 15 seconds and 72°C for 30 seconds. Reactions were finished with a 72°C incubation for 5 minutes. PCR products were cleaned using AMPure beads at a 1:1 ratio and resuspended in a total volume of 9 µl. PCR2 amplification occurred using 0.5 µl of the forward i7 and reverse i5 primers (as in MHC methods) using the same conditions as PCR1, but with 20 cycles. PCR2 products were again cleaned with a 1:1 ratio of AMPure beads. Concentrations of each pool were quantified using a Qubit Fluorometer and the correct peak distribution was checked with the Agilent Bioanalyzer DNA HS chip. PCR2 Products were then pooled on an equimolar basis. The amplicon pool was purified, and size selected by Pippin prep, then checked via qPCR before sequencing using 2x250 bp paired-end Illumina MiSeq (Illumina, San Diego).

For each Illumina run, raw reads were first trimmed using Cutadapt 1.2.1 (Martin 2011) to remove Illumina adapter sequences. Reads were further trimmed using Sickle 1.200

with a minimum window quality score of 20, resulting in totals of 12,308,047, 9,397,303, 9,831,508 demultiplexed reads for the three runs (mean ± SEM per sample: 102,567.1 ± 10454.8, 67,123.6 ± 6633.1, 70,225.1 ± 5423.5). Sequences were then analysed using QIIME2 2019.10 (Caporaso et al. 2010). Based on overall quality scores the first 10 bases of each read were trimmed, and sequences truncated to 210 bp for both forward and reverse reads. The DADA2 plugin 2019.10.0 was used to join paired-end reads, denoise, remove chimeras and residual PhiX reads, dereplicate and call amplicon sequence variants (ASVs) (Callahan et al. 2016, Callahan et al. 2017). Following this, results from the three separate runs were merged, resulting in a total of 22,997,693 reads (mean ± SEM per sample: 57,494.3 ± 3424.8) with 36,182 ASVs. A mid-point rooted phylogeny was then constructed using the masked alignment MAFFT (Katoh et al. 2002) and the Fast Tree approach (Price et al. 2009). Taxonomic assignment of ASVs was performed by training a naïve-Bayes classifier on the SILVA 132 16S dataset using 99% sequence similarity (Quast et al. 2012, Yilmaz et al. 2013). Plastid-like sequences and archaeal sequences were removed, as well as singletons which likely represent sequencing errors. Additionally, two ASVs – one from the genus *Defltia* (relative abundance of 90.5% in a negative extraction control from the first run) and one from the genus *Limnobacter* (relative abundance of 99.9% in a negative extraction control from the third run) were removed as probable contaminants. This resulted in a total of 21,904,965 reads (mean ± SEM per sample: 54,899.7 ± 3429.5) with 35,428 ASVs. The final sample metadata, ASV and taxonomy tables were all exported from QIIME2 into R 3.6.1 where they were processed using phyloseq 1.28.0 (McMurdie and Holmes 2013). Sample completeness and rarefaction curves were generated using iNEXT 2.0.20 (Hsieh et al. 2016); completeness plateaued at approximately 10,000 reads and samples with fewer reads were excluded from downstream analyses. Following sequence processing, 34 samples (including all six negative extraction controls) were removed from downstream analysis due to insufficient read depth. Overall, 320 unique samples (93%) remained from 224 individuals.

The repeatability of GM sequencing was tested by comparing the 37 samples that were sequenced multiply within and across sequencing runs. Dissimilarity between pairs of samples were compared using one metric of alpha diversity (the Shannon diversity index), and two metrics of beta diversity (unweighted and weighted UniFrac of between and within duplicated samples) using Kruskal-Wallis tests.

4.3.3. Statistical analysis

Unless otherwise stated, all analyses were conducted in R 3.6.1.

To characterise the Seychelles warbler GM, samples sequenced twice for repeatability analysis were filtered such that only the sample with the greatest read-depth was retained for downstream analysis. Samples collected from the same catch were also filtered to retain a single sample based on their potential exposure to external contamination i.e., samples collected from cleaned trays were prioritised over those collected from other substrates, then the sample with the highest read depth was prioritised. The removal of sample and catch duplicates resulted in 281 samples (from 224 individuals). For all alpha diversity, beta diversity and differential abundance analyses, microbiome samples taken from chicks were excluded due to a small sample size (n = 11). Individuals with incomplete MHC genotype data (n = 25) were also excluded. Lastly, to prevent pseudo-replication, a single sample was selected at random for each individual that had multiple samples taken from different catches, giving an overall sample size of 195 samples from 195 individuals from Cousin Island. This resulted in a total of 10,998,587 reads (mean ± SEM per sample: 56,403 ± 4181.0) with 35,428 ASVs in the un-rarefied dataset.

4.3.3.1. Alpha diversity

All 195 samples were rarefied to a depth of 10,000 reads, based on sample completeness curves, leaving a total of 1,950,000 reads and 27,861 ASVs across samples in the rarefied dataset. Analyses were run using both rarefied and non-rarefied data, however, as results were comparable between datasets and library size was highly variable across samples, only the outcome of models using the rarefied dataset are presented here. Three metrics of alpha diversity were calculated: Chao1 (Chao 1984) (a measure of microbial species richness), Shannon diversity index (Shannon 1948) (a measure of species richness, taking into account sample evenness), and Faith's phylogenetic diversity index (Faith 1992) (a measure of the phylogenetic diversity of a sample). Chao1 and Shannon diversity indices were calculated using phyloseq 1.28.0 (McMurdie and Holmes 2013), and Faith's phylogenetic diversity was calculated using btools 0.0.1 (Battaglia 2018). Both Chao1 and Faith's phylogenetic diversity were log-transformed to improve residual fit.

Linear models with a Gaussian distribution were constructed using glmmTMB 0.2.3 (Brooks et al. 2017) to determine whether the alpha diversity of the GM differed with: 1) the presence or absence of individual immune genes, and 2) immune gene diversity. The first set of models contained the presence or absence of all MHC-I and MHC-II alleles that were present in at least 15% of sampled individuals, and that were the correct length (see above). This included the following alleles: *Ase-ua1/10, Ase-ua3, Ase-ua4, Ase-*

ua5, Ase-ua6, Ase-ua7, Ase-ua8, Ase-ua9, ASe-ua11, Ase-dab3, Ase-dab4, Ase-dab5. The second set of models contained MHC-I diversity, MHC-II diversity and a squared term for both of these terms, since optimal, rather than a greater diversity of MHC alleles could be more beneficial (Nowak et al. 1992) and may result in optimal GM alpha diversity. Both sets of models also included individual heterozygosity (Hs), and TLR3 genotype ($TLR3^{AA}$, $TLR3^{AC}$ or $TLR3^{CC}$). The following factors were also included in all models; field period sampled (major 2017, major 2018 or minor 2018), sex (male, female), age (fledgling, old fledging, sub-adult or adult), as these have been shown to be important factors in determining GM variation in other wild studies. Alpha diversity (Shannon, logChao1, or logFaith) was entered as the response variable. In all models, continuous factors were standardised (scaled and centred) using arm 1.10-1 (Gelman et al. 2018). All biologically relevant interactions were initially included in models but were removed prior to model averaging to prevent model over-parametrisation as all were non-significant ($P < 0.1$). Field season and territory quality were correlated (linear model; $F_{2,185} = 117.2$, $P < 0.001$), and so only field season was included in the models. Collinearity between independent variables was tested using variance inflation factors with an upper limit of three. Collinearity between the presence and absence of immune alleles was further assessed using GGally 2.0.0 (Schloerke et al. 2011). The alleles *Ase-ua1* and *Ase-ua10* were perfectly correlated, and so only *Ase-ua1* was included in the analysis. DHARMA 0.2.4 (Hartig 2017) was used to confirm that there was no over or under dispersion, or residual spatial or temporal autocorrelation in the models. Model averaging - an information-theoretic approach using the dredge function in MUMIn 1.43.6 (Barton and Barton 2019) - was used to select plausible models. All models within seven AICc of the top model were included in the averaged model, to get the final conditional model (Burnham, Anderson et al. 2011).

4.3.3.2. Beta diversity

The unrarefied dataset was further filtered to remove ASVs which appeared in fewer than five samples, and that had a total read depth of <50 across samples, based on an assessment of overall ASV prevalence and abundance (Fig S4.1. Overall, 1,944 out of 35,428 ASVs were retained following filtering. To account for uneven sequencing depth across samples, reads were normalised using the cumulative sum scaling function (Paulson et al. 2013a) in metagenomeSeq 1.26.3 (Paulson et al. 2013b). Two beta diversity metrics which incorporate phylogenetic distance were then calculated using phyloseq 1.28.0 (Lozupone et al. 2006); these were unweighted UniFrac distance (based on the presence/absence of microbial taxa) and weighted UniFrac distance (a quantitative measure which also accounts for differences in the abundances of microbial taxa) (Lozupone et al. 2007). Marginal Permutational Analysis of Variance tests

(PERMANOVAs) were used to assess whether GM community composition differed with immune gene characteristics, using the adonis2 function in Vegan 2.5.6 (Oksanen et al. 2007) with 10,000 permutations. As with alpha diversity models, two sets of PERMANOVA tests were constructed for each beta diversity metric, with the first set of models including the presence/absence of MHC alleles, and the second set of models including MHC diversity, other variables were included as in alpha diversity models. To clarify whether significant differences detected in PERMANOVA tests were caused by differences in mean values, rather than variation in dispersion across groups (Anderson 2001), homogeneity of group dispersions was tested using the betadisper function in Vegan 2.5.6 (Oksanen et al. 2007). Principle coordinate analysis (PCoA), based on weighted and unweighted UniFrac distances, was used to visualise the differences in composition between groups.

4.3.3.3. Differential abundance analysis

To assess whether particular ASVs were differentially abundant between groups of individuals with different immune gene characteristics, DESeq2 1.24.0 (Love et al. 2014) was used. For this analysis, unrarefied reads were filtered but untransformed, as DeSeq2 uses its own variance stabilising transformation to account for variation in library size across samples. DeSeq2 estimates the log2-fold change in microbial abundance between sample groups, using a negative binomial distribution to model ASV counts. Only variables that were associated with significant compositional shifts in PERMANOVA tests (see above) were included in this analysis to avoid over-parametrization (Table S4.3). To account for the large number of zero counts for individual ASVs, the poscounts estimator was included when estimating size factors. Differential ASV abundance was assessed using negative binomial Wald tests and P values were adjusted using the Benjamini and Hochberg false-discovery rate correction, with a significance cut-off of $P < 0.01$. Two ASVs did not converge due to a high number of zero counts across samples; these were removed from the analysis.

4.4 Results

4.4.1. Seychelles warbler GM profile

The overall profile of the Seychelles warbler GM was similar to other passerine bird species (Kropáčková et al. 2017). A total of 40 bacterial phyla were identified across the 281 samples, however, of these, Proteobacteria (42% of total reads), Firmicutes (22%) and Actinobacteria (17%) dominated, with all other phyla being present at lower relative

abundances (summing to <5% of the total read count). The dominant bacterial classes were Gammaproteobacteria (25%), Alphaproteobacteria (16%), Actinobacteria (16%), Bacilli (16%) and Clostridia (6%) (Fig 4.1).

The core microbiome was further characterised at the family level by extracting bacterial families that appeared in at least 50% of samples with a minimum relative abundance of 0.1%. This resulted in the detection of 28 core families, with ASVs from these families making up 74% of all reads. Of the core families, eight were present in at least 80% of samples, and four accounted for >5% of all reads; this latter group consisted of Enterobacteriaceae (23% of total reads), Streptococcaceae (10%), Rhizobiaceae (6%) and Enterococcaceae (5%). Of the assigned genera, 20 were present in at least 50% of samples and ASVs from these genera made up a total of 28% of all reads. However, of these, only two genera (*Microbacterium* and *Enterococcus*) were present in more than 80% of samples.

Despite the presence of a core microbiome, the abundance of individual bacterial taxa was highly variable across individuals (Fig 4.1). Additionally, there was considerable individual variation in alpha diversity when measured as Chao1 (mean = 323.2 ± 14.63 SEM), Shannon (mean = 4.0 ± 0.07), and Faith's phylogenetic diversity (mean = 18.8 ± 0.62).

Repeatability of GM sequencing was tested using three metrics of diversity (Shannon, unweighted, and weighted UniFrac). Between-sample comparisons, i.e., pairwise distances between different samples, were significantly greater (or more dissimilar) than for within-sample comparisons, i.e., pairwise distances when gDNA from the same sample had been sequenced twice (n = 37, $P < 0.001$) (Fig S4.2).

4.4.2. MHC characteristics

244 individuals were successfully genotyped at MHC-I exon 3 and MHC-II exon 2 genes. The repeatability of MHC-I genotypes was 95.0% and repeatability of MHC-II was 90.1%, based on 26 and 24 duplicate samples respectively (Table S4.1, S4.2). The great reed warbler positive control sample had four MHC-I and four MHC-II alleles – all of which mapped with 100% similarity to previously sequenced great reed warbler MHC alleles. Barring one MHC-I sample, which was adjacent to a positive control sample during sequencing and was subsequently removed from the analysis, no other Seychelles warbler samples contained these alleles.

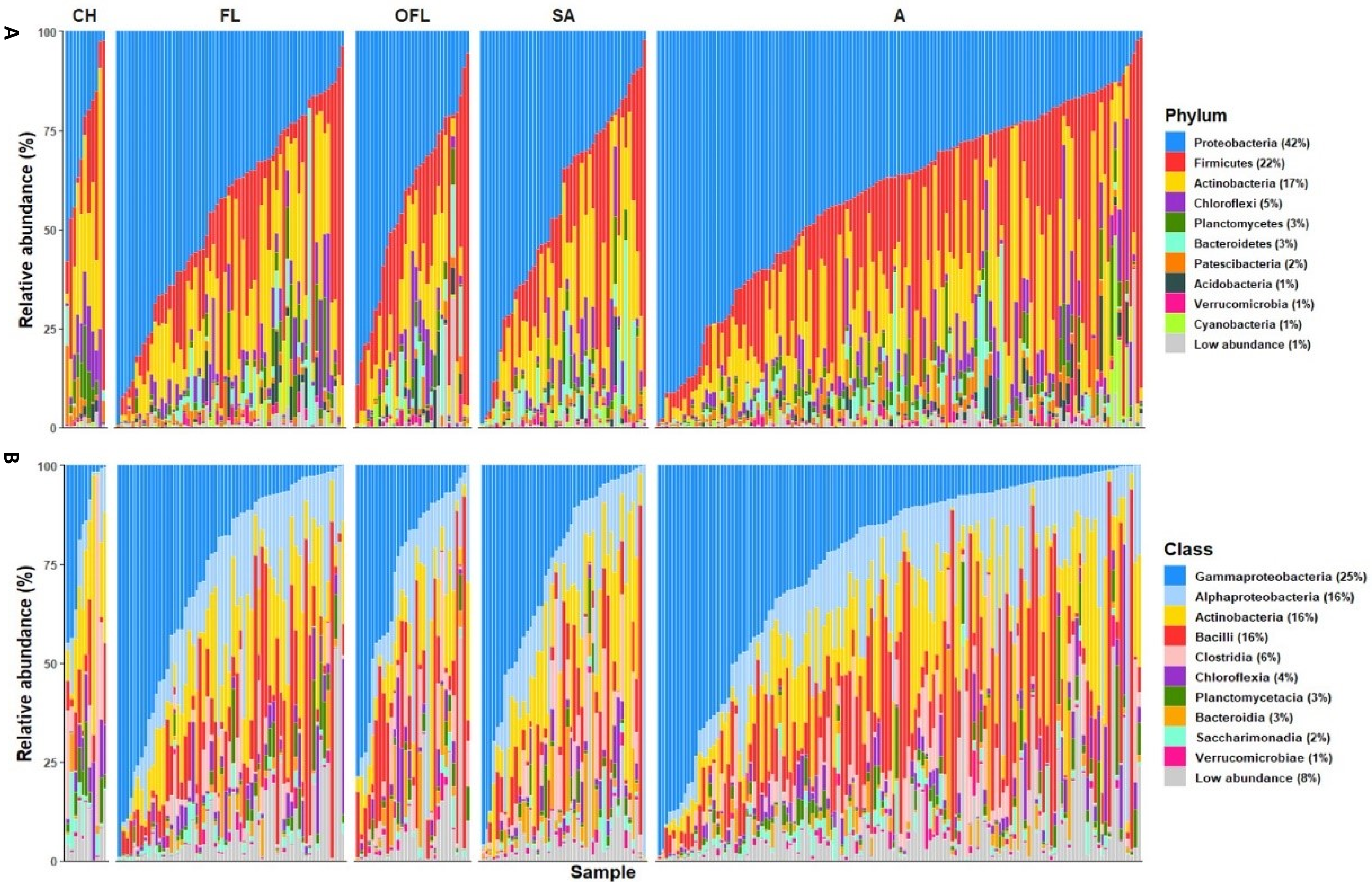


Fig 4.1: The relative abundance (%) of bacterial **A)** phyla and **B)** classes in 281 faecal samples, collected from 224 Seychelles warblers. Each column represents a single faecal sample. Samples have been separated by age class: CH= chick, FL= fledgling, OFL= old fledgling, SA= sub-adult, A= adult. Y-axis values show the relative abundances (%) of the 10 most abundant bacterial taxa. All other taxa are collapsed into the Low abundance category.

On average, individuals had 5.0 MHC-I alleles: 2-7 alleles per individual. Of these, 10 MHC-I alleles were present in >5% but <95% individuals, and another 10 were present in ≤5% of individuals (Fig 4.2). Comparing these 20 alleles to previously sequenced data (Richardson and Westerdahl 2003) nine of the ten common alleles matched previously sequenced alleles, with an average of 98% sequence similarity. One allele, *Ase-ua2*, was not present in the current sequencing cohort. When comparing 29 individuals also genotyped using reference strand-mediated conformation analysis (Wright, Spurgin et al. 2014) and excluding *Ase-ua2*, there is 95% similarity between genotyping methods.

Including all MHC-II alleles, individuals had on average 5.8 alleles (range 3-11) out of a total of 24 alleles (Fig 4.2a). However, of these 24 MHC-II alleles, only 14 were of the full exon length (270 bp), six alleles had a 1 bp deletion (269 bp in length), two alleles had a 2 bp deletion (268 bp) and two alleles had a 3 bp deletion (267 bp). Of the 10 alleles which contained indels, three of these also contained stop codons, and all were missing the Cys74 residue, which along with Cys10 residue creates a disulfide bridge which is important for conformation of the mature MHC protein, therefore, these alleles were removed as putative pseudo or non-functional alleles. Concentrating on putative functional MHC-II alleles, there were 2.9 alleles on average per individual (range from 1-5 alleles per individual). Of these only 3 were present in >5% but < 95% of individuals (Fig 4.2b). Of the other putatively functional alleles, 2 were present in virtually all individuals, while 9 alleles were at a frequency of <5%.

For downstream analysis only alleles of the full, correct length (i.e., MHC-I: 262 bp, MHC-II: 270 bp) were included when calculating diversity or for presence/absence. Additionally, for the presence/absence of alleles, only alleles present in >5% but <95% of individuals were included. Of those alleles included in the final presence/absence analyses, all 10 of the MHC-I and three of the MHC-II alleles translated into unique amino acid sequences.

