Evolutionary forces shaping innate immune gene variation in a bottlenecked population of the Seychelles warbler

Danielle Louisa Gilroy **MBiol. Sci. (Hons)**

A thesis submitted for the degree of Doctor of Philosophy

School of Biological Sciences

University of East Anglia, UK

September 2015

Word Count: 63 075

Thesis abstract

In this thesis, I investigated different evolutionary forces in shaping genetic variation within a bottlenecked population of an island species, the Seychelles warbler (*Acrocephalus sechellensis*). I specifically explore pathogen-mediated selection within this system by using avian beta-defensins and toll-like receptor genes to examine functional variation. First, I characterise variation within both gene groups in this population and show that this species' demographic history has had an overriding effect on selection and random drift is the predominant evolutionary force. I characterise variation within these gene groups across several other *Acrocephalus* species, in addition to looking at a specific locus in a pre-bottlenecked population in order to directly compare genetic variation pre- and post-bottleneck. I use population genetic statistical methods to detect selection at several polymorphic genes and evaluate the robustness of these methods when applied to single-locus sequence data, which may be lacking in power and not meet the demographic assumptions that come with these tests. To overcome this, I designed forward-in-time simulations based on microsatellite markers used in pre- and post-bottleneck populations of the Seychelles warbler. I am able to delineate the evolutionary effects of selection from drift and show that some toll-like receptor genes are indeed under positive balancing selection in spite of the recent bottleneck. I further explore how this variation is maintained by conducting association analyses investigating innate immune gene variation and its relationship with individual survival and malarial susceptibility / resistance. Environmental factors are also considered. By investigating the consequences of functional variation in a bottlenecked species we are able to assess its long-term viability and adaptive potential, whilst elucidating the evolutionary importance of maintaining genetic variation in natural populations.

Acknowledgements

I thank my supervisor Prof David S Richardson for his endless and patient support in all aspects of the PhD and my secondary supervisor Dr Cock van Oosterhout for his never-ending enthusiasm and involvement in the project. You both make a fantastic supervisory team and I feel privileged to have been your PhD student. I thank Nature Seychelles for facilitating the long-term study of the species and providing permission to work on Cousin Island and thanks to the Department of Environment and Seychelles Bureau of Standards for giving the permission for fieldwork and sampling. This work was supported by a VH-C Dean Studentship from the School of Biological Sciences at the University of Anglia, Norwich, and an additional grant provided by Prof Jan Komdeur at the University of Groningen, the Netherlands.

This PhD would not have been possible without the fantastic group effort and support from the Seychelles warbler research group. I would like to thank Dr David Wright and Dr Sjouke Kingma for putting up with just me for company on a little island….I hope it wasn't too terrible! I would like to thank the numerous field assistants (past and present) for catching the birds and collecting data that was ultimately used in my analyses. Thanks go to Eleanor Fairfield and Dr