4.4.3. The effect of MHC and other host variables on GM alpha diversity

The presence of four MHC alleles was associated with reduced diversity and richness of the GM (Fig 4.3). Individuals with the *Ase-ua5* allele had significantly lower alpha diversity for all calculated metrics, compared to individuals without a copy (Table S4.3, Fig 4.3), indicating that *Ase-ua5* negatively influences the richness, evenness and the phylogenetic diversity of the GM. Presence of the *Ase-ua3* allele was also associated with a decrease in both Shannon diversity and Chao1 richness (Table S4.4a, Fig 4.3), but there was no significant difference in the phylogenetic diversity of the GM between

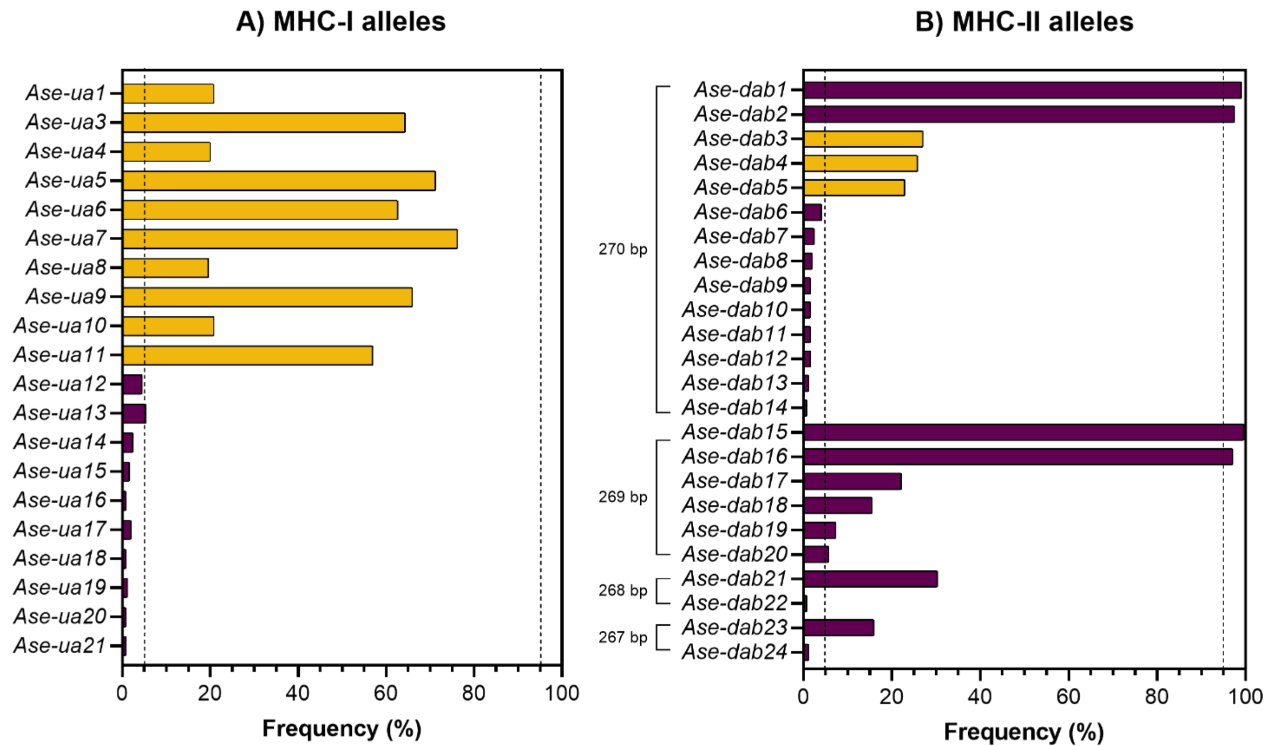


Fig 4.2: Variation in MHC-I exon 3 and MHC-II exon 2 in 244 Seychelles warblers. Each bar represents the frequency (%) of each **A)** MHC-I allele and **B)** MHC-II allele. Colours in bars represent MHC alleles included (in yellow) or not included (in purple) in presence/absence analysis. Dashed lines indicate 5% or 95% frequency cut offs.

individuals with or without the *Ase-ua3* allele (Table S4.4a, Fig 4.3). Presence of the *Ase-ua4* allele was associated with reduced GM richness (Table S4.4a, Fig 4.3) and this effect was approaching significance when taking phylogenetic diversity into account (logFaith's phylogenetic diversity: P = 0.059, Table S4.4a, Fig 4.3). None of the remaining MHC-I alleles or the TLR3 genotype were associated with alpha diversity metrics (Table S4.4a, Fig 4.3). Likewise, the majority of MHC-II alleles were not associated with changes in GM diversity. However, the presence of one allele, *Ase-dab4*, was associated with a reduction in Shannon diversity (Table S4.4a, Fig 4.3), but not Chao1 richness or Faith's phylogenetic diversity (Table S4.4a, Fig 4.3). There was no significant effect of MHC-I or MHC-II diversity, or diversity^2 on alpha diversity (Table S4.4b). In contrast, individual heterozygosity was positively associated with Shannon diversity (Fig 4.3e, Table S4.4a, Fig 4.3), but not Chao1 or Faith's phylogenetic diversity, though these both show the same pattern.

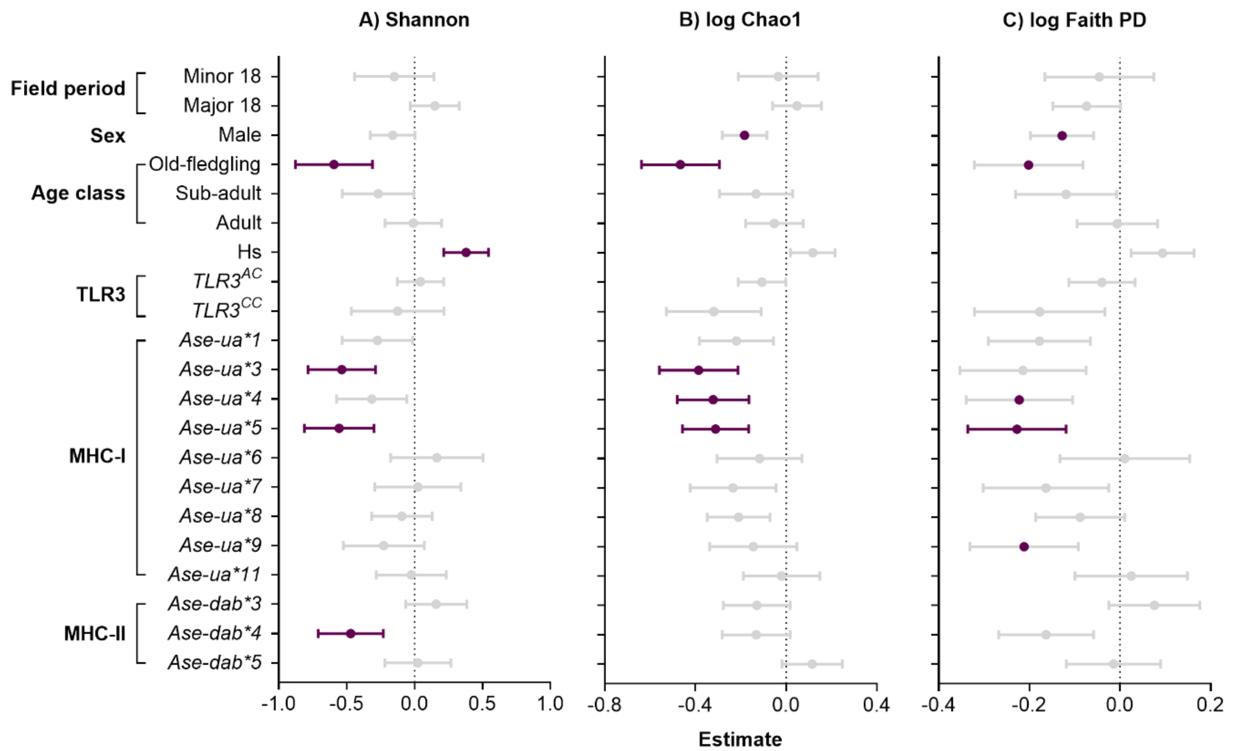


Fig 4.3: Effects of the presence of specific MHC alleles, TLR3 genotype and genome-wide heterozygosity and other host variables on alpha diversity in 195 Seychelles warblers. Alpha diversity metrics are **A)** Shannon diversity, **B)** Chao1, and **C)** faiths phylogenetic diversity. Estimates and standard errors are plotted based on linear conditional mode- averaged estimates. An estimate size >0 indicates increased alpha diversity, while <0 indicates decreased alpha diversity. Significant terms ($P < 0.05$) are highlighted in purple, and terms approaching significance ($P < 0.1$) are indicated with a purple point. Estimates are in reference to MHC allele = Absent, *TLR3* genotype = *TLR3*^{AA}, Sex = Female, Age class = fledgling, Field period = Major 2017.

In regard to other factors, males had a borderline tendency to have reduced richness compared to females (logChao1 $P = 0.062$, logFaith $P = 0.066$, Table S4.4, Fig 4.3). There was also an association between age and GM diversity, with old fledglings having reduced Shannon diversity and Chao1 richness compared to all other age classes (Table S4.4, Fig 4.3), although the phylogenetic diversity of the GM did not change between age classes (Table S4.4, Fig 4.3). There was no association between GM diversity and field period (Table S4.4a, Fig 4.3), suggesting that environmental variation across field periods had little influence on the observed variation in alpha diversity values across individuals.

4.4.4. The effect of host variables on GM composition

In addition to effects on alpha diversity, compositional differences in the GM of individuals with, or without, specific MHC-I alleles were identified, although these alleles were not

the same as those associated with shifts in GM alpha diversity. PERMANOVA tests showed that the overall composition of the GM was significantly different for individuals with the *Ase-ua11* allele versus those without it (Table S4.3, Fig 4.4a). GM composition was also significantly different for individuals possessing either the *Ase-ua7* allele (Table S4.3, Fig 4.4b) or the combined *Ase-ua1/10* alleles (Table S4.3, Fig 4.4c) compared to individuals without these alleles, but only when weighted UniFrac distances were used as a beta diversity metric. None of the differences detected in PERMANOVA tests were due to differential dispersion (All betadisper tests: $P > 0.05$), indicating that results reflected differences in mean values across groups. However, although statistically significant, the presence and absence of specific alleles only explained between a minimum of 0.4% and a maximum of 1.7% per allele of the variation in GM composition, suggesting that each allele only had a small individual influence on GM composition overall (Table S4.3). The remaining MHC-I and MHC-II alleles, as well as MHC-I and MHC-II diversity (or diversity^2) had no effect on GM composition (Table S4.3). Additionally, TLR3 genotype and Hs were not associated with any of the beta diversity metrics (Table S4.3).

Looking at other host factors, age class was associated with a compositional shift in the GM, in PERMANOVA's based on unweighted UniFrac distances (Table S4.3), explaining 1.9% of the variance in GM composition. Based on the Principal Coordinate analysis plots, this difference was due to old fledglings being slightly more differentiated compared to other age classes (Fig S4.3). However, this effect was absent in models based on weighted UniFrac which takes the abundance of ASVs into account (Table S4.3). This suggests that the changes in composition with age class may be due to differences in the presence/absence of different bacterial taxa in the GM, rather than differing abundances of these taxa. There were no differences in beta diversity between males and females (Table S4.3). Focusing on the influence of extrinsic factors on the GM, there were significant compositional differences in the GM between field periods which overall explained either 1.7% or 2.0% variance for unweighted and weighted UniFrac distance, respectively (Table S4.3).

4.4.5. The influence of host variables on the abundance of specific ASVs

The correlated *Ase-ua1/10* alleles were associated with the greatest change in ASV abundance, with 32 ASVs (across 15 orders) being significantly more abundant when

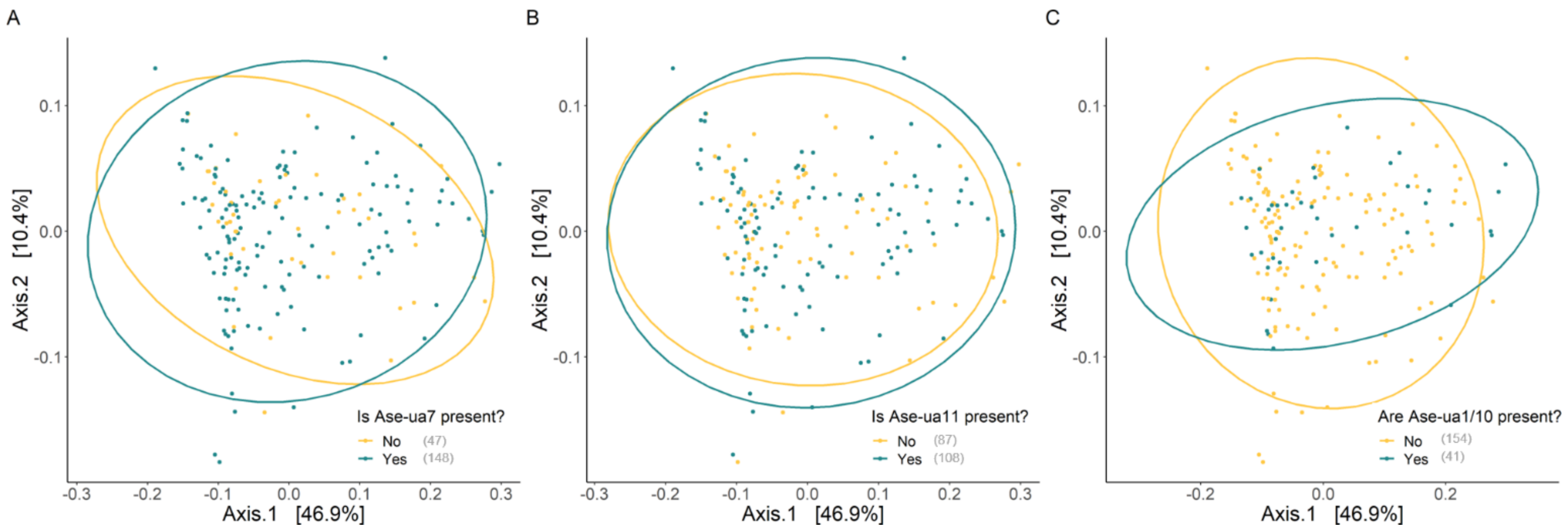


Fig 4.4: Beta diversity of Seychelles warbler gut microbiome composition according to the presence or absence of the MHC-I **A)** *Ase-ua7* allele, **B)** *Ase-ua11* allele, or **C)** *Ase-ua1/10* allele. The principal coordinate plots are based on weighted UniFrac distances. Points represent a single faecal sample from a single individual (n = 195). Sample sizes are specified in brackets in the legend, and colours indicate the presence (blue) or absence (yellow) of the MHC-I allele. Ellipses represent a 95% confidence interval around the cluster centroids.

the alleles were absent and 70 ASVs (across 33 orders) being more abundant when they were present (Fig 4.5c). Fewer taxa were differentially abundant between groups of individuals with/without *Ase-ua11*. In this instance, 32 ASVs (across 17 orders) were significantly more abundant when the allele was absent, and 12 ASVs (across 7 orders) were more abundant when the allele was present (Fig 4.5a). Overall, 29 ASVs were significantly more abundant when the allele *Ase-ua7* was present and 22 ASVs were more abundant when the allele was absent (Fig 4.5b).

Many ASVs were significantly more abundant in old fledglings when compared to fledglings (149 of the 175 differentially abundant taxa were more abundant in old fledglings, Fig S4.4), sub-adults (133 out of 190 taxa), or adults (169 out of 188 taxa). In comparison, the numbers of differentially abundant taxa between other age groups were more even (fledglings compared to sub-adults: 55 compared to 48; fledglings compared to adults: 22 compared to 35; sub-adults compared to adults: 104 compared to 22, Fig S4.4).

Concentrating on extrinsic associations with GM, 229, 225, and 146 ASVs differed significantly in abundance between the three field periods (Fig S4.5). The majority of ASVs were overrepresented in the minor 2018 season compared to either major season (166 in the Minor 2018 compared to 59 in the Major 2017 season, and 192 in the minor 2018 season and 37 in the major 2018 season). Of these, 150 ASVs from the minor season were differently abundant across analysis.

4.5. Discussion

In this study we have screened both the GM variation and the MHC class I and II characteristics of individuals in a natural population of the Seychelles warbler. This has also allowed us to test how individual immunogenetic variation - i.e., MHC genes and the TLR3 loci - is linked to the bacterial diversity and composition of the GM. Our results indicate that presence of certain MHC alleles within an individual, result in differences in the bacterial diversity within the GM (specifically, four out of the 13 tested MHC alleles are associated with lower diversity). Furthermore, the presence/absence of four (different) MHC-I alleles was linked to differences in GM composition, including the differential abundance of certain bacterial taxa. While we found no effect of MHC diversity or TLR3 genotype on GM diversity or composition, we did find a positive association between individual genome-wide heterozygosity and bacterial GM diversity.

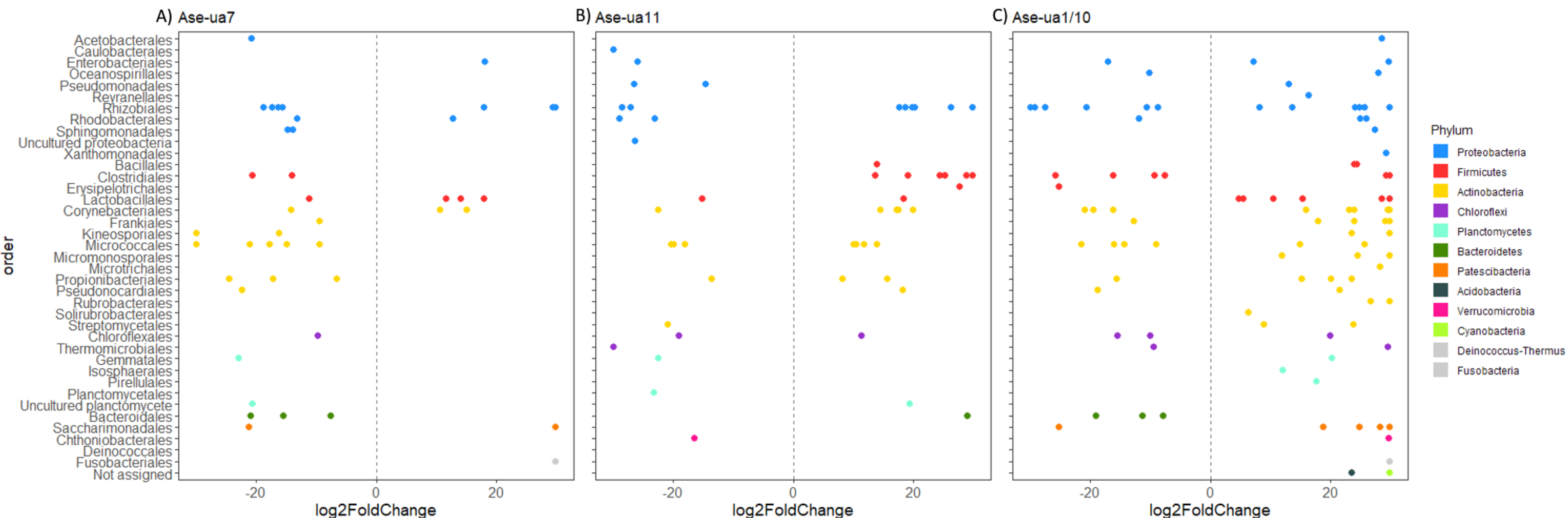


Fig 4.5: Differentially abundant ASV's (P_{adj} < 0.01) in the gut microbiome of Seychelles warblers, according to the presence/absence of the MHC-I alleles **A)** *Ase-ua7* **B)** *Ase-ua11* or **C)** *Ase-ua1/10*. ASVs are grouped at the level of bacterial order and coloured according to bacterial phylum. ASVs shown with a log_2 fold change greater than zero are significantly more abundant in individuals with this allele and ASVs with a log_2 fold change smaller than zero are significantly more abundant in individuals without a copy of this allele.

Lastly, several other extrinsic and host specific variables were also associated with GM characteristics, namely sex, age and sampling period.

4.5.1. MHC variation in the Seychelles warbler

Although previous studies have already characterised MHC variation in the Seychelles warbler (Richardson and Westerdahl 2003, Hutchings 2009), here we screened variation at the MHC-I exon 3 and MHC-II exon 2 genes using next-generation sequencing for the first time. We found reduced functional allelic diversity at MHC-II compared to MHC-I, consistent with what has been found in other passerines (Minias et al. 2018). Previous studies on the Seychelles warbler have provided evidence that balancing selection has maintained variation at both the MHC-I (Richardson and Westerdahl 2003) and MHC-II (Hutchings 2009). However, the latter study –did not resolve individual MHC-II characteristics because of difficulties with the cloning and reference strand-mediated conformation analysis techniques used. In the present study we are able to confirm (class-I) and fully characterise (class-II) individual MHC variation. Our results, showing that variation has been maintained at both sites in this species, despite reduced genome-wide variation (Spurgin et al. 2014), supports the idea that balancing selection could be maintaining variation at both MHC classes. Given the MHCs role in pathogen detection, this is likely to be pathogen-mediated selection. Previous work on the Seychelles warbler has found only one MHC-I allele, *Ase-ua4* predicts survival (Brouwer et al. 2010). Fitness associations with MHC-II have not yet been explored. Though, as survival is a definitive measure, which can be confounded by factors other than an individual's fitness, such as stochastic mortality, it could be that we are missing more minimal detrimental costs of MHC variation. Correlating non-lethal accumulative costs with MHC variation would provide more resolution of the impact of immunogenetic variation on individual fitness, and if this could in part, be driven by pathogen-mediated selection.

We detected a number of non-functional MHC-II alleles (with stop codons, or frameshift deletions) at relatively high frequency. Although such pseudogenes are not uncommon in passerines (Westerdahl 2007) these alleles had as high variability and frequency as the putatively functional MHC-II alleles, suggesting drift, as well as balancing selection, could still be playing a large part in driving variation at MHC-II. We did not detect any non-functional alleles in MHC-I, likely because the primers were designed to avoid the amplification of pseudogenes (Westerdahl et al. 2004).

4.5.2. Does MHC variation shape GM variation?

We found that the presence of three (out of 10) MHC-I alleles (*Ase-ua3, Ase-ua4, Ase-ua5*), and one (out of three) MHC-II allele (*Ase-dab4*) were negatively associated with GM diversity. All three MHC-I alleles were consistently associated with a reduction in GM bacterial richness and *Ase-ua5* was additionally associated with reduced phylogenetic diversity in the GM. This suggests that these alleles may lead to the selective elimination of certain bacterial taxa from the gut, but the tolerance of a reduced community of species with a narrower phylogenetic range. This is similar to the findings of another study which identified MHC-II motifs associated with reduced GM diversity in wild threespine sticklebacks (Bolnick et al. 2014). It is likely that MHC-II alleles - such as *Ase-dab4* in the Seychelles warbler - directly impact GM diversity, since MHC-II molecules are produced in antigen-presenting cells, which are abundant in the lamina propria behind the gut epithelium (Hooper et al. 2012). Such dendritic cells can extend between epithelial cells and phagocytose particles, including microbes, from the gut lumen (Niess et al. 2005). Antigens from these microbes are then exported to the cell surface by MHC-II molecules, so that they can be presented to B- and T-cell populations (Hooper and Macpherson 2010) and thus instigate an immune response.

Our study expands on previous work by also investigating how variation at MHC-I genes may impact GM variation. To our knowledge it is the first-time individual MHC-I variation has been shown to be associated with differences in GM characteristics in a wild population. MHC-I molecules typically respond to intracellular pathogens, rather than extracellular microbes and play a central role in anti-viral and anti-tumour immunity (Hughes and Yeager 1998). As such we would not necessarily expect them to recognise bacterial antigens, although cross reactivity via the presentation of exogenously derived antigens can occur (Ackerman and Cresswell 2004). Alternatively, MHC-I variation may indirectly affect GM characteristics by impacting other aspects of the host's health and physiology. In the Seychelles warbler we know that a specific MHC-I allele (A*se-ua4*) is linked to differential survival (Brouwer et al. 2010) and, although we do not know what drives this effect (i.e., we have not identified the host-pathogen interaction responsible), any such interaction could also shape changes in the GM. In our study, *Ase-ua4* allele was associated with a reduction in GM richness but did not significantly change the composition of the GM. Numerous studies have linked MHC-I variation and susceptibility to malarial infection in passerines (Westerdahl et al. 2005, Biedrzycka et al. 2018) and other taxa (Hill et al. 1991). Malaria infections can alter the GM, via disruption to immune homeostasis (Mukherjee et al. 2020). Pathogen screening of the Seychelles warbler has identified a single strain of *Haemoproteus nucleocondensus* (Fairfield et al. 2016), but no evidence of gastrointestinal macroparasites (Hutchings 2009). Although the malarial strain is not highly pathogenic in the Seychelles warbler (Hammers et al. 2016) it likely

has some impact on individual condition and fitness across the lifespan - as shown in a close congener (Asghar et al. 2015). In the Seychelles warbler malarial infection has been linked to more rapid telomere shortening (Brown et al. 2021), although not reduced survival (Hammers et al. 2016). Thus, it is possible that a malarial-host MHC interaction can consequently impact the host GM. Alternatively the MHC-I may alter an individual's susceptibility to other, as yet unidentified pathogens, such as viruses, that could in turn, drive differences in the diversity of the GM (Rosshart et al. 2017, Li et al. 2019).

Given the negative, but lack of positive, associations identified between MHC alleles and GM alpha diversity, it is surprising that there was no significant effect of overall MHC diversity on the GM. One might expect the cumulative negative effect of each of those alleles (Fig 4.3) to cause at least a weak negative correlation between MHC and GM diversity. However, given the multitude of factors involved in determining both the host GM and immune response, this lack of association between MHC diversity and GM could simply be due limited power to detect weak effects, as is often the case when examining associations between immunocompetence and MHC variation (Gaigher et al. 2019). Assessing how the functional divergence of MHC alleles within an individual – which provides information about the range of antigens that can be detected (Wakeland et al. 1990) - rather than just the number of MHC genes has provided additional resolution in other MHC studies e.g. (Pineaux et al. 2020). This approach could provide extra clarity, particularly when considering the diversity of bacterial taxa that are present within the GM.

We also observed compositional differences in the GM associated with four MHC-I alleles (*Ase-ua7*, *Ase-ua11*, and the linked *Ase-ua1/10* alleles) but no MHC-II alleles. These alleles were different to those negatively associated with GM alpha diversity. Changes in GM composition linked to these MHC-I variants may not have translated into changes in alpha diversity because similar numbers of taxa, at comparable abundances, can be present overall, even with compositional shifts. It is likely that MHC-I genes are having an indirect, rather than direct effect on GM composition (see discussion above). The biggest compositional shift, causing the greatest number of differentially abundant ASVs, was associated with the co-occurring *Ase-ua1/10* allele combination. This is perhaps not surprising as individuals with both of these alleles would be able to recognise a larger number of antigens, thus providing a broader immune response compared to a single allele. However, the presence/absence of the *Ase-ua1/10* alleles only explained 0.5–1.4% (depending on the metric of beta diversity) of the variance in GM composition, suggesting each allele only has a relatively small impact on the GM overall.

Several ASVs were not assigned beyond the level of bacterial family and many bacterial taxa have not been fully characterised, making it difficult to draw conclusions about the functional significance of compositional changes for the host. However, there were several potentially interesting, shared candidate taxa that were differentially abundant between individuals with/without the *Ase-ua1/10* and *Ase-ua11* alleles. For example, individuals with these alleles had greater abundance of ASVs from the order Lactobacillales, a lactic acid producing bacteria, generally thought to be beneficial members of the GM. Indeed, members of this order are used as probiotics in poultry farms to boost the immune response of chickens (Brisbin et al. 2011). Although the ASVs present were largely different between individuals with *Ase-ua1/10,* and *Ase-ua11* alleles. In contrast, individuals with the *Ase-ua1/10* and *Ase-ua11* alleles had reduced abundance of ASVs from Bacteroidales, an order commonly associated with chronic intestinal inflammation (Zitomersky et al. 2013). Two of these ASVs were from the genus *bacteroides*; while species from this genus can be mutualistic, opportunistic pathogenic infections can occur in humans and other animals (Wexler 2007). The third ASV was from the genus *Alistipes,* which has a pathogenic role in various diseases (Parker et al. 2020), this was only differentially abundant for individuals with the *Ase-ua1/10* alleles. The patterns of change associated with *Ase-ua7* were different to those arising from the presence/absence of *Ase-ua1/10* and *Ase-ua11*, with fewer ASVs from the orders Lactobacillales and Bacteriodales differentially abundant. Instead, several ASVs in the order Clostridiales were significantly more abundant when the *Ase-ua7* allele was present (and less abundant when the *Ase-ua11* or *Ase-ua1/10* alleles were present), suggesting that this order could have been selectively tolerated, or that ASVs in this order proliferated when other competing taxa were removed.

Cumulatively, the variance in composition explained by overall MHC allele presence or absence was 6.3% or 9% when unweighted or weighted UniFrac were used, respectively. Although this is low, it is not unusual when looking at factors causing compositional changes between individuals within a single population e.g., environmental and host factors explained between 0.4 – 10% variance in a population of North American red squirrel (*Tamiasciurus hudsonicus*) (Ren et al. 2017). Even sampling period, the most significant predictor of beta diversity in our study, only explained around 2% of variation in the GM. One explanation for the low level of explained variance, could be due to a greater presence of transitionary microbiota in the avian gut (Song et al. 2020). Adaptations for flight have placed constraints on avian morphology, one of which is a comparatively shorter gut and, consequently, shorter food retention times (Caviedes-Vidal et al. 2007); this may constrain the potential for adaptation of bacterial species to the avian gut and to variation in host ecology. A second explanation is that if a large

number of bacterial taxa carry out the same function in the host gut, there could be a high turn-over of species without any consequences for the host (Huttenhower et al. 2012). This can give rise to high inter-individual variation in the GM and may explain why the variables analysed here (or indeed in most within population studies of the GM e.g., (Bolnick et al. 2014, Ren et al. 2017)) explain a low proportion of the overall variance. Indeed, functional diversity in GM may be more important than species diversity; differences in GM characteristics associated with host variables, do not necessarily equate to a biologically difference in microbial function between individuals (Louca et al. 2018). To address this, future work incorporating metagenomic analysis would allow greater resolution of bacterial species and an accurate assessment of the functional composition of the GM (Ranjan et al. 2016).

The GM may also be affected by the innate immune response (underpinned by genes such as the toll-like receptors), as well as the acquired immune response (Thaiss et al. 2016), However, we detected no effect of TLR3 genotype on the GM in the Seychelles warbler - one of the few TLRs to have functional variation in this system (Gilroy et al. 2017). This is perhaps not surprising given TLR3's role in recognising viral dsRNA (Barton 2007) rather than any bacterial conserved structures. Although, as the influenza A virus causes changes in the cloacal microbiome in wild waterfowl (Ganz et al. 2017), it is possible that TLR3-viral interactions could have led to indirect effects on the GM in the present study. Not least as variation in TLR3 is significantly associated with survival and reproductive success in the Seychelles warbler (chapter 2).

Individual heterozygosity (Hs) was positively correlated with GM bacterial richness in the Seychelles warbler, though this was not associated with differences in GM phylogenetic diversity or composition. A decrease in individual heterozygosity (or increase in homozygosity) may reflect increased inbreeding. In the Seychelles warbler increased inbreeding has been linked to individual condition, via reduced telomere length (Bebbington et al. 2016), and reproductive success, with maternal homozygosity negatively predicting offspring survival (Richardson et al. 2004, Brouwer et al. 2007) in this species. This suggests we may be detecting an overall indirect effect of increased inbreeding, resulting in decreased fitness or health of individuals, which in turn negatively impacts GM diversity (Grosser et al. 2019). Although it is also possible this association could be driven by a direct effect of heterozygosity of currently unknown, functional loci. In future studies, it could be informative to use either quantitative trait locus mapping (Snijders et al. 2016), or genome-wide association studies to identify candidate genes associated with GM variation e.g. (Bonder et al. 2016, Suzuki et al. 2019). The Seychelles warbler could be particularly useful for this as it underwent a bottleneck in

the 1960's, resulting in 25% reduction in genome-wide variation (Spurgin et al. 2014), thus making it a more tractable study system to disentangle the associations between host genetic variation and the GM.

What impact the identified relationships between MHC alleles and GM variation have on the host is difficult to assess. Typically, greater GM alpha diversity is thought to be beneficial as it correlates with increased health and survival in humans (Claesson et al. 2012), and other captive animals. Thus, the decreased GM diversity associated with particular MHC alleles may be detrimental to individuals in the Seychelles warbler. However, other studies have shown that high alpha diversity can indicate dis-regulation and GM instability (Coyte et al. 2015). Thus, lower GM diversity and a difference in composition could be indicative of a stable network of beneficial bacteria which can proliferate once competing, non-beneficial bacterial strains have been removed. Given the key role MHC genes play in pathogen resistance, it is possible the reduction in alpha diversity we observe is beneficial to the host. Indeed, in mice, MHC-mediated changes to GM composition influence susceptibility to enteric infection, and can hence, be important for host health (Kubinak et al. 2015). In the Seychelles warbler, the negative correlation between *Ase-ua4* and GM alpha diversity could be beneficial, due to the survival advantage conferred by *Ase-ua4* (Brouwer et al. 2010). However, to fully unpick the consequences of these GM/MHC relationships in the Seychelles warbler, further work is needed to test whether there are fitness differences between individuals with particular MHC alleles and GM characteristics. This is no small undertaking and will require extensive and powerful datasets, which are not yet available.

4.5.3. Can the GM drive the evolution of immune genes?

Variation in the GM can affect traits important to the host's own fitness (Suzuki 2017) including host immune function (van Veelen et al. 2020), modulating severity of diseases (Villarino et al. 2016), and ultimately, survival (Benskin et al. 2015), providing the potential for evolutionary adaptation. If immune genes can regulate fitness through modulation of the microbiome, then the microbiome can also influence selection on immune genes. Our results indicate that immune genes can play a role in regulating GM composition and diversity in the Seychelles warbler. However, to explore coevolution of immune genes and the GM we would need to first establish the degree to which GM characteristics and their interaction with immune genes are associated with the fitness (see 4.5.1.). Coevolution between host immune genes and their microbiota can occur by direct pathways (i.e., by MHC genes actively recognising and removing non-beneficial bacteria or by tolerating commensal and mutualistic bacteria), or indirect pathways (i.e.,

by protection from pathogens through competition by commensal bacteria (Hooper et al. 2012, Kubinak et al. 2015). Path analysis could be a powerful way of determining connections between host variables and the direct and indirect influences these have on the GM, and vice versa (Woodhams et al. 2020). Linking host genotype to microbiome variation and exploring how this effects host fitness will help in understanding the evolution of host–microbial interactions.

Pathogen-mediated selection is thought to be central in maintaining diversity at MHC genes (Hedrick 2002, Bernatchez and Landry 2003, Eizaguirre et al. 2012). If components of the GM are acting as a driver of MHC variation, this could explain how variation at MHC genes is maintained in the previously bottlenecked Seychelles warbler (Richardson and Westerdahl 2003, Spurgin et al. 2014), despite the very limited macroparasite fauna in this population (Hutchings 2009). We detected no effect of either MHC diversity (or optimality) associated with the GM, and therefore no evidence of MHC heterozygote advantage in relation to the GM (Doherty and Zinkernagel 1975). However, we did find an effect of specific alleles in determining GM composition, which would be consistent with either rare allele (Slade and McCallum 1992) or fluctuating selection (Hill et al. 1991) mechanism at play in the interaction between the MHC genes and the GM, although differentiating between these two mechanisms is extremely difficult (Spurgin and Richardson 2010). Identifying the exact taxa that are associated with MHC alleles, whether they be pathogenic, beneficial, or commensal, could help infer the significance and direction of these associations.

4.5.4. Effects of age, sex, and field period on the GM

In addition to genetic factors, several other key variables which have been shown to influence individual variation in the GM were included in the analysis. Our results indicated a relationship between age class and GM composition in the Seychelles warbler. Specifically, the GM of newly independent old fledglings had reduced alpha diversity and compositional differences compared to all other age group comparisons (which did not differ from one another). In the Seychelles warbler old fledglings have to forage for themselves for the first time and so may be eating different - perhaps lower quality - food items compared to older birds. This may explain the greater number of differentially abundant taxa present in this age group, including the greater abundance of environmentally derived Planctomycetes order, which are typically transient colonisers of the gut. Alternatively, increased stress in these young individuals encountering new situations and pathogens could contribute to differences between age groups; indeed, mortality is greatest at this point in the Seychelles warbler (Brouwer et al. 2007).

Exposure to stress via glucocorticoids has been shown to alter host GM in other species (Noguera et al. 2018).

While sex is an important determinant of individual variation in wild populations, its importance as a driver of GM variation is unclear. In wild populations, sex often has a weak, if any link to the GM, (Tung et al. 2015, Ren et al. 2017, Teyssier et al. 2018) although some species show more pronounced sex differences in GM, especially sexually dimorphic vertebrates, such as southern and northern elephant seals (*Mirounga Leonine* and *Mirounga angustirostris* respectively) (Nelson et al. 2013, Stoffel et al. 2020). Sex was only associated with a minor difference in the GM in the Seychelles warbler, with males having marginally reduced diversity, but no difference in composition compared to females. It is, perhaps, not surprising that the effect of sex on the GM was so limited, given that Seychelles warblers of both sexes have the same diet, and exhibit limited differences in morphology and behaviour. In threespine stickleback's MHC-GM associations were also found to be sex dependent (Bolnick et al. 2014), however we found no evidence of this in the Seychelles warbler.

Within a species, seasonal changes in diet can be an important factor in driving GM variation (Bolnick et al. 2014, Ren et al. 2017, Michel et al. 2018). In the Seychelles warbler, we found that field period explains 1.7 – 2% of the variance in GM composition. Although the temperature on Cousin Island is relatively stable, there are measurable differences between seasons and years (van de Crommenacker et al. 2011) which could cause a resulting change in diet between seasons, i.e., in the type and abundance of insect prey species. This could explain the observed difference in GM composition, but not diversity, between field periods. For example, average island-wide territory quality increased by 80% in the major 2017 field period and 75% in the minor 2018 field period, compared to the later major 2018 field period. Alternatively, increases in food availability between season could also act indirectly on the GM by buffering individuals against stress or susceptibility to pathogens.

Lastly, it is important to note that the present study is cross-sectional. However, to fully separate and understand how extrinsic and host-intrinsic factors (such as prey availability and age respectively) are driving GM characteristics, longitudinal data and analyses accounting for within- and between-individual differences need to be incorporated (Björk et al. 2019).

4.5.5. Conclusion

Our results show that variation has been maintained at MHC-I and MHC-II genes in the Seychelles warbler, and that the presence of specific alleles, but not MHC diversity, was associated with differences in GM diversity and composition. It is possible that such MHC-GM interactions may explain previous results in this population showing that specific MHC alleles are associated with greater survival. However, further data are needed to establish whether these associations equate to fitness differences between individuals and to better understand host immunogenetic-GM coevolution.

4.6. References

Ackerman, A. L. and Cresswell, P. (2004). Cellular mechanisms governing cross-presentation of exogenous antigens. *Nature Immunology,* **5**(7), 678-684.

Amato, K.R. (2013). Co-evolution in context: The importance of studying gut microbiomes in wild animals. *Microbiome Science and Medicine,* **1**(1).

Anderson, M. J. (2001). A new method for non-parametric multivariate analysis of variance. *Australian ecology,* **26**(1), 32-46.

Asghar, M., Hasselquist, D., Hansson, B., Zehtindjiev, P., Westerdahl, H., and Bensch, S. (2015). Hidden costs of infection: Chronic malaria accelerates telomere degradation and senescence in wild birds. *Science,* **347**(6220), 436-438.

Barton, G. M. (2007). Viral recognition by Toll-like receptors. *Seminars in Immunology,* ***19***(1), 33-40.

Barton, K. and Barton, M. K. (2019). Package ‘MuMIn’.

Battaglia, T. (2018). Package ‘btools’, A Suite of R Function for All Types of Microbial Diversity Analyses.

Bebbington, K., Spurgin, L. G., Fairfield, E. A., Dugdale, H. L., Komdeur, J., Burke, T. and Richardson, D. S. (2016). Telomere length reveals cumulative individual and transgenerational inbreeding effects in a passerine bird. *Molecular Ecology,* **25**(12), 2949-2960.

Benskin, C. M. H., Rhodes, G., Pickup, R. W., Mainwaring, M. C., Wilson, K. Hartley, and I. R. (2015). Life history correlates of fecal bacterial species richness in a wild population of the blue tit Cyanistes caeruleus. *Ecology and evolution,* **5**(4), 821-835.

Bernatchez, L. and Landry, C. (2003). MHC studies in nonmodel vertebrates, what have we learned about natural selection in 15 years? *Journal of Evolutionary Biology*, **16**(3), 363-377.

Biedrzycka, A., Bielański, W., Ćmiel, A., Solarz, W., Zając, T., Migalska, M., . . . Radwan, J. (2018). Blood parasites shape extreme major histocompatibility complex diversity in a migratory passerine. *Molecular ecology,* **27**(11), 2594-2603.

Björk, J. R., Dasari, M., Grieneisen, L., and Archie, E. A. (2019). Primate microbiomes over tim:, Longitudinal answers to standing questions in microbiome research. *American Journal of Primatology,* **81**(10-11), e22970.

Bolnick, D. I., Snowberg, L. K., Caporaso, J. G., Lauber, C., Knight, R. and Stutz, W. E. (2014). Major Histocompatibility Complex class IIb polymorphism influences gut microbiota composition and diversity. *Molecular Ecology,* **23**(19), 4831-4845.

Bolnick, D. I., Snowberg, L. K., Hirsch, P. E., Lauber, C. L., Knight, R., Caporaso, J. G., and Svanbäck, R. (2014). Individuals' diet diversity influences gut microbial diversity in two freshwater fish (threespine stickleback and Eurasian perch). *Ecology letters,* **17**(8), 979-987.

Bonder, M. J., Kurilshikov, A., Tigchelaar, E. F., Mujagic, Z., Imhann, F., Vila, A. V., . . . Zhernakova, A. (2016). The effect of host genetics on the gut microbiome. *Nature Genetics,* **48**(11), 1407-1412.

Brisbin, J. T., Gong, J., Orouji, S., Esufali, J., Mallick, A. I., Parvizi, P., . . . Sharif, S. (2011). Oral treatment of chickens with lactobacilli influences elicitation of immune responses. *Clinical and vaccine immunology,***18**(9), 1447-1455.

Brooks, M. E., Kristensen, K., van Benthem, K. J., Magnusson, A., Berg, C. W., Nielsen, A. , . . . Bolker, B. M. (2017). glmmTMB balances speed and flexibility among packages for zero-inflated generalized linear mixed modeling. *The R Journal*, **9**(2), 378-400.

Brouwer, L., Barr, I., Van De Pol, M., Burke, T., Komdeur, J., and Richardson, D. S. (2010). MHC-dependent survival in a wild population, evidence for hidden genetic benefits gained through extra-pair fertilizations. *Molecular Ecology*, **19**(16), 3444-3455.

Brouwer, L., Komdeur, J., and Richardson, D. S. (2007). Heterozygosity–fitness correlations in a bottlenecked island species, a case study on the Seychelles warbler. *Molecular Ecology,* **16**(15), 3134-3144.

Brouwer, L., Tinbergen, J. M., Both, C., Bristol, R., Richardson, D. S., and Komdeur, J. (2009). Experimental evidence for density-dependent reproduction in a cooperatively breeding passerine. *Ecology,* **90**(3), 729-741.

Brown, T. J., Dugdale, H., Spurgin, L., Komdeur, J., Burke, T., and Richardson, D. S., (2021). Causes and Consequences of Telomere Lengthening in a Wild Vertebrate Population. *Authorea Preprints*.

Burnham, K. P., Anderson, D. R. and Huyvaert, K. P. (2011). AIC model selection and multimodel inference in behavioral ecology: some background, observations, and comparisons. *Behavioral Ecology and Sociobiology.* **65**(1): 23-35.

Callahan, B. J., McMurdie, P. J. and Holmes, S. P. (2017). Exact sequence variants should replace operational taxonomic units in marker-gene data analysis. *The ISME Journal.* **11**(12): 2639.

Callahan, B. J., McMurdie, P. J., Rosen, M. J., Han, A. W., Johnson, A. J. A. and Holmes, S. P. (2016). DADA2: high-resolution sample inference from Illumina amplicon data. *Nature methods.* **13**(7): 581.

Caporaso, J. G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F. D., Costello, E. K., . . . Gordon, J. I. (2010). QIIME allows analysis of high-throughput community sequencing data. *Nature methods.* **7**(5): 335.

Caporaso, J. G., Lauber, C. L., Walters, W. A., Berg-Lyons, D., Lozupone, C. A., Turnbaugh, P. J., . . . Knight, R. (2011). Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. *Proceedings of the National Academy of Sciences.* **108**(Supplement 1): 4516-4522.

Caviedes-Vidal, E., McWhorter, T. J., Lavin, S. R., Chediack, J. G., Tracy, C. R. and Karasov, W. H. (2007). The digestive adaptation of flying vertebrates: High intestinal paracellular absorption compensates for smaller guts. *Proceedings of the National Academy of Sciences.* **104**(48): 19132-19137.

Chao, A. (1984). Nonparametric estimation of the number of classes in a population. *Scandinavian Journal of statistics.* 265-270.

Cho, I. and Blaser, M. J. (2012). The human microbiome: at the interface of health and disease. *Nature Reviews Genetics.* **13**(4): 260-270.

Claesson, M. J., Jeffery, I. B., Conde, S., Power, S. E., O/'Connor, E. M., Cusack, S., . . . O/'Toole, P. W. (2012). Gut microbiota composition correlates with diet and health in the elderly. *Nature.* **488**(7410): 178-184.

Clayton, J. B., Vangay, P., Huang, H., Ward, T., Hillmann, B. M., Al-Ghalith, G. A., . . . Knights, D. (2016). Captivity humanizes the primate microbiome. *Proceedings of the National Academy of Sciences.* **113**(37): 10376-10381.

Couch, C. E., Arnold, H. K., Crowhurst, R. S., Jolles, A. E., Sharpton, T. J., Witczak, M. F., . . . Beechler, B. R. (2020). Bighorn sheep gut microbiomes associate with genetic and spatial structure across a metapopulation. *Scientific Reports.* **10**(1): 6582.

Coulon, A. (2010). genhet: an easy-to-use R function to estimate individual heterozygosity. *Molecular ecology resources.* **10**(1): 167-169.

Coyte, K. Z., Schluter, J. and Foster, K. R. (2015). The ecology of the microbiome: Networks, competition, and stability. *Science.* **350**(6261): 663-666.

Davidson, G. L., Cooke, A. C., Johnson, C. N. and Quinn, J. L. (2018). The gut microbiome as a driver of individual variation in cognition and functional behaviour. *Philosophical Transactions of the Royal Society B: Biological Sciences.* **373**(1756): 20170286.

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

Eizaguirre, C., Lenz, T. L., Kalbe, M. and Milinski, M. (2012). Rapid and adaptive evolution of MHC genes under parasite selection in experimental vertebrate populations. *Nature Communications.* **3**: 621.

Fairfield, E. A., Hutchings, K., Gilroy, D. L., Kingma, S. A., Burke, T., Komdeur, J. and Richardson, D. S. (2016). The impact of conservation-driven translocations on blood parasite prevalence in the Seychelles warbler. *Scientific Reports.* **6**.

Faith, D. P. (1992). Conservation evaluation and phylogenetic diversity. *Biological Conservation.* **61**(1): 1-10.

Gaigher, A., Burri, R., San-Jose, L. M., Roulin, A. and Fumagalli, L. (2019). Lack of statistical power as a major limitation in understanding MHC-mediated immunocompetence in wild vertebrate populations. *Molecular Ecology.* **28**(23): 5115-5132.

Ganz, H. H., Doroud, L., Firl, A. J., Hird, S. M., Eisen, J. A. and Boyce, W. M. (2017). Community-Level Differences in the Microbiome of Healthy Wild Mallards and Those Infected by Influenza A Viruses. *mSystems.* **2**(1): e00188-00116.

Gelman, A., Su, Y., Masanao, Y., Zheng, T. and Dorie, V. (2018). arm: Data Analysis Using Regression and Multilevel/Hierarchical Models, version 1.10-1.

Gilroy, D., van Oosterhout, C., Komdeur, J. and Richardson, D. S. (2017). Toll-like receptor variation in the bottlenecked population of the endangered Seychelles warbler. *Animal Conservation.* **20**(3): 235-250.

Greene, L. K., Williams, C. V., Junge, R. E., Mahefarisoa, K. L., Rajaonarivelo, T., Rakotondrainibe, H., . . . Drea, C. M. (2020). A role for gut microbiota in host niche differentiation. *The ISME Journal.* **14**(7): 1675-1687.

Griffiths, R., Double, M. C., Orr, K. and Dawson, R. J. G. (1998). A DNA test to sex most birds. *Molecular Ecology.* **7**(8): 1071-1075.

Grosser, S., Sauer, J., Paijmans, A. J., Caspers, B. A., Forcada, J., Wolf, J. B. W. and Hoffman, J. I. (2019). Fur seal microbiota are shaped by the social and physical environment, show mother–offspring similarities and are associated with host genetic quality. *Molecular Ecology.* **28**(9): 2406-2422.

Hammers, M. and Brouwer, L. (2017). Rescue behaviour in a social bird: removal of sticky ‘bird-catcher tree’seeds by group members. *Behaviour.* **154**(4): 403-411.

Hammers, M., Kingma, S. A., Spurgin, L. G., Bebbington, K., Dugdale, H. L., Burke, T., . . . Richardson, D. S. (2019). Breeders that receive help age more slowly in a cooperatively breeding bird. *Nature Communications.* **10**(1): 1301.

Hammers, M., Komdeur, J., Kingma, S. A., Hutchings, K., Fairfield, E. A., Gilroy, D. L. and Richardson, D. S. (2016). Age-specific haemosporidian infection dynamics and survival in Seychelles warblers. *Scientific Reports.* **6**: 29720.

Hartig, F. (2017). DHARMa: residual diagnostics for hierarchical (multi-level/mixed) regression models. *R package version 0.2.4.* **5**.

Hedrick, P. W. (1994). Evolutionary genetics of the major histocompatibility complex. *The American Naturalist.* **143**(6): 945-964.

Hedrick, P. W. (2002). Pathogen resistance and genetic variation at MHC loci. *Evolution.* **56**(10): 1902-1908.

Hernández-Gómez, O., Briggler, J. T. and Williams, R. N. (2018). Influence of immunogenetics, sex and body condition on the cutaneous microbial communities of two giant salamanders. *Molecular Ecology.* **27**(8): 1915-1929.

Hill, A. V. S., Allsopp, C. E. M., Kwiatkowski, D., Anstey, N. M., Twumasi, P., Rowe, P. A., . . . Greenwood, B. M. (1991). Common West African HLA antigens are associated with protection from severe malaria. *Nature.* **352**(6336): 595-600.

Hird, S. M. (2017). Evolutionary Biology Needs Wild Microbiomes. *Frontiers in Microbiology.* **8**(725).

Hird, S. M., Sánchez, C., Carstens, B. C. and Brumfield, R. T. (2015). Comparative Gut Microbiota of 59 Neotropical Bird Species. *Frontiers in Microbiology.* **6**: 1403-1403.

Hooper, L. V., Littman, D. R. and Macpherson, A. J. (2012). Interactions Between the Microbiota and the Immune System. *Science.* **336**(6086): 1268-1273.

Hooper, L. V. and Macpherson, A. J. (2010). Immune adaptations that maintain homeostasis with the intestinal microbiota. *Nature Reviews Immunology.* **10**(3): 159-169.

Hooper, L. V., Midtvedt, T. and Gordon, J. I. (2002). How host-microbial interactions shape the nutrient environment of the mammalian intestine. *Annu Rev Nutr.* **22**: 283-307.

Hsieh, T., Ma, K. and Chao, A. (2016). iNEXT: an R package for rarefaction and extrapolation of species diversity (H ill numbers). *Methods in Ecology and Evolution.* **7**(12): 1451-1456.

Hughes, A. L. and Yeager, M. (1998). Natural selection at major histocompatibility complex loci of vertebrates. *Annual review of genetics.* **32**(1): 415-435.

Hutchings, K. (2009). *Parasite-mediated selection in an island endemic, the Seychelles warbler (Acrocephalus sechellensis)*. University of East Anglia.

Huttenhower, C., Gevers, D., Knight, R., Abubucker, S., Badger, J. H., Chinwalla, A. T., . . . The Human Microbiome Project, C. (2012). Structure, function and diversity of the healthy human microbiome. *Nature.* **486**(7402): 207-214.

Jensen, P. E. (2007). Recent advances in antigen processing and presentation. *Nat Immunol.* **8**(10): 1041-1048.

Katoh, K., Misawa, K., Kuma, K. i. and Miyata, T. (2002). MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic acids research.* **30**(14): 3059-3066.

Khan, M. A. W., Stephens, W. Z., Mohammed, A. D., Round, J. L. and Kubinak, J. L. (2019). Does MHC heterozygosity influence microbiota form and function. *PLoS ONE.* **14**(5).

Knutie, S. A. and Gotanda, K. M. (2018). A Non-invasive Method to Collect Fecal Samples from Wild Birds for Microbiome Studies. *Microbial Ecology.*

Komdeur, J. (1992). Importance of habitat saturation and territory quality for evolution of cooperative breeding in the Seychelles warbler. *Nature.* **358**(6386): 493-495.

Komdeur, J., Burke, T., Dugdale, H. L. and Richardson, D. S. (2016). Seychelles warblers: Complexities of the helping paradox. *Cooperative breeding in vertebrates: Studies of ecology, evolution, and behavior.* 197-216.

Komdeur, J. and Daan, S. (2005). Breeding in the monsoon: semi-annual reproduction in the Seychelles warbler (Acrocephalus sechellensis). *Journal of Ornithology.* **146**(4): 305-313.

Komdeur, J., Piersma, T., Kraaijeveld, K., Kraaijeveld-Smit, F. and Richardson, D. S. (2004). Why Seychelles Warblers fail to recolonize nearby islands: unwilling or unable to fly there? *Ibis.* **146**(2): 298-302.

Kropáčková, L., Těšický, M., Albrecht, T., Kubovčiak, J., Čížková, D., Tomášek, O., . . . Kreisinger, J. (2017). Codiversification of gastrointestinal microbiota and phylogeny in passerines is not explained by ecological divergence. *Molecular Ecology.*

Kubinak, J. L., Stephens, W. Z., Soto, R., Petersen, C., Chiaro, T., Gogokhia, L., . . . Morrison, L. (2015). MHC variation sculpts individualized microbial communities that control susceptibility to enteric infection. *Nature Communications.* **6**.

Kurilshikov, A., Wijmenga, C., Fu, J. and Zhernakova, A. (2017). Host Genetics and Gut Microbiome: Challenges and Perspectives. *Trends in Immunology.* **38**(9): 633-647.

Leclaire, S., Strandh, M., Dell'Ariccia, G., Gabirot, M., Westerdahl, H. and Bonadonna, F. (2019). Plumage microbiota covaries with the major histocompatibility complex in blue petrels. *Molecular Ecology.* **28**(4): 833-846.

Li, N., Ma, W.-T., Pang, M., Fan, Q.-L. and Hua, J.-L. (2019). The Commensal Microbiota and Viral Infection: A Comprehensive Review. *Frontiers in Immunology.* **10**(1551).

Lin, P., Bach, M., Asquith, M., Lee, A. Y., Akileswaran, L., Stauffer, P., . . . Rosenbaum, J. T. (2014). HLA-B27 and Human β2-Microglobulin Affect the Gut Microbiota of Transgenic Rats. *PLoS ONE.* **9**(8): e105684.

Loiseau, C., Zoorob, R., Garnier, S., Birard, J., Federici, P., Julliard, R. and Sorci, G. (2008). Antagonistic effects of a Mhc class I allele on malaria-infected house sparrows. *Ecology Letters.* **11**(3): 258-265.

Louca, S., Polz, M. F., Mazel, F., Albright, M. B. N., Huber, J. A., O'Connor, M. I., . . . Parfrey, L. W. (2018). Function and functional redundancy in microbial systems. *Nature Ecology & Evolution.* **2**(6): 936-943.

Love, M. I., Huber, W. and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology.* **15**(12): 550.

Lozupone, C., Hamady, M. and Knight, R. (2006). UniFrac–an online tool for comparing microbial community diversity in a phylogenetic context. *BMC Bioinformatics.* **7**(1): 371.

Lozupone, C. A., Hamady, M., Kelley, S. T. and Knight, R. (2007). Quantitative and Qualitative β Diversity Measures Lead to Different Insights into Factors That Structure Microbial Communities. *Applied and environmental microbiology.* **73**(5): 1576-1585.

Mallott, E. K., Borries, C., Koenig, A., Amato, K. R. and Lu, A. (2020). Reproductive hormones mediate changes in the gut microbiome during pregnancy and lactation in Phayre's leaf monkeys. *Scientific Reports.* **10**(1): 9961.

Marietta, E., Rishi, A. and Taneja, V. (2015). Immunogenetic control of the intestinal microbiota. *Immunology.* **145**(3): 313-322.

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet. journal.* **17**(1): pp. 10-12.

Maurice, C. F., Cl Knowles, S., Ladau, J., Pollard, K. S., Fenton, A., Pedersen, A. B. and Turnbaugh, P. J. (2015). Marked seasonal variation in the wild mouse gut microbiota. *The ISME Journal.* **9**(11): 2423-2434.

McFall-Ngai, M., Hadfield, M. G., Bosch, T. C. G., Carey, H. V., Domazet-Lošo, T., Douglas, A. E., . . . Wernegreen, J. J. (2013). Animals in a bacterial world, a new imperative for the life sciences. *Proceedings of the National Academy of Sciences.* **110**(9): 3229-3236.

McMurdie, P. J. and Holmes, S. (2013). phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS ONE.* **8**(4): e61217.

Michel, A. J., Ward, L. M., Goffredi, S. K., Dawson, K. S., Baldassarre, D. T., Brenner, A., . . . Chaves, J. A. (2018). The gut of the finch: uniqueness of the gut microbiome of the Galápagos vampire finch. *Microbiome.* **6**(1): 167.

Minias, P., Pikus, E., Whittingham, L. A. and Dunn, P. O. (2018). A global analysis of selection at the avian MHC. *Evolution.* **72**(6): 1278-1293.

Miya, M., Sato, Y., Fukunaga, T., Sado, T., Poulsen, J. Y., Sato, K., . . . Iwasaki, W. (2015). MiFish, a set of universal PCR primers for metabarcoding environmental DNA from fishes: detection of more than 230 subtropical marine species. *Royal Society Open Science.* **2**(7): 150088.

Mukherjee, D., Chora, Â. F. and Mota, M. M. (2020). Microbiota, a Third Player in the Host–Plasmodium Affair. *Trends in Parasitology.* **36**(1): 11-18.

Nelson, T. M., Rogers, T. L., Carlini, A. R. and Brown, M. V. (2013). Diet and phylogeny shape the gut microbiota of Antarctic seals: a comparison of wild and captive animals. *Environmental Microbiology.* **15**(4): 1132-1145.

Niess, J. H., Brand, S., Gu, X., Landsman, L., Jung, S., McCormick, B. A., . . . Reinecker, H.-C. (2005). CX₃CR1-Mediated Dendritic Cell Access to the Intestinal Lumen and Bacterial Clearance. *Science.* **307**(5707): 254-258.

Noguera, J. C., Aira, M., Pérez-Losada, M., Domínguez, J. and Velando, A. (2018). Glucocorticoids modulate gastrointestinal microbiome in a wild bird. *Royal Society Open Science.* **5**(4): 171743.

Nowak, M. A., Tarczy-Hornoch, K. and Austyn, J. M. (1992). The optimal number of major histocompatibility complex molecules in an individual. *Proc Natl Acad Sci U S A.* **89**(22): 10896-10899.

Oksanen, J., Kindt, R., Legendre, P., O'Hara, B., Stevens, M. H. H., Oksanen, M. J. and Suggests, M. (2007). The vegan package. *Community ecology package.* **10**(631-637): 719.

Parker, B. J., Wearsch, P. A., Veloo, A. C. M. and Rodriguez-Palacios, A. (2020). The Genus Alistipes: Gut Bacteria With Emerging Implications to Inflammation, Cancer, and Mental Health. *Frontiers in Immunology.* **11**(906).

Paulson, J. N., Pop, M. and Bravo, H. C. (2013). metagenomeSeq: Statistical analysis for sparse high-throughput sequencing. *Bioconductor package.* **1**(0): 191.

Paulson, J. N., Stine, O., Bravo, H. and Pop, M. (2013). Robust methods for differential abundance analysis in marker gene surveys. *Nature methods.* **10**: 1200 - 1202.

Pearce, D. S., Hoover, B. A., Jennings, S., Nevitt, G. A. and Docherty, K. M. (2017). Morphological and genetic factors shape the microbiome of a seabird species (Oceanodroma leucorhoa) more than environmental and social factors. *Microbiome.* **5**(1): 146.

Perofsky, A. C., Lewis, R. J., Abondano, L. A., Fiore, A. D. and Meyers, L. A. (2017). Hierarchical social networks shape gut microbial composition in wild Verreaux's sifaka. *Proceedings of the Royal Society B: Biological Sciences.* **284**(1868): 20172274.

Pickard, J. M., Zeng, M. Y., Caruso, R. and Núñez, G. (2017). Gut microbiota: Role in pathogen colonization, immune responses, and inflammatory disease. *Immunol Rev.* **279**(1): 70-89.

Piertney, S. B. and Oliver, M. K. (2005). The evolutionary ecology of the major histocompatibility complex. *Heredity.* **96**(1): 7-21.

Pineaux, M., Merkling, T., Danchin, E., Hatch, S., Duneau, D., Blanchard, P. and Leclaire, S. (2020). Sex and hatching order modulate the association between MHC-II diversity and fitness in early-life stages of a wild seabird. *Molecular Ecology.* **29**(17): 3316-3329.

Price, M. N., Dehal, P. S. and Arkin, A. P. (2009). FastTree: computing large minimum evolution trees with profiles instead of a distance matrix. *Molecular Biology and Evolution.* **26**(7): 1641-1650.

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., . . . Glöckner, F. O. (2012). The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic acids research.* **41**(D1): D590-D596.

Ranjan, R., Rani, A., Metwally, A., McGee, H. S. and Perkins, D. L. (2016). Analysis of the microbiome: Advantages of whole genome shotgun versus 16S amplicon sequencing. *Biochemical and biophysical research communications.* **469**(4): 967-977.

Ren, T., Boutin, S., Humphries, M. M., Dantzer, B., Gorrell, J. C., Coltman, D. W., . . . Wu, M. (2017). Seasonal, spatial, and maternal effects on gut microbiome in wild red squirrels. *Microbiome.* **5**(1): 163.

Richardson, D., Jury, F., Blaakmeer, K., Komdeur, J. and Burke, T. (2001). Parentage assignment and extra-group paternity in a cooperative breeder: The Seychelles warbler (Acrocephalus sechellensis). *Molecular Ecology.* **10**(9): 2263-2273.

Richardson, D. and Westerdahl, H. (2003). MHC diversity in two Acrocephalus species: the outbred Great reed warbler and the inbred Seychelles warbler. *Molecular Ecology.* **12**(12): 3523-3529.

Richardson, D. S., Komdeur, J. and Burke, T. (2004). Inbreeding in the Seychelles warbler: environment-dependent maternal effects. *Evolution.* **58**(9): 2037-2048.

Richardson, D. S., Komdeur, J., Burke, T. and von Schantz, T. (2005). MHC-based patterns of social and extra-pair mate choice in the Seychelles warbler. *Proceedings of the Royal Society B: Biological Sciences.* **272**(1564): 759-767.

Rosshart, S. P., Vassallo, B. G., Angeletti, D., Hutchinson, D. S., Morgan, A. P., Takeda, K., . . . Rehermann, B. (2017). Wild Mouse Gut Microbiota Promotes Host Fitness and Improves Disease Resistance. *Cell.* **171**(5): 1015-1028.e1013.

Round, J. L. and Mazmanian, S. K. (2009). The gut microbiota shapes intestinal immune responses during health and disease. *Nat Rev Immunol.* **9**(5): 313-323.

Schloerke, B., Crowley, J., Cook, D., Hofmann, H., Wickham, H., Briatte, F., . . . Larmarange, J. (2011). Ggally: Extension to ggplot2.

Sebastian, A., Herdegen, M., Migalska, M. and Radwan, J. (2016). amplisas: a web server for multilocus genotyping using next-generation amplicon sequencing data. *Molecular ecology resources.* **16**(2): 498-510.

Shannon, C. E. (1948). A mathematical theory of communication. *The Bell System Technical Journal.* **27**(3): 379-423.

Slade, R. and McCallum, H. (1992). Overdominant vs. frequency-dependent selection at MHC loci. *Genetics.* **132**(3): 861.

Snijders, A. M., Langley, S. A., Kim, Y.-M., Brislawn, C. J., Noecker, C., Zink, E. M., . . . Huang, Y. (2016). Influence of early life exposure, host genetics and diet on the mouse gut microbiome and metabolome. *Nature Microbiology.* **2**: 16221.

Song, S. J., Sanders, J. G., Delsuc, F., Metcalf, J., Amato, K., Taylor, M. W., . . . Knight, R. (2020). Comparative Analyses of Vertebrate Gut Microbiomes Reveal Convergence between Birds and Bats. *mBio.* **11**(1): e02901-02919.

Sparks, A. M., Spurgin, L. G., van der Velde, M., Fairfield, E. A., Komdeur, J., Burke, T., . . . Dugdale, H. (2020). Telomere heritability and parental age at conception effects in a wild avian population.

Spor, A., Koren, O. and Ley, R. (2011). Unravelling the effects of the environment and host genotype on the gut microbiome. *Nature Reviews Microbiology.* **9**(4): 279-290.

Spurgin, L. G. and Richardson, D. S. (2010). How pathogens drive genetic diversity: MHC, mechanisms and misunderstandings. *Proceedings of the Royal Society B: Biological Sciences.* **277**(1684): 979-988.

Spurgin, L. G., Wright, D. J., van der Velde, M., Collar, N. J., Komdeur, J., Burke, T. and Richardson, D. S. (2014). Museum DNA reveals the demographic history of the endangered Seychelles warbler. *Evolutionary Applications.* **7**(9): 1134-1143.

Stoffel, M. A., Acevedo-Whitehouse, K., Morales-Durán, N., Grosser, S., Chakarov, N., Krüger, O., . . . Hoffman, J. I. (2020). Early sexual dimorphism in the developing gut microbiome of northern elephant seals. *Molecular Ecology.* **29**(11): 2109-2122.

Suzuki, T. A., Phifer-Rixey, M., Mack, K. L., Sheehan, M. J., Lin, D., Bi, K. and Nachman, M. W. (2019). Host genetic determinants of the gut microbiota of wild mice. *Molecular Ecology.* **28**(13): 3197-3207.

Teyssier, A., Rouffaer, L. O., Saleh Hudin, N., Strubbe, D., Matthysen, E., Lens, L. and White, J. (2018). Inside the guts of the city: Urban-induced alterations of the gut microbiota in a wild passerine. *Science of The Total Environment.* **612**(Supplement C): 1276-1286.

Thaiss, C. A., Zmora, N., Levy, M. and Elinav, E. (2016). The microbiome and innate immunity. *Nature.* **535**(7610): 65-74.

Toivanen, P., Vaahtovuo, J. and Eerola, E. (2001). Influence of Major Histocompatibility Complex on Bacterial Composition of Fecal Flora. *Infection and immunity.* **69**(4): 2372-2377.

Tung, J., Barreiro, L. B., Burns, M. B., Grenier, J.-C., Lynch, J., Grieneisen, L. E., . . . Archie, E. A. (2015). Social networks predict gut microbiome composition in wild baboons. *eLife.* **4**: e05224.

van de Crommenacker, J., Komdeur, J., Burke, T. and Richardson, D. S. (2011). Spatio-temporal variation in territory quality and oxidative status: a natural experiment in the Seychelles warbler (Acrocephalus sechellensis). *Journal of Animal Ecology.* **80**(3): 668-680.

van Veelen, H. P. J., Falcão Salles, J., Matson, K. D., van der Velde, M. and Tieleman, B. I. (2020). Microbial environment shapes immune function and cloacal microbiota dynamics in zebra finches Taeniopygia guttata. *Animal Microbiome.* **2**(1): 21.

Videvall, E., Song, S. J., Bensch, H. M., Strandh, M., Engelbrecht, A., Serfontein, N., . . . Cornwallis, C. K. (2018). The development of gut microbiota in ostriches and its association with juvenile growth. *bioRxiv.*

Villarino, N. F., LeCleir, G. R., Denny, J. E., Dearth, S. P., Harding, C. L., Sloan, S. S., . . . Schmidt, N. W. (2016). Composition of the gut microbiota modulates the severity of malaria. *Proceedings of the National Academy of Sciences.* **113**(8): 2235-2240.

Wakeland, E. K., Boehme, S., She, J. X., Lu, C.-C., McIndoe, R. A., Cheng, I., . . . Potts, W. K. (1990). Ancestral polymorphisms of MHC class II genes: Divergent allele advantage. *Immunologic Research.* **9**(2): 115-122.

Westerdahl, H. (2007). Passerine MHC: genetic variation and disease resistance in the wild. *Journal of Ornithology.* **148**(2): 469-477.

Westerdahl, H., Waldenström, J., Hansson, B., Hasselquist, D., von Schantz, T. and Bensch, S. (2005). Associations between malaria and MHC genes in a migratory

songbird. *Proceedings of the Royal Society B: Biological Sciences.* **272**(1571): 1511-1518.

Westerdahl, H., Wittzell, H., Schantz, T. v. and Bensch, S. (2004). MHC class I typing in a songbird with numerous loci and high polymorphism using motif-specific PCR and DGGE. *Heredity.* **92**: 534-542.

Westerdahl, H., Wittzell, H., Schantz, T. v. and Bensch, S. (2004). MHC class I typing in a songbird with numerous loci and high polymorphism using motif-specific PCR and DGGE. *Heredity.*

Westerdahl, H., H. Wittzell, T. v. Schantz and S. Bensch (2004). MHC class I typing in a songbird with numerous loci and high polymorphism using motif-specific PCR and DGGE. Heredity **92**(6), 534-542.

Wexler, H. M. (2007). Bacteroides: the good, the bad, and the nitty-gritty. *Clinical microbiology reviews.* **20**(4): 593-621.

Woodhams, D. C., Bletz, M. C., Becker, C. G., Bender, H. A., Buitrago-Rosas, D., Diebboll, H., . . . Whetstone, R. (2020). Host-associated microbiomes are predicted by immune system complexity and climate. *Genome Biology.* **21**(1): 23.

Wright, D. J. (2014). *Evolutionary and conservation genetics of the Seychelles warbler (Acrocephalus sechellensis)*. University of East Anglia.

Wright, D. J., Spurgin, L. G., Collar, N. J., Komdeur, J., Burke, T. and Richardson, D. S. (2014). The impact of translocations on neutral and functional genetic diversity within and among populations of the Seychelles warbler. *Molecular Ecology.* **23**(9): 2165-2177.

Yilmaz, P., Parfrey, L. W., Yarza, P., Gerken, J., Pruesse, E., Quast, C., . . . Glöckner, F. O. (2013). The SILVA and All-species Living Tree Project (LTP) taxonomic frameworks. *Nucleic acids research.* **42**(D1): D643-D648.

Zitomersky, N. L., Atkinson, B. J., Franklin, S. W., Mitchell, P. D., Snapper, S. B., Comstock, L. E. and Bousvaros, A. (2013). Characterization of Adherent Bacteroidales from Intestinal Biopsies of Children and Young Adults with Inflammatory Bowel Disease. *PLoS ONE.* **8**(6): e63686.

4.7. Supplementary material

Table S4.1: Testing repeatability of MHC-I (n = 26) and MHC-II (n = 24) genotyping for different dominant frequency thresholds while keeping minimum amplicon frequency at 0.3%. Threshold with greatest repeatability is in bold and underlined.

dominant frequency threshold (%)	MHC-I		MHC-II	
	Repeatability (%)	average number of alleles	Repeatability (%)	average number of alleles
5	77.83	10.17	78.63	29.00
10	77.67	10.13	84.38	34.42
15	77.48	8.79	85.58	36.31
20	77.89	7.85	86.09	36.52
25 (default)	**78.69**	7.48	**86.37**	36.44

Table S4.2: Testing repeatability of MHC-I (*n* = 26) and MHC-II (*n* = 24) genotyping for different minimum amplicon frequencies while keeping minimum dominant frequency threshold at 25%. Minimum amplicon frequency with the greatest repeatability for each MHC class is in bold and underlined. – indicates frequency not run.

Minimum amplicon frequency (%)	MHC-I		MHC-II	
	Repeatability (%)	average number of alleles	Repeatability (%)	average number of alleles
0	70.79	13.85	91.09	58.94
0.1	70.74	10.58	88.12	54.19
0.2	74.97	8.96	87.99	42.81
0.3	78.69	7.48	86.37	36.44
0.4	83.54	6.77	81.94	30.52
0.5	86.45	6.31	82.03	25.27
0.6	87.01	6.13	84.00	21.31
0.7	90.93	5.79	84.01	18.00
0.8	91.58	5.71	85.05	15.88
0.9	91.46	5.52	84.19	14.10
1	90.9	5.44	83.93	12.46
1.5	94.11	5.06	87.18	7.71
1.6	**95.48**	**4.96**	-	-
1.7	94.78	4.92	88.87	6.73
1.8	93.03	4.87	**90.12**	**13.04**
1.9	-	-	89.70	12.79
2	93.34	4.75	87.80	6.02
3.00 (default)	90.3	4.13	79.89	3.50

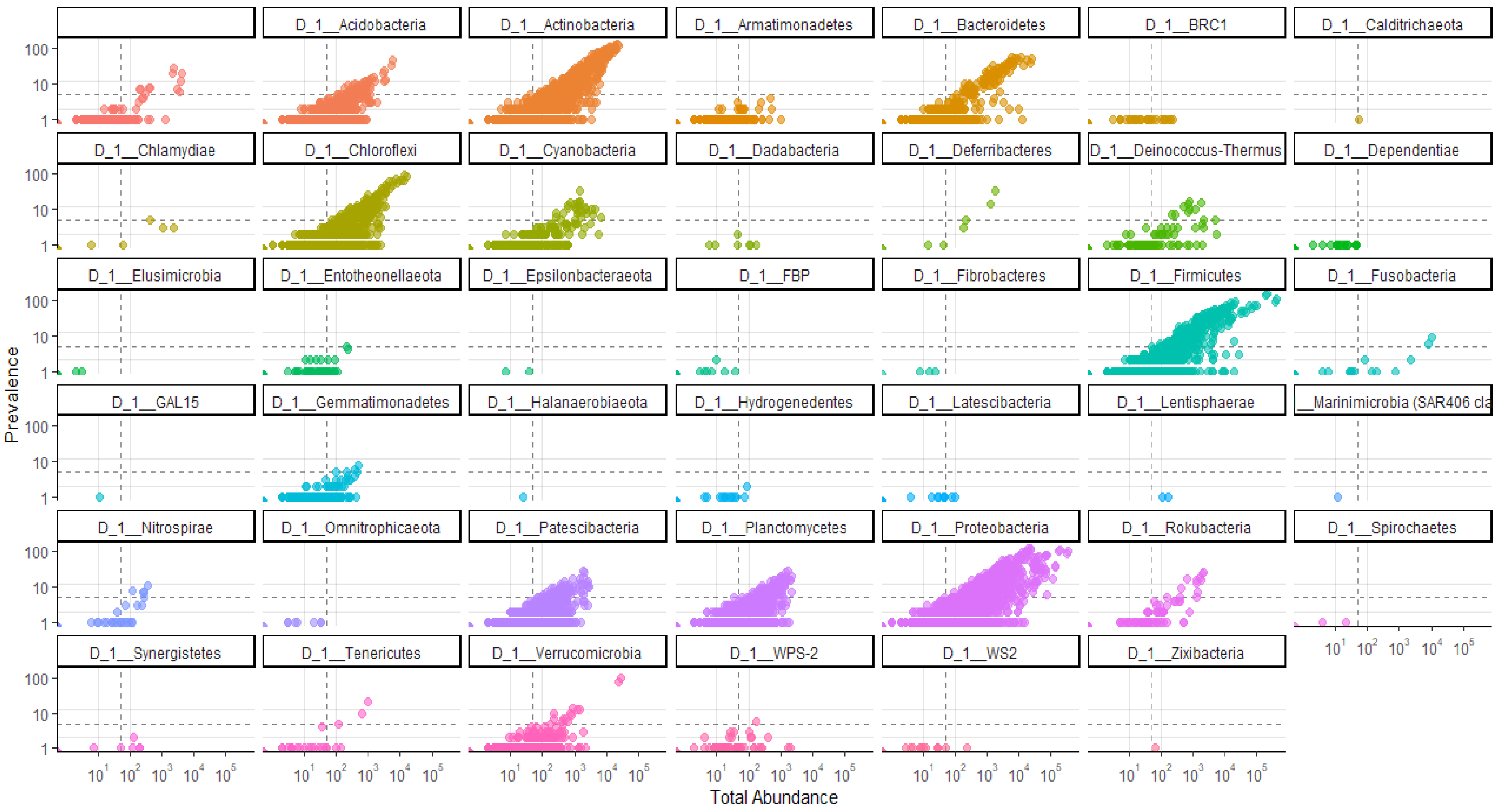
Table S4: The effect of host intrinsic and extrinsic variables on GM diversity in the Seychelles warbler (n = 195) for three metrics of alpha diversity: Shannon diversity, Chao 1(log transformed) and Faiths phylogenetic diversity (log transformed): **A**) including the presence/absence of MHC alleles or **B**) MHC diversity. A Linear model was used to generate conditional model-averaged estimates (β), their standard error (SE), z value, *p* value, and relative importance (ω) are shown for all predictors featuring in the top model set ($\Delta AIC_c \leq 7$). All continuous factors were standardised. Estimates are in reference to MHC allele = Absent *TLR3* genotype = $TLR3^{AA}$, sex = female, age class = fledgling, field period = Major 2017. Significant terms are in bold and underlined. *** $P < 0.001$, ** $P < 0.01$, * $P < 0.05$.

Model	Factor	Shannon						log Chao1						log Faiths Phylogenetic diversity					
		ω	β	SE	z	p		ω	β	SE	z	p		ω	β	SE	z	p	
A) Presence /absence of MHC alleles	(Intercept)		**4.98**	**0.35**	**14.04**	**<0.001**	***		**6.38**	**0.25**	**25.8**	**<0.001**	***		**3.38**	**0.17**	**19.51**	**<0.001**	***
	Field period: Major 2018	0.09	0.15	0.18	0.83	0.407		0.02	0.05	0.11	0.44	0.659		0.05	-0.07	0.07	0.99	0.324	
	Field period: Minor 2018		-0.15	0.29	0.51	0.61			-0.03	0.17	0.2	0.843			-0.05	0.12	0.38	0.703	
	Sex	0.28	-0.16	0.16	0.99	0.325		0.75	-0.18	0.1	1.86	0.062	.	0.73	-0.13	0.07	1.84	0.066	.
	Age: Old fledglings	**0.5**	**-0.59**	**0.28**	**2.09**	**0.036**	*	**0.88**	**-0.47**	**0.17**	**2.72**	**0.007**	**	0.18	-0.2	0.12	1.68	0.093	.
	Age: Sub-adults		-0.27	0.26	1.02	0.309			-0.13	0.16	0.82	0.41			-0.12	0.11	1.06	0.288	
	Age: Adults		-0.01	0.21	0.04	0.965			-0.05	0.13	0.41	0.682			-0.01	0.09	0.07	0.946	
	Individual Hs	**0.95**	**0.38**	**0.16**	**2.3**	**0.022**	*	0.36	0.12	0.1	1.19	0.233		0.44	0.09	0.07	1.35	0.178	
	TLR3 : AC	0.02	0.04	0.17	0.26	0.798		0.21	-0.11	0.1	1.02	0.307		0.08	-0.04	0.07	0.55	0.586	
	TLR3 : CC		-0.13	0.34	0.37	0.714			-0.32	0.21	1.53	0.126			-0.18	0.14	1.23	0.219	
	Ase-ua1/10	0.3	-0.27	0.26	1.06	0.292		0.39	-0.22	0.16	1.34	0.179		0.55	-0.18	0.11	1.58	0.115	
	Ase-ua3	**0.87**	**-0.54**	**0.25**	**2.16**	**0.031**	*	**0.87**	**-0.39**	**0.17**	**2.23**	**0.026**	*	0.49	-0.21	0.14	1.54	0.124	
	Ase-ua4	0.38	-0.32	0.26	1.22	0.223		**0.84**	**-0.32**	**0.16**	**2.04**	**0.042**	*	0.75	-0.22	0.12	1.89	0.059	.
	Ase-ua5	**0.94**	**-0.56**	**0.25**	**2.17**	**0.03**	*	**0.87**	**-0.31**	**0.15**	**2.13**	**0.033**	*	**0.9**	**-0.23**	**0.11**	**2.1**	**0.036**	*
	Ase-ua6	0.21	0.16	0.34	0.48	0.63		0.23	-0.12	0.19	0.63	0.529		0.2	0.01	0.14	0.07	0.941	
	Ase-ua7	0.17	0.02	0.32	0.08	0.939		0.37	-0.23	0.19	1.24	0.215		0.38	-0.16	0.14	1.18	0.238	
	Ase-ua8	0.16	-0.09	0.22	0.42	0.675		0.49	-0.21	0.14	1.52	0.13		0.26	-0.09	0.1	0.89	0.371	
	Ase-ua9	0.24	-0.23	0.3	0.76	0.447		0.24	-0.15	0.19	0.75	0.451		0.64	-0.21	0.12	1.77	0.077	.
	Ase-ua11	0.16	-0.02	0.26	0.09	0.925		0.15	-0.02	0.17	0.12	0.903		0.18	0.02	0.12	0.2	0.842	
	Ase-dab3	0.21	0.16	0.22	0.71	0.481		0.22	-0.13	0.15	0.88	0.379		0.23	0.08	0.1	0.76	0.45	
	Ase-dab4	**0.79**	**-0.47**	**0.24**	**1.96**	**0.05**	*	0.25	-0.13	0.15	0.88	0.378		0.55	-0.16	0.1	1.56	0.119	
	Ase-dab5	0.15	0.02	0.24	0.1	0.92		0.21	0.11	0.13	0.86	0.389		0.16	-0.01	0.1	0.14	0.891	
B) MHC Diversity	(Intercept)		**4.16**	**0.17**	**24.89**	**<0.001**	***		**5.74**	**0.13**	**44.84**	**<0.001**	***		**2.95**	**0.07**	**42.4**	**<0.001**	***
	Field period: Major 2018	0.15	0.16	0.18	0.89	0.373		0.08	0.05	0.11	0.42	0.674		0.12	-0.07	0.07	0.89	0.376	
	Field period: Minor 2018		-0.08	0.29	0.28	0.779			-0.02	0.18	0.1	0.92			-0.01	0.12	0.12	0.907	
	Sex	0.28	-0.11	0.16	0.68	0.499		0.54	-0.15	0.1	1.54	0.124		0.53	-0.1	0.07	1.52	0.13	
	Age: Old fledglings	0.46	-0.56	0.28	1.96	0.051	.	**0.72**	**-0.44**	**0.17**	**2.59**	**0.01**	**	0.23	-0.18	0.12	1.51	0.132	
	Age: Sub-adults		-0.33	0.26	1.27	0.206			-0.13	0.16	0.85	0.396			-0.11	0.11	1.05	0.295	
	Age: Adults		-0.02	0.21	0.09	0.926			-0.04	0.12	0.31	0.755			0.01	0.09	0.09	0.932	
	Individual Hs	**0.76**	**0.33**	**0.16**	**2**	**0.045**	*	0.4	0.11	0.1	1.18	0.239		0.42	0.08	0.07	1.24	0.216	
	TLR3 : AC	0.06	-0.03	0.17	0.16	0.871		0.47	-0.17	0.1	1.61	0.107		0.29	-0.08	0.07	1.14	0.254	
	TLR3 : CC		-0.1	0.34	0.29	0.771			-0.3	0.2	1.48	0.139			-0.19	0.14	1.34	0.181	
	MHC-I diversity	0.51	-0.2	0.66	0.31	0.76		0.5	-0.17	0.44	0.39	0.697		0.49	0.07	0.44	0.17	0.869	
	MHC-I $diversity^2$	0.61	-0.51	0.55	0.92	0.36		0.66	-0.4	0.33	1.2	0.229		0.76	-0.38	0.31	1.23	0.22	
	MHC-II diversity	0.45	-0.11	0.62	0.17	0.864		0.46	0.23	0.56	0.41	0.679		0.61	0.41	0.47	0.87	0.385	
	MHC-II $diversity^2$	0.54	-0.43	0.51	0.83	0.407		0.58	-0.42	0.46	0.9	0.366		0.82	-0.49	0.41	1.21	0.226	

Table S4.4: PERMANOVA tests investigating the effect of host variables on the composition of the Seychelles warbler gut microbiome (*n* = 195). Unweighted and weighted UniFrac were used as beta diversity metrics in separate models. **A)** including the presence/absence of MHC alleles or, **B**) MHC diversity. Significant terms are in bold and underlined. ** *P* < 0.01, * *P* < 0.05.

Model	Factors	Unweighted UniFrac					Weighted UniFrac				
		Df	R^2	F	*P*		Df	R^2	F	*P*	
A) Presence/absence of MHC alleles	Field period	**2**	**0.017**	**1.69**	**0.002**	**	**2**	**0.02**	**2.04**	**0.042**	*
	Sex	1	0.006	1.26	0.105		1	0.005	0.99	0.352	
	Age class	**3**	**0.019**	**1.26**	**0.035**	*	3	0.018	1.22	0.231	
	Individual Hs	1	0.006	1.22	0.129		1	0.009	1.85	0.096	.
	TLR3	2	0.011	1.04	0.337		2	0.01	0.99	0.39	
	Ase-ua1/10	1	0.005	1.04	0.339		**1**	**0.014**	**2.81**	**0.03**	*
	Ase-ua3	1	0.005	0.93	0.556		1	0.008	1.56	0.147	
	Ase-ua4	1	0.006	1.11	0.227		1	0.006	1.29	0.225	
	Ase-ua5	1	0.005	0.92	0.59		1	0.003	0.69	0.611	
	Ase-ua6	1	0.005	1.04	0.339		1	0.003	0.7	0.597	
	Ase-ua7	1	0.005	1.07	0.285		**1**	**0.017**	**3.47**	**0.012**	*
	Ase-ua9	1	0.004	0.88	0.682		1	0.004	0.76	0.529	
	Ase-ua8	1	0.004	0.79	0.883		1	0.004	0.71	0.581	
	Ase-ua11	**1**	**0.008**	**1.54**	**0.024**	*	**1**	**0.016**	**3.28**	**0.015**	*
	Ase-dab3	1	0.006	1.14	0.199		1	0.005	1.01	0.344	
	Ase-dab4	1	0.004	0.77	0.921		1	0.002	0.44	0.889	
	Ase-dab5	1	0.006	1.25	0.11		1	0.007	1.34	0.202	
	Residual	173	0.876				173	0.864			
	Total	194	1				194	1			
B) MHC Diversity	Field period	**2**	**0.017**	**1.72**	**0.002**	**	2	0.020	1.96	0.052	.
	Sex	1	0.006	1.19	0.149		1	0.005	1.07	0.3031	
	Age class	3	0.018	1.21	0.063	.	3	0.017	1.14	0.2879	
	Individual Hs	1	0.006	1.24	0.116		1	0.009	1.81	0.1028	
	TLR3	2	0.011	1.08	0.234		2	0.010	1.03	0.3629	
	MHC-I diversity	1	0.005	1.04	0.339		1	0.004	0.87	0.4341	
	MHC-I $diversity^2$	1	0.005	1.08	0.271		1	0.004	0.88	0.4271	
	MHC-II diversity	1	0.004	0.89	0.668		1	0.003	0.60	0.7013	
	MHC-II $diversity^2$	1	0.005	0.91	0.62		1	0.003	0.57	0.7304	
	Residual	181	0.919				181	0.920			
	Total	194	1				194	1			

Fig S4.1: Prevalence and total abundance of all ASV's separated by phylum, each phylum is shown in a separate plot, and a different colour. Dashed lines represent the values used as cut-offs for filtering rare taxa before beta diversity analyses (prevalence = 0.5; minimum abundance = 50).



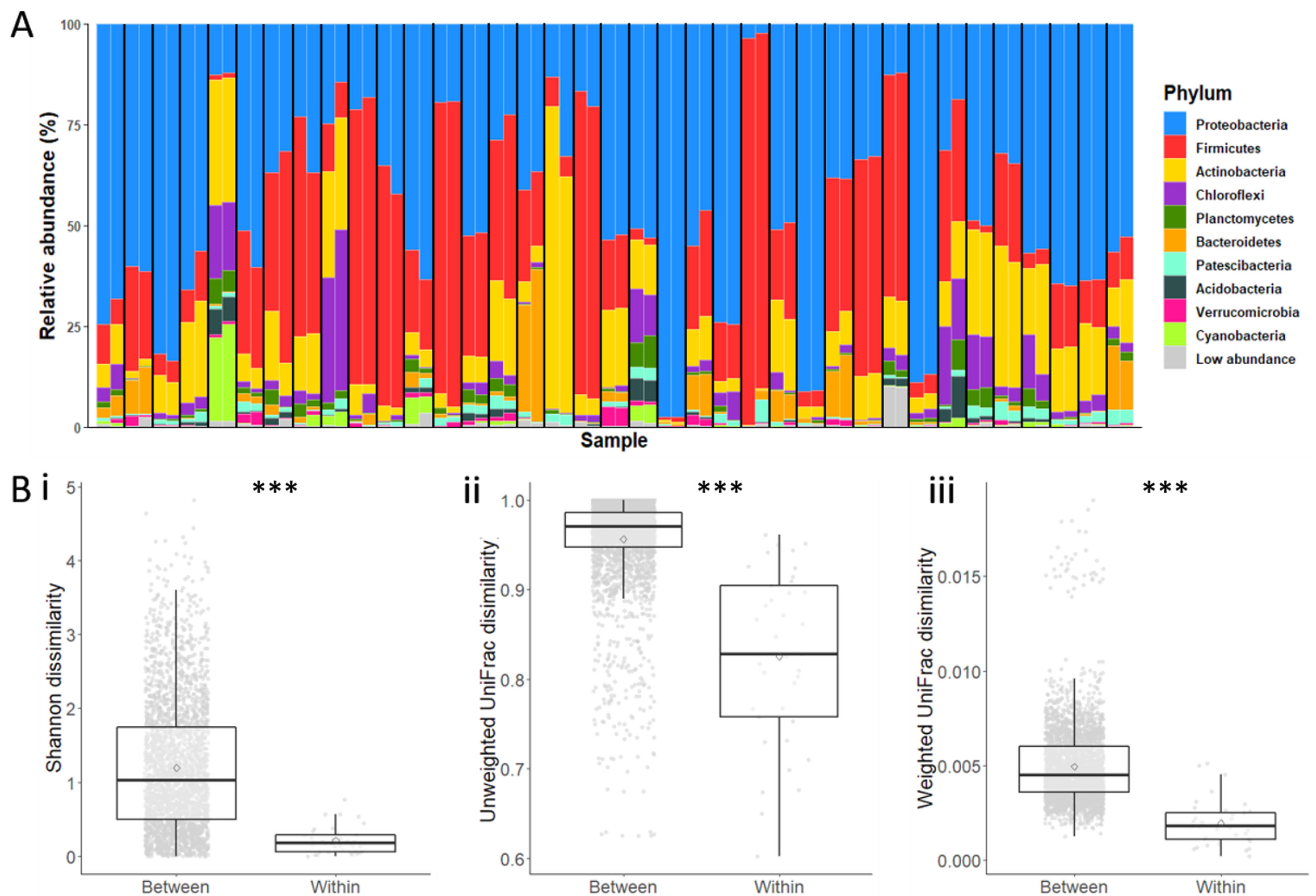


Fig S4.2: Sequencing repeatability of the gut microbiome tested using 37 faecal samples taken from Seychelles warblers, sequenced twice. **A**) Relative abundance (%) of the 10 most abundant taxa at the Phylum level for the 37 duplicated samples. Each column represents one sample, black lines separate duplicated samples. All other taxa within each sample are collapsed into the low abundance category. **B**) The pairwise dissimilarity between different samples, versus within pairs of duplicated samples (same DNA sequenced twice) for **i**. Shannon dissimilarity **ii**. Unweighted uniFrac dissimilarity and **iii**. Weighted uniFrac dissimilarity.

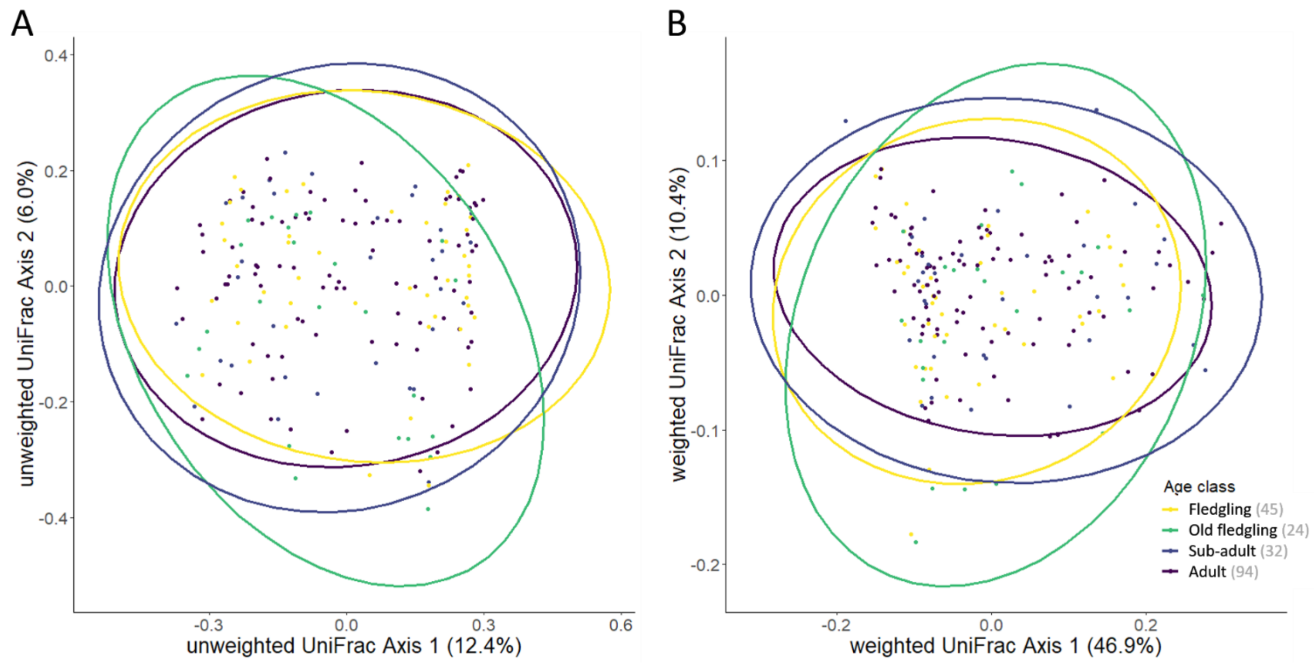


Fig S4.3: Beta diversity of Seychelles warbler gut microbiome composition in different age classes. The principal coordinate plots are based on **A)** unweighted UniFrac distances, and **B)** weighted UniFrac distances. Points represent a single faecal sample from a different individual (*n* = 195). Sample sizes are specified in brackets in the legend, and colours indicate the age class which was either fledgling (yellow), old-fledgling (green), sub-adult (indigo) and adult (purple). Ellipses represent a 95% confidence interval around the cluster centroids.

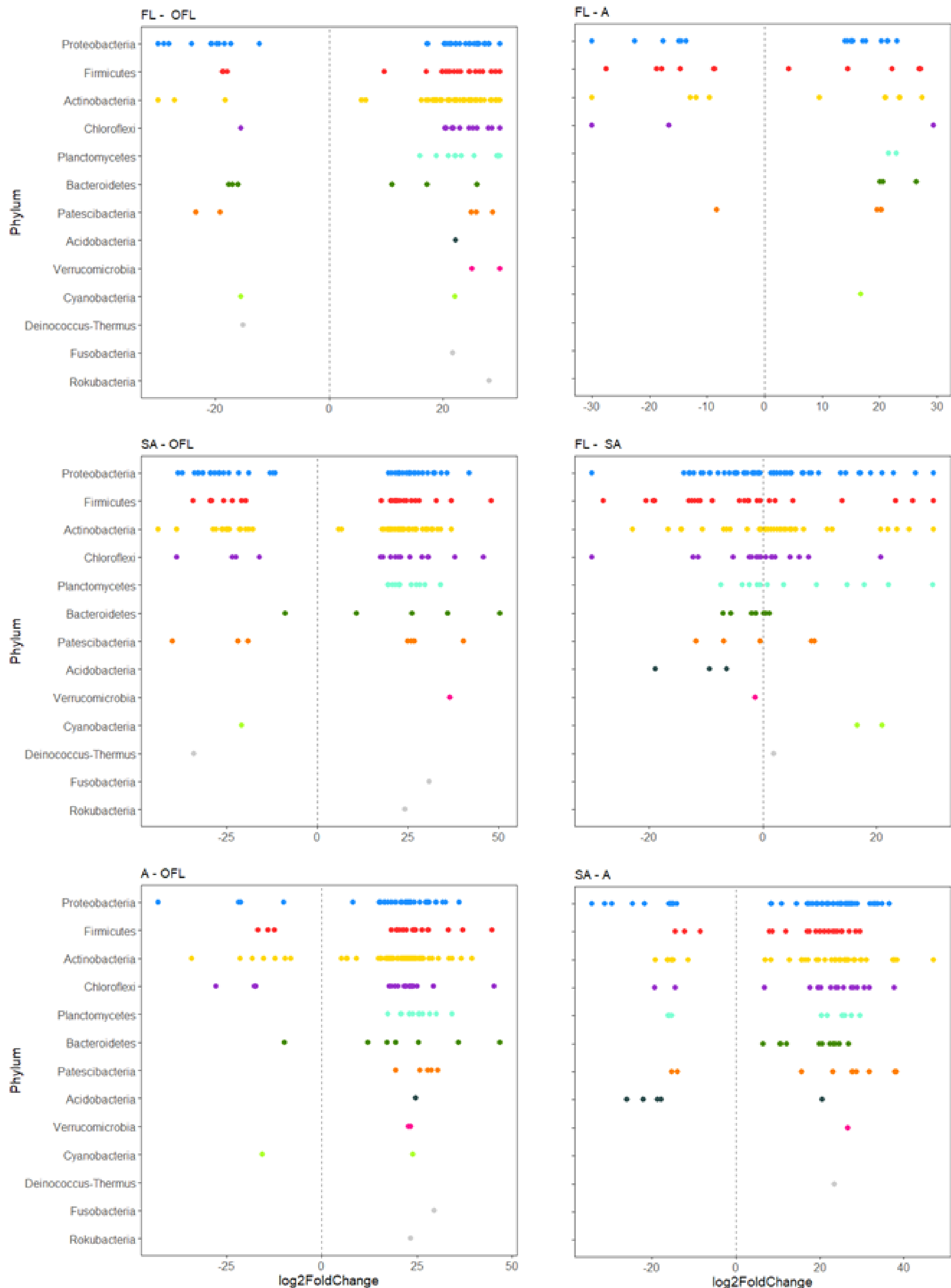


Fig S4.4: Differentially abundant ASV's (P_{adj} < 0.01) in the gut microbiome of Seychelles warblers between different age categories seasons. ASVs are grouped at the level of bacterial order and coloured according to bacterial phylum. ASVs shown with a $\log_2$ fold change greater than zero are significantly more abundant in seasons on the left and ASVs with a $\log_2$ fold change smaller than zero are significantly more abundant in seasons on the right.

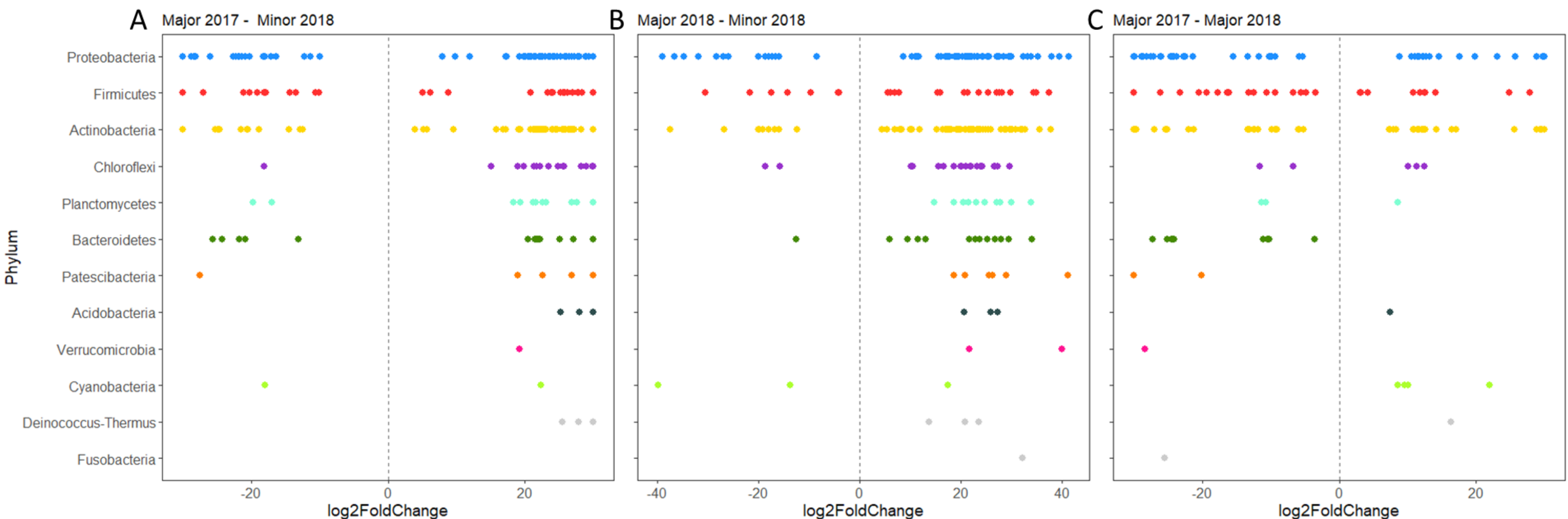


Fig S4.5: Differentially abundant ASV's ($P_{adj} < 0.01$) in the gut microbiome of Seychelles warblers, between seasons. Comparisons are **A)** Major 2017 vs Minor 2018, **B)** Major 2018 vs Minor 2017, or **C)** Major 2017 vs Major 2018. ASVs are grouped at the level of bacterial order and coloured according to bacterial phylum. ASVs shown with a $\log_2$ fold change greater than zero are significantly more abundant in seasons on the left and ASVs with a $\log_2$ fold change smaller than zero are significantly more abundant in seasons on the right.

Chapter 5

General discussion

Blood moon eclipse from Cousin, 2018

5.1. Thesis synthesis

The overall objective of this thesis has been to investigate the forces that shape immunogenetic variation in small populations, utilizing data from the long-term, intensively monitored Seychelles warbler (*Acrocephalus sechellensis)* system. First, focusing on a functional variant at the Toll-like receptor 3 (*TLR3)* locus, I asked whether the allele frequencies changed over time and space, and whether selection was driving any such evolution (Chapter 2). Second, based on my results from Chapter 2 (discussed below), I tested whether pre- and/or post-copulatory sexual selection was also acting antagonistically or synergistically with natural selection to shape *TLR3* variation in the population of Seychelles warblers on Cousin (Chapter 3). Third, I investigated whether host-microbiome coevolution may play a role in maintaining functional immunogenetic variation in this same population. Specifically, I looked at the potential role of key candidate immune genes that are important in the Seychelles warbler, i.e., Major histocompatibility complex (MHC) genes and the *TLR3* gene, in determining characteristics of the gut microbiome (GM) (Chapter 4). In the current final chapter, I discuss the broad implications of my findings, and outline some directions for future research.

5.2. Understanding immunogenetic variation in the Seychelles warbler

Throughout this thesis, I have used the candidate gene approach to investigate the role of different evolutionary forces in shaping genetic diversity. I specifically focused on immune genes which still harbour functional variation in the Seychelles warbler, despite the population bottleneck this species has been through (Richardson and Westerdahl 2003, Gilroy et al. 2017). However, first I had to develop efficient methods to characterise variation at these loci.

In Chapters 2-4 I used a single SNP assay to investigate variation at the viral sensing *TLR3* locus (Barton 2007). This SNP was previously found be the only functionally relevant TLR SNP showing a signature of past positive selection, and with a minor allele frequency sufficient for further analysis, in the Seychelles warbler (Gilroy et al. 2017). In my thesis, use of the single SNP assays allowed a large number of individuals (ca 1600) to be rapidly sequenced at relatively low cost. Thus, *TLR3* variation could be combined

with the long-term, individual-based dataset, including fitness component measures, to investigate evolutionary processes in depth.

In Chapter 4, I used next-generation sequencing (NGS) to characterise MHC class I (MHC-I) and class II (MHC-II) genes in the same individuals. Importantly, I found MHC-I genotyping using NGS to be highly consistent with previously sequenced alleles (98% similarity (Richardson and Westerdahl 2003)), and samples genotyped using reference strand-mediated conformation analysis (95% similarity; Wright et al. 2014). Reproducibility is important when comparing results that are based on genotypes derived using different protocols, for example, the survival results from Chapter 2 and (Brouwer et al. 2010), and changes in GM diversity (Chapter 4). Consistent with other studies in passerines (Minias et al. 2018), I found reduced functional allelic diversity at MHC-II, compared to MHC-I in the Seychelles warbler. Indeed, two of the MHC-II alleles were present in virtually every individual. Intriguingly, I also found evidence of several non-functional MHC-II alleles, which were present at high frequency. Although this is not completely unexpected in passerines (Westerdahl 2007), this does suggest that genetic drift could be playing a larger role in driving variation at MHC-II genes, compared to MHC-I genes.

Although I targeted MHC regions of known importance (Richardson and Westerdahl 2003), use of specific primers which only amplify a short fragment from each gene mean that some putatively functional MHC alleles may be missed. Additionally, high levels of MHC gene duplication and conversion in passerines mean that alleles cannot be assigned to specific loci using short-read amplicon sequencing. Together this hampers the accurate estimation of functional allele numbers, and thus diversity (heterozygosity) (reviewed in O'Connor et al. 2019). Using long-read and linked-read sequencing technologies such as PacBio and Oxford Nanopore would allow full-length MHC alleles to be sequenced, and fully characterised (e.g., Westbrook et al. 2015, He et al. 2020). Utilising this approach in the Seychelles warbler, would aid knowledge of the structure of MHC genes, and allow accurate locus-specific genotyping, thus improving our understanding of how variation is maintained by selection – including at the non-functional genes found in MHC-II (Chapter 4). This could also give further insight into how evolutionary process such as gene conversion have shaped the MHC.

Throughout this thesis, I have used a bottom-up approach, based on candidate genes we knew had functional variation in this species. However, there are many other genes involved in the immune system which could also be under selection (Acevedo-Whitehouse and Cunningham 2006, Ekblom et al. 2010). While it is not feasible to

sequence all these immune genes using the methods utilized in this thesis; low-coverage whole-genome resequencing would capture a large number of SNPS throughout the genome. This is currently being done for ~1500 individuals as part of a different project in the Seychelles warbler. A top-down approach using SNPs identified across the entire genome would provide greater resolution on relative contributions different evolutionary forces exert on each SNP (Monnahan et al. 2021). For example, utilizing the extensive pedigree (Sparks et al. 2020) and conducting gene dropping simulations on all SNPs, would provide insights to evolutionary forces governing temporal allele frequency dynamics (as in Chen et al. 2019). This approach would allow identification of specific alleles which are changing more than expected (i.e., are also under contemporary selection, not just drift). Once identified, selected loci could then be correlated with temporal and/or spatial change, and associations with various fitness metrics (e.g., survival, reproductive success) and selective pressures (e.g., gut microbiome and malaria infection), to try and disentangle which evolutionary forces and mechanisms could be shaping variation.

5.3. Contemporary selection of *TLR3*

While various studies have identified signatures of past selection at key immune genes (such as the MHC), very few have mapped how evolutionary forces are shaping genetic variation over multiple generations in contemporary wild populations (but see Chen et al. 2019). In Chapter 2, I track evolution of the *TLR3* variants across ca 10 generations in the Seychelles warbler. The evidence shows that the minor *TLR3*^{C} allele consistently declined in frequency over a 25-year period in the Cousin Island population. Furthermore, similar temporal declines occurred in the four other warbler populations that have been established by translocation from Cousin island (Komdeur 1994, Richardson et al. 2006, Wright et al. 2014). Intriguingly, the rate of *TLR3*^{C} allele frequency change between island populations appears to differ – although sample size is small. The rate of change in the *TLR3* alleles was weaker on Aride and Cousine, compared to Cousin, Frégate and Denis. Why the rate of change of this TLR3 allele differs between islands is not known. Indeed, only when we know what causes the contemporary evolution of this locus on Cousin (see below) will it be possible to then determine why it may differ between islands.

I then tested whether selection is currently acting on this locus in the Cousin population. I found that *TLR3*^{CC} individuals of both sexes had a 37% increase in overall mortality risk, and that this difference in mortality was largely in the first year of life before reaching

adulthood. Thus, the *TLR3*^{A} allele conferred a viability advantage. Indeed, *TLR3* genotype had the largest effect upon survival of the factors tested in this analysis, including MHC-1 diversity, and presence of the *Ase-ua4* allele – both previously been shown to be important determinants of survival in this population (Brouwer et al. 2010). But what could be driving this?

To date - and despite extensive screening - only one pathogen (a single strain of intracellular protozoan blood parasite - *Haemoproteus nucleocondensus* (GRW1)) has been observed in any Seychelles warbler populations (Hutchings 2009, Fairfield et al. 2016). Furthermore, I found no association between *TLR3* genotypes and the bacteria in the gut (Chapter 4). However, given *TLR3's* role as a specific receptor for viral dsRNA (Barton 2007), genotypic differences in survival are likely due to differential recognition of a viral pathogen. Yet there is no evidence of any virus circulating in the Seychelles warblers, although Influenza A has been reported in Procellariformes (petrels and shearwaters) in the Seychelles (Lebarbenchon et al. 2015). Dead warblers are rarely found, and no visible signs of disease have been observed over the entire 35+ years this bird has been studied in the Seychelles. Nonetheless, individuals may be asymptomatic, or only show visible symptoms during the acute phase of infection when they are also least active and least likely to be caught or observed (LaPointe et al. 2009). Therefore, the exact nature of the selective factor remains unknown. In the absence of any obvious symptoms, virome screening (see below) may enable us to determine what viruses may drive the *TLR3* selection observed in Chapter 2. This could also provide insight into what could be maintaining MHC-I variation (Chapter 2 and 4). Currently virome analysis is difficult for a range of reasons, including the absence of universal primers, difficulty in nucleic acid extraction and lack of comprehensive viral databases (reviewed in Garmaeva et al. 2019). However, with the fast pace of genomic technologies this is becoming an important avenue of future research. For example, using metagenomic, or metatranscriptomic sequencing would enable characterisation of the faecal or blood virome (see François and Pybus 2020, for recommendations), and has been successfully used in wild species (Ramírez-Martínez et al. 2018, Vibin et al. 2018, Wang et al. 2019, Bergner et al. 2020).

Another key result I found was that males - but not females - that were heterozygous (*TLR3*^{AC}) had a reduced rate of reproduction. Furthermore, combining the survival and reproductive rate results, males that were *TLR3*^{AA} produced double the number of offspring over their lifetimes than males with at least one copy of the *TLR3*^{C} allele.
This difference in LRS could be due to offspring mortality – as we know that *TLR3*^{CC} offspring that have reduced survival are never produced by *TLR3*^{AA} males. Unfortunately,

nests on Cousin are generally too high to reach. Therefore, it is difficult to accurately determine pre-fledgling mortality in this system. However, as there were no *TLR3* related differences in female lifetime reproductive success (LRS), offspring mortality is highly unlikely to be the sole cause of reduced male LRS. Instead, the differential LRS in males is likely due to another mechanism – for example sexual selection (Chapter 3). Alternatively, it is possible that *TLR3*^{AA} males are better at providing parental care. This could be resolved using social and extra-pair offspring and testing whether offspring fledgling or recruitment success is dependent on parental (genetic or social), or offspring *TLR3* genotype. Incorporating nest guarding, incubation and provisioning rates into this analysis would enable further insight into whether reproductive success is due to pre-copulatory choice, or early offspring mortality through reduced male parental care.

Taken together, the survival and reproductive success results provide solid evidence that strong positive selection is currently causing the evolution of the *TLR3* locus in the Seychelles warbler. In time this should lead to the loss of the *TLR3*^{C} allele in the Cousin population of Seychelles warbler, and indeed the species overall. However, it is surprising that this functional variant exists at all in the Seychelles warbler considering the severe bottleneck that this species suffered, which has resulted in decreased genome-wide genetic variation (Spurgin et al. 2014). Indeed, variation is limited or non-existent across many other immunogenetic loci (Gilroy et al. 2016, Gilroy et al. 2017). It may be that variation in the *TLR3* locus has been maintained through the bottleneck by balancing selection, but we are investigating too small a window of (evolutionary) time to detect this.

It is important to consider the different mechanism that could be acting to cause balancing selection in the Seychelles warbler. Based on the above survival and reproductive success results, there is no evidence that heterozygote advantage (Doherty and Zinkernagel 1975) is driving variation at this locus. But what about rare allele advantage (Slade and McCallum 1992) and/or fluctuating selection (Hill et al. 1991)? Differentiating the relative importance of these two mechanisms in driving genetic variation, and separating them from other evolutionary mechanisms, is extremely complicated (Spurgin and Richardson 2010). To conclusively assign rare-allele or fluctuating selection as the mechanism driving *TLR3* variation would require an extensive temporal study, across multiple replicate populations. Under a rare-allele mechanism, different alleles would be selected for in different populations with the same pathogen. Selected alleles would fluctuate in frequency, with rare-alleles conferring resistance initially selected for, before the pathogen gains resistance and a different rare-allele is selected for. Under a fluctuating selection mechanism, spatio-temporal variation in

pathogen abundance would equate to spatio-temporal variation in alleles. To do so, the selective agent responsible needs to be identified (as already discussed), and its presence in time and space tracked in relation to *TLR3* variation, across populations. The present 25-year time period would need to be extended by including past, or future population samples of Seychelles warbler, to capture any potential change in the rate of *TLR3*^{C} allele frequency loss. Forward extrapolation from the current temporal pattern suggests that it will take ca 40 years before the *TLR3*^{C} allele reaches <5% frequency in the adult population on Cousin. Likewise, backwards extrapolation suggests that *TLR3* alleles were at roughly equal frequency in the mid-1970s. Previously, it has been possible to use museum samples (collected between 1876-1940) - from 26 warblers to examine pre-bottleneck diversity of microsatellite markers, MHC class I alleles (Spurgin et al., 2014), and avian β-defensin genes (Gilroy et al., 2016). In the future, it would be useful to sequence these samples to determine what *TLR3* variation existed prior to the bottleneck. A previous study by Gilroy et al. (2017) including six other closely related species only found the A variant at this site, thus suggesting that the *TLR3*^{A} allele may be ancestral, and the *TLR3*^{C} allele a recent mutation which increased, then decreased in frequency. But further phylogenetic analysis across a wide range of bird species would be needed to confirm this.

5.4. Is *TLR3* under sexual selection?

Given that the TLR3 locus is under viability selection and linked to differences in reproductive success (Chapter 2), it is possible that this locus could also be being shaped by sexual selection. Importantly, I found that TLR3 genotype conferred sex specific differences in reproductive success. Which, along with evidence of heterozygote deficiency in offspring could be explained by various mechanisms, including non-random mating. Therefore, in Chapter 3, I asked a suite of questions to test this idea. Based on the results from Chapter 2, I predicted that *TLR3*^{AA} males who had the greatest LRS would have an advantage in terms of sexual selection compared to *TLR3*^{AC} or *TLR3*^{CC} males, but that there would be no differences associated with female *TLR3* genotype.

I found no overall difference in likelihood of gaining a dominant breeding position based on *TLR3* genotype in either sex. However, *TLR3*^{AA} males in larger groups, and *TLR3*^{CC} males in smaller groups were more likely to gain a dominant position, with heterozygous *TLR3*^{AC} males showing an intermediate pattern. Male warblers need to become dominant to successfully breed, although female subordinates can gain reproductive success by co-breeding (Richardson et al. 2002, Raj Pant et al. 2019). Thus, male Seychelles

warblers have stronger intra-sexual competition for territories (Eikenaar et al. 2009), than do females. Although not significant, effect sizes provided indication that beneficial genetic characteristics (Chapter 3) could potentially be positively correlated with likelihood of males gaining a dominant position. In future, increasing sample size, and incorporating *TLR3* (and other genetic metrics) into analysis on dispersal strategy and success, could provide clarity on the importance of genetic characteristics in gaining a dominant breeding position. Indeed, studies based on the MHC have suggested that rather large sample sizes are usually needed to detect significant allele effects – particularly in wild systems (Kamiya et al. 2014, Gaigher et al. 2019)

In Chapter 2, a heterozygote deficiency was observed at the *TLR3* locus in young offspring. In Chapter 3 I found evidence that *TLR3*-based assortative social pairing might help explain this. Whether this heterozygote deficiency was caused by directional or assortative social-mate choice is unclear. $TLR3^{AA}$ individuals were more likely to pair together, than were $TLR3^{AC}$ females and $TLR3^{AA}$ males, but $TLR3^{CC}$ females exhibited random-pairing – which suggests that directional social mate choice is more likely.

Based on the analysis in this thesis, there is no apparent physiological difference between individuals of the two genotypes subject to directional social pairing ($TLR3^{AA}$ and $TLR3^{AC}$); both have similar survival probabilities (Chapter 2), and there were no clear differences in adult size or condition (Chapter 3). Therefore, the cue underlying this social choice, and indeed who is making the choice, is unknown. Gaining a finer understanding of phenotypic differences between *TLR3* individuals would be beneficial, and aid in understanding the selective benefits of these preferences. Non-lethal accumulative costs of a given trait can be detrimental in the long-term, and so be under selection. Biomarkers such as telomere length, which have been shown to be an accurate measure of somatic costs in the Seychelles warbler (Bebbington et al. 2016), could be used to provide a more precise proxy of phenotypic quality. Correlating non-lethal costs with *TLR3* and MHC variation could provide a stronger indication of the benefits of immune characteristics confer in the long-term, and if this is driven by morbidity effects. As immune gene variation should have a direct effect on immunocompetence, incorporating direct measures of baseline immune function (e.g., measuring complement and natural antibodies from plasma samples (Salvante 2006, Hegemann et al. 2017, Smyth et al. 2018) could also provide insight into the sub-lethal costs conferred by different immune genotypes (Gaigher et al. 2019). These measures would be particularly informative in looking at whether genes were sexually selected via condition dependent traits, where small differences between individuals could result in reduced reproductive success.

I found no evidence of genetic mate choice occurring; the likelihood of being cuckolded, and number annual extra-pair offspring were no different between male *TLR3* genotypes (Chapter 3). $TLR3^{CC}$ males did have increased mortality (Chapter 2) and produced fewer within-pair offspring (Chapter 3), which could explain their reduced LRS. But as $TLR3^{AC}$ males did not have different mortality, or within-pair offspring success compared to $TLR3^{AA}$ males, and there is no significant difference in male *TLR3* genotype in gaining a social position, or of gaining and losing paternity, I am still unable to fully explain the reduced reproductive success results for $TLR3^{AC}$ males found in Chapter 2. It's worth noting that, although not significant, $TLR3^{AC}$ males tended to do slightly worse in gaining dominance, paternity, and were cuckolded more often, than $TLR3^{AA}$ males. It is possible the combined accumulation of slightly deleterious effects across life stages could cause the observed reduction in male $TLR3^{AC}$ reproductive success compared to other genotypes. But until more data is available, this is speculative.

Intriguingly, in Chapter 3 I found evidence that within-male post-copulatory sexual selection was occurring; when the genetic father assigned was heterozygous at the *TLR3* locus, offspring were significantly more likely to inherit the paternal $TLR3^{A}$ allele rather than the $TLR3^{C}$ allele. Along with the assortative mating found in Chapter 3 this preference could contribute to the heterozygote deficiency found in young offspring (Chapter 2). This bias in within-male paternal allele inheritance may be due to cryptic mate-choice (Eberhard 1996) or sperm competition (Birkhead and Møller 1998). There is now a wider appreciation of the role that gametic selection – including within-males - may play in organisms (reviewed in Sutter and Immler 2020). Within-male variation in sperm can be expressed via differential longevity, immunity, performance, and morphology (Holt and Look 2004), which could be due to genotypic differences. However, the mechanism behind the result in the Seychelles warbler is unknown. One interesting avenue of future research is the reproductive microbiome, which can influence post-copulatory sexual selection through a range of mechanisms (reviewed in Rowe et al. 2020). Of particular relevance here, sexually transmitted infections can affect sperm success (Wigby et al. 2019). If $TLR3^{C}$ sperm are either of worse quality, or more vulnerable to pathogens in the reproductive tract, then this could explain the allelic bias observed. Future work could investigate this through comprehensive analysis of the reproductive microbiome (including the virome – using techniques discussed in 5.2 and 5.4) which could indicate presence of pathogens. Correspondingly, as it is possible to collect sperm from passerines via various methods (Wolfson 1952, Immler and Birkhead 2005), differences in morphology and longevity from *TLR3* homozygous males could be quantified– although transporting these samples from field to lab before degredation occurs might be tricky.

5.5. Could the gut microbiome be a selective pressure?

There is now considerable evidence showing that selection is acting to shape immunogenetic variation in the Seychelles warbler e.g., at the MHC (Richardson and Westerdahl 2003, Brouwer et al. 2010) and the *TLR3* (Chapters 2 and 3), but which pathogen(s) are acting as the selective agent in the Seychelles warbler remains unknown. Variation in the GM directly affect hosts fitness in many ways, for example through proliferation of pathogenic microbes (Uddin et al. 2017), or by enabling nutrient uptake (Hooper et al. 2002). Therefore, the GM has the potential to act as a selective pressure (Amato 2013). In Chapter 4 I found associations between specific MHC alleles and gut microbiome characteristics in the Seychelles warbler. The presence of specific MHC alleles was associated with a reduction in GM diversity, and the differential abundance of certain bacteria. This is consistent with the few other studies (Bolnick et al. 2014, Hernández-Gómez et al. 2018) that have investigated associations between MHC and the GM in wild systems – although these were only based on MHC-II diversity.

Presence of associations between specific MHC alleles and the GM suggests that the GM could act as a selective pressure in shaping immuno-genetic vacation in the Seychelles warbler. Investigating fitness associations with the GM is needed to test the GMs role as a possible selective pressure – this work is already being conducted. In future, it would be beneficial to test for direct, or indirect associations between all three components (immune genes, GM, and fitness), this would help place associations into context – e.g., we found presence of MHC alleles resulted in reduced GM diversity – but is this detrimental or beneficial to the host?

MHC-I molecules mainly recognise intracellular pathogens, while MHC-II molecules mainly recognise extracellular microbes such as bacteria (Hughes and Yeager 1998). Thus, it was unexpected that most of the associations I found between GM and immunogenetic variation were with MHC-I alleles, although the presence of one MHC-II (*Ase-dab4*) was associated with reduction in GM diversity. It is possible the MHC-GM interactions I have detected may be caused by indirect effects e.g., reduced health or condition. Given that MHC-I do not usually recognise extracellular microbes, this is likely the case.

As well as virome screening, investigating associations between MHC-I and MHC-II variation and malaria infection could provide resolution on which alleles may result in susceptibility to infection – and thus be detrimental. This would have the dual purpose of providing insight into whether pathogen-mediating selection could be shaping diversity

at these loci (reviewed in Westerdahl, Asghar et al. 2011). Although presence of this malarial strain does not result in reduced survival in the Seychelles warbler (Hammers et al. 2016), it has been linked to more rapid telomere shortening (Brown et al. 2021). Suggesting it does have some impact on individual condition and fitness across the lifespan - as shown in a close congener (Asghar et al. 2015). If this is the case in the Seychelles warbler, then malaria infection could detrimentally impact the GM (Mukherjee et al. 2020). However, the screening currently done in the Seychelles warbler only tests whether an individual is infected or not, rather than the severity of infection. In related species, it has been possible to use qPCR to successfully measure intensity, as well as prevalence, of malaria infections, (Biedrzycka et al. 2015). By incorporating intensity, it is possible to separate out those birds which are susceptible and those which have a quantitative form of resistance (i.e., are infected but able to limit any deleterious effects selection, Westerdahl et al. 2011, Hayward et al. 2014). Variation at MHC-1 has previously been linked to malarial intensity (Westerdahl et al. 2013, Aguilar, Westerdahl et al. 2016, Biedrzycka et al. 2018) in other species, suggesting malaria could be an important selective pressure shaping variation at these genes. Thus, in future this would be preferred method for malaria screening in the Seychelles warbler.

In this thesis, I screened the bacteria of the GM, however this is only one taxonomic component of the microbiome. The GM is also comprised of archaea, viruses, and microbial eukaryotes (including fungi), all of which could be important for host biological functions and fitness (Harrison et al. 2020). In addition to bacteria and viruses, other components of the microbiome can also be important for the host. For example, many microbial eukaryotes such as nematodes are parasitic, (reviewed in Wegener Parfrey et al. 2011), and various fungi taxa have been related to host disease and mucosal health (reviewed in Huffnagle and Noverr 2013). Host-archaea associations also occur, although the role and importance of these are still unclear (reviewed in Borrel et al. 2020).

While most research has focused on the bacterial component, there are several methodologies which enable sequencing of the whole microbiome (reviewed in Knight et al. 2018, Vemuri et al. 2020). Primers are already optimised for metabarcoding marker genes in different taxa e.g., 18S rRNA for eukaryotic microbes, internal transcribed spacer for fungi, and 16S rRNA for bacteria and archaea (e.g., Kittelmann et al. 2013, Hoggard et al. 2018). An alternate to the metagenetic approach is whole genome sequencing. This would capture the entire diversity of the microbiome, enabling greater taxonomic resolution. This would also allow functional information to be collected (Quince et al. 2017). As well as providing a fuller knowledge of the impact changes in GM could infer on the host, this would allow identification of likely pathogenic microbes.

While metagenomic and metatranscriptomic analysis of microbiomes is becoming increasingly popular, this technology is still very costly and complicated. In future, a combination of metagenetic and metagenomic approaches may be recommended. For example, sequencing a large number of samples at targeted GM marker genes and using integrated statistical approaches (e.g., Sweeny et al. 2020) would elucidate which host factors are associated with changes in the GM – such as key alleles e.g., *Ase-ua1*, or host factors e.g., age (as seen in Chapter 4), and the direction of these broad scale changes Based on these results, targeted metagenomic sequencing of the GM could be used to help disentangle significance of host traits in shaping the GM e.g., comparing individuals from different genetic backgrounds, or comparing those that survived vs those that died, or comparing within-individual changes with age. Combining the various methodologies outlined above, with the detailed knowledge of life-history in the Seychelles warbler system would allow many exciting research questions to be investigated in the future (e.g., Antwis et al. 2017)

5.6. Conclusion

The Seychelles warbler as a long-term study system, provides a wealth of accurate individual-based data, without which it would not have been possible to study evolutionary processes shaping genetic variation in such detail. Utilizing this extensive dataset, I found evidence of contemporary evolution - through both viability and sexual selection - at one innate immune receptor (*TLR3*). I then characterised MHC-1, MHC-2 and the GM and found associations between specific alleles and diversity and composition of the GM, raising the possibility that the GM could act as a selective pressure, in turn shaping immunogenetic variation. Taken together, these results emphasize the importance of immune gene variation within a population. More generally, this thesis opens the way for future research into GM variation and its association with traits such as condition, senescence, and inbreeding, and ultimately its effect on fitness.

5.7. References

Acevedo-Whitehouse, K. and Cunningham, A. A. (2006). Is MHC enough for understanding wildlife immunogenetics? *Trends in Ecology & Evolution.* **21**(8): 433-438.

Aguilar, J. R.-d., Westerdahl, H., Puente, J. M.-d. I., Tomás, G., Martínez, J. and Merino, S. (2016). MHC-I provides both quantitative resistance and susceptibility to blood parasites in blue tits in the wild. *Journal of Avian Biology.* **47**(5): 669-677.

Amato Katherine, R. (2013). Co-evolution in context: The importance of studying gut microbiomes in wild animals. *Microbiome Science and Medicine.* **1**(1).

Antwis, R. E., Griffiths, S. M., Harrison, X. A., Aranega-Bou, P., Arce, A., Bettridge, A. S., . . . Sutherland, W. J. (2017). Fifty important research questions in microbial ecology. *FEMS microbiology ecology.* **93**(5).

Asghar, M., Hasselquist, D., Hansson, B., Zehtindjiev, P., Westerdahl, H. and Bensch, S. (2015). Hidden costs of infection: Chronic malaria accelerates telomere degradation and senescence in wild birds. *Science.* **347**(6220): 436-438.

Barton, G. M. (2007). Viral recognition by Toll-like receptors. *Seminars in Immunology.* **19**(1): 33-40.

Bebbington, K., Spurgin, L. G., Fairfield, E. A., Dugdale, H. L., Komdeur, J., Burke, T. and Richardson, D. S. (2016). Telomere length reveals cumulative individual and transgenerational inbreeding effects in a passerine bird. *Molecular Ecology.* **25**(12): 2949-2960.

Bergner, L. M., Orton, R. J., Benavides, J. A., Becker, D. J., Tello, C., Biek, R. and Streicker, D. G. (2020). Demographic and environmental drivers of metagenomic viral diversity in vampire bats. *Molecular Ecology.* **29**(1): 26-39.

Biedrzycka, A., Bielański, W., Ćmiel, A., Solarz, W., Zając, T., Migalska, M., . . . Radwan, J. (2018). Blood parasites shape extreme major histocompatibility complex diversity in a migratory passerine. *Molecular Ecology.* **27**(11): 2594-2603.

Biedrzycka, A., Migalska, M. and Bielański, W. (2015). A quantitative PCR protocol for detecting specific Haemoproteus lineages: molecular characterization of blood parasites in a Sedge Warbler population from southern Poland. *Journal of Ornithology.* **156**(1): 201-208.

Birkhead, T. R. and Møller, A. (1998). Sperm competition. *New York: Academic.*

Bolnick, D. I., Snowberg, L. K., Caporaso, J. G., Lauber, C., Knight, R. and Stutz, W. E. (2014). Major Histocompatibility Complex class IIb polymorphism influences gut microbiota composition and diversity. *Molecular Ecology.* **23**(19): 4831-4845.

Borrel, G., Brugère, J. F., Gribaldo, S., Schmitz, R. A. and Moissl-Eichinger, C. (2020). The host-associated archaeome. *Nat Rev Microbiol.* **18**(11): 622-636.

Brouwer, L., Barr, I., Van De Pol, M., Burke, T., Komdeur, J. and Richardson, D. S. (2010). MHC-dependent survival in a wild population: evidence for hidden genetic benefits gained through extra-pair fertilizations. *Molecular Ecology.* **19**(16): 3444-3455.

Brown, T. J., Dugdale, H., Spurgin, L., Komdeur, J., Burke, T. and Richardson, D. (2021). Causes and Consequences of Telomere Lengthening in a Wild Vertebrate Population. *Authorea Preprints.*

Chen, N., Juric, I., Cosgrove, E. J., Bowman, R., Fitzpatrick, J. W., Schoech, S. J., . . . Coop, G. (2019). Allele frequency dynamics in a pedigreed natural population. *Proceedings of the National Academy of Sciences.* **116**(6): 2158-2164.

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

Eberhard, W. (1996). *Female control: sexual selection by cryptic female choice.* Princeton University Press.

Eikenaar, C., Richardson, D. S., Brouwer, L., Bristol, R. and Komdeur, J. (2009). Experimental evaluation of sex differences in territory acquisition in a cooperatively breeding bird. *Behavioral Ecology.* **20**(1): 207-214.

Ekblom, R., French, L., Slate, J. and Burke, T. (2010). Evolutionary Analysis and Expression Profiling of Zebra Finch Immune Genes. *Genome Biology and Evolution.* **2**: 781-790.

Fairfield, E. A., Hutchings, K., Gilroy, D. L., Kingma, S. A., Burke, T., Komdeur, J. and Richardson, D. S. (2016). The impact of conservation-driven translocations on blood parasite prevalence in the Seychelles warbler. *Scientific Reports.* **6**.

François, S. and Pybus, O. G. (2020). Towards an understanding of the avian virome. *Journal of General Virology.* **101**(8): 785-790.

Gaigher, A., Burri, R., San-Jose, L. M., Roulin, A. and Fumagalli, L. (2019). Lack of statistical power as a major limitation in understanding MHC-mediated immunocompetence in wild vertebrate populations. *Molecular Ecology.* **28**(23): 5115-5132.

Garmaeva, S., Sinha, T., Kurilshikov, A., Fu, J., Wijmenga, C. and Zhernakova, A. (2019). Studying the gut virome in the metagenomic era: challenges and perspectives. *BMC Biology.* **17**(1): 84.

Gilroy, D., van Oosterhout, C., Komdeur, J. and Richardson, D. S. (2016). Avian β-defensin variation in bottlenecked populations: the Seychelles warbler and other congeners. *Conservation Genetics.* **17**(3): 661-674.

Gilroy, D., van Oosterhout, C., Komdeur, J. and Richardson, D. S. (2017). Toll-like receptor variation in the bottlenecked population of the endangered Seychelles warbler. *Animal Conservation.* **20**(3): 235-250.

Hammers, M., Komdeur, J., Kingma, S. A., Hutchings, K., Fairfield, E. A., Gilroy, D. L. and Richardson, D. S. (2016). Age-specific haemosporidian infection dynamics and survival in Seychelles warblers. *Scientific Reports.* **6**: 29720.

Harrison, X. A., McDevitt, A. D., Dunn, J. C., Griffiths, S., Benvenuto, C., Birtles, R., . . . Antwis, R. E. (2020). Host-associated fungal communities are determined by host phylogeny and exhibit widespread associations with the bacterial microbiome *bioRxiv.* 2020.2007.2007.177535.

Hayward, A. D., Nussey, D. H., Wilson, A. J., Berenos, C., Pilkington, J. G., Watt, K. A., . . . Graham, A. L. (2014). Natural Selection on Individual Variation in Tolerance of Gastrointestinal Nematode Infection. PLOS Biology. 12(7): e1001917..

He, K., Minias, P. and Dunn, P. O. (2020). Long-Read Genome Assemblies Reveal Extraordinary Variation in the Number and Structure of MHC Loci in Birds. *Genome Biology and Evolution.* **13**(2).

Hegemann, A., Pardal, S. and Matson, K. D. (2017). Indices of immune function used by ecologists are mostly unaffected by repeated freeze-thaw cycles and methodological deviations. *Frontiers in zoology.* **14**(1): 43.

Hernández-Gómez, O., Briggler, J. T. and Williams, R. N. (2018). Influence of immunogenetics, sex and body condition on the cutaneous microbial communities of two giant salamanders. *Molecular Ecology.* **27**(8): 1915-1929.

Hill, A. V. S., Allsopp, C. E. M., Kwiatkowski, D., Anstey, N. M., Twumasi, P., Rowe, P. A., . . . Greenwood, B. M. (1991). Common West African HLA antigens are associated with protection from severe malaria. *Nature.* **352**(6336): 595-600.

Hoggard, M., Vesty, A., Wong, G., Montgomery, J. M., Fourie, C., Douglas, R. G., . . . Taylor, M. W. (2018). Characterizing the Human Mycobiota: A Comparison of Small Subunit rRNA, ITS1, ITS2, and Large Subunit rRNA Genomic Targets. *Frontiers in Microbiology.* **9**(2208).

Holt, W. V. and Look, K. J. W. V. (2004). Concepts in sperm heterogeneity, sperm selection and sperm competition as biological foundations for laboratory tests of semen quality. *Reproduction.* **127**(5): 527.

Hooper, L. V., Midtvedt, T. and Gordon, J. I. (2002). How host-microbial interactions shape the nutrient environment of the mammalian intestine. *Annu Rev Nutr.* **22**: 283-307.

Huffnagle, G. B. and Noverr, M. C. (2013). The emerging world of the fungal microbiome. *Trends in microbiology.* **21**(7): 334-341.

Hughes, A. L. and Yeager, M. (1998). Natural selection at major histocompatibility complex loci of vertebrates. *Annual review of genetics.* **32**(1): 415-435.

Hutchings, K. (2009). *Parasite-mediated selection in an island endemic, the Seychelles warbler (Acrocephalus sechellensis)*. University of East Anglia.

Kamiya, T., O'Dwyer, K., Westerdahl, H., Senior, A. and Nakagawa, S. (2014). A quantitative review of MHC-based mating preference: the role of diversity and dissimilarity. *Molecular Ecology.* **23**(21): 5151-5163.

Kittelmann, S., Seedorf, H., Walters, W. A., Clemente, J. C., Knight, R., Gordon, J. I. and Janssen, P. H. (2013). Simultaneous Amplicon Sequencing to Explore Co-Occurrence Patterns of Bacterial, Archaeal and Eukaryotic Microorganisms in Rumen Microbial Communities. *PLoS ONE.* **8**(2): e47879.

Knight, R., Vrbanac, A., Taylor, B. C., Aksenov, A., Callewaert, C., Debelius, J., . . . Dorrestein, P. C. (2018). Best practices for analysing microbiomes. *Nature Reviews Microbiology.* **16**(7): 410-422.

Komdeur, J. (1994). Conserving the Seychelles warbler Acrocephalus sechellensis by translocation from Cousin Island to the islands of Aride and Cousine. *Biological Conservation.* **67**(2): 143-152.

LaPointe, D. A., Hofmeister, E. K., Atkinson, C. T., Porter, R. E. and Dusek, R. J. (2009). Experimental infection of Hawaii amakihi (Hemignathus virens) with West Nile virus and competence of a co-occurring vector, Culex quinquefasciatus: potential impacts on endemic Hawaiian avifauna. *Journal of wildlife diseases.* **45**(2): 257-271.

Lebarbenchon, C., Jaeger, A., Feare, C., Bastien, M., Dietrich, M., Larose, C., . . . Dellagi, K. (2015). Influenza A Virus on Oceanic Islands: Host and Viral Diversity in Seabirds in the Western Indian Ocean. *PLOS Pathogens.* **11**(5): e1004925.

Minias, P., Pikus, E., Whittingham, L. A. and Dunn, P. O. (2018). A global analysis of selection at the avian MHC. *Evolution.* **72**(6): 1278-1293.

Monnahan, P. J., Colicchio, J., Fishman, L., Macdonald, S. J. and Kelly, J. K. (2021). Predicting evolutionary change at the DNA level in a natural Mimulus population. *PLoS Genetics.* **17**(1): e1008945.

Mukherjee, D., Chora, Â. F. and Mota, M. M. (2020). Microbiota, a Third Player in the Host–Plasmodium Affair. *Trends in Parasitology.* **36**(1): 11-18.

O'Connor, E. A., Westerdahl, H., Burri, R. and Edwards, S. V. (2019). Avian MHC Evolution in the Era of Genomics: Phase 1.0. *Cells.* **8**(10): 1152.

Quince, C., Walker, A. W., Simpson, J. T., Loman, N. J. and Segata, N. (2017). Shotgun metagenomics, from sampling to analysis. *Nature Biotechnology.* **35**(9): 833-844.

Raj Pant, S., Komdeur, J., Burke, T. A., Dugdale, H. L. and Richardson, D. S. (2019). Socio-ecological conditions and female infidelity in the Seychelles warbler. *Behavioral Ecology.* **30**(5): 1254-1264.

Ramírez-Martínez, L. A., Loza-Rubio, E., Mosqueda, J., González-Garay, M. L. and García-Espinosa, G. (2018). Fecal virome composition of migratory wild duck species. *PLoS ONE.* **13**(11): e0206970.

Richardson, D. and Westerdahl, H. (2003). MHC diversity in two Acrocephalus species: the outbred Great reed warbler and the inbred Seychelles warbler. *Molecular Ecology.* **12**(12): 3523-3529.

Richardson, D. S., Bristol, R. and Shah, N. J. (2006). Translocation of the Seychelles warbler Acrocephalus sechellensis to establish a new population on Denis Island, Seychelles. *Conservation Evidence.* **3**: 54-57.

Richardson, D. S., Burke, T., Komdeur, J. and Dunn, P. (2002). Direct benefits and the evolution of female-biased cooperative breeding in Seychelles warblers. *Evolution.* **56**(11): 2313-2321.

Rowe, M., Veerus, L., Trosvik, P., Buckling, A. and Pizzari, T. (2020). The Reproductive Microbiome: An Emerging Driver of Sexual Selection, Sexual Conflict, Mating Systems, and Reproductive Isolation. *Trends in Ecology & Evolution.* **35**(3): 220-234.

Salvante, K. G. (2006). Techniques for Studying Integrated Immune Function in Birds. *The Auk.* **123**(2): 575-586.

Slade, R. and McCallum, H. (1992). Overdominant vs. frequency-dependent selection at MHC loci. *Genetics.* **132**(3): 861.

Smyth, K. N., Caruso, N. M., Davies, C. S., Clutton-Brock, T. H. and Drea, C. M. (2018). Social and endocrine correlates of immune function in meerkats: implications for the immunocompetence handicap hypothesis. *Royal Society Open Science.* **5**(8).

Sparks, A. M., Spurgin, L. G., van der Velde, M., Fairfield, E. A., Komdeur, J., Burke, T., . . . Dugdale, H. (2020). Telomere heritability and parental age at conception effects in a wild avian population.

Spurgin, L. G. and Richardson, D. S. (2010). How pathogens drive genetic diversity: MHC, mechanisms and misunderstandings. *Proceedings of the Royal Society B: Biological Sciences.* **277**(1684): 979-988.

Spurgin, L. G., Wright, D. J., van der Velde, M., Collar, N. J., Komdeur, J., Burke, T. and Richardson, D. S. (2014). Museum DNA reveals the demographic history of the endangered Seychelles warbler. *Evolutionary Applications.* **7**(9): 1134-1143.

Sutter, A. and Immler, S. (2020). Within-ejaculate sperm competition. *Philosophical Transactions of the Royal Society B: Biological Sciences.* **375**(1813): 20200066.

Sweeny, A. R., Lemon, H., Ibrahim, A., Nussey, D. H., Free, A. and McNally, L. (2020). A mixed model approach for estimating drivers of microbiota community composition and differential taxonomic abundance. *bioRxiv.* 2020.2011.2024.395715.

Uddin, W., Menke, S., Melzheimer, J., Thalwitzer, S., Heinrich, S., Wachter, B. and Sommer, S. (2017). *Gut microbiomes of free-ranging and captive Namibian cheetahs: Diversity, putative functions and occurrence of potential pathogens.*

Vemuri, R., Shankar, E. M., Chieppa, M., Eri, R. and Kavanagh, K. (2020). Beyond Just Bacteria: Functional Biomes in the Gut Ecosystem Including Virome, Mycobiome, Archaeome and Helminths. *Microorganisms.* **8**(4): 483.

Vibin, J., Chamings, A., Collier, F., Klaassen, M., Nelson, T. M. and Alexandersen, S. (2018). Metagenomics detection and characterisation of viruses in faecal samples from Australian wild birds. *Scientific Reports.* **8**(1): 8686.

Wang, Y., Yang, S., Liu, D., Zhou, C., Li, W., Lin, Y., . . . Delwart, E. (2019). The fecal virome of red-crowned cranes. *Archives of Virology.* **164**(1): 3-16.

Wegener Parfrey, L., Walters, W. and Knight, R. (2011). Microbial Eukaryotes in the Human Microbiome: Ecology, Evolution, and Future Directions. *Frontiers in Microbiology.* **2**(153).

Westbrook, C. J., Karl, J. A., Wiseman, R. W., Mate, S., Koroleva, G., Garcia, K., . . . Palacios, G. (2015). No assembly required: Full-length MHC class I allele discovery by PacBio circular consensus sequencing. *Human Immunology.* **76**(12): 891-896.

Westerdahl, H. (2007). Passerine MHC: genetic variation and disease resistance in the wild. *Journal of Ornithology.* **148**(2): 469-477.

Westerdahl, H., Asghar, M., Hasselquist, D. and Bensch, S. (2011). Quantitative disease resistance: to better understand parasite-mediated selection on major histocompatibility complex. *Proceedings of the Royal Society B: Biological Sciences.*

Westerdahl, H., Stjernman, M., Råberg, L., Lannefors, M. and Nilsson, J.-Å. (2013). MHC-I Affects Infection Intensity but Not Infection Status with a Frequent Avian Malaria Parasite in Blue Tits. *PLoS ONE.* **8**(8): e72647.

Wigby, S., Suarez, S. S., Lazzaro, B. P., Pizzari, T. and Wolfner, M. F. (2019). Sperm success and immunity. *Current topics in developmental biology.* **135**: 287-313.

Wright, D. J., Shah, N. J. and Richardson, D. S. (2014). Translocation of the Seychelles warbler Acrocephalus sechellensis to establish a new population on Frégate Island, Seychelles. *Conservation Evidence.* **11**: 20-24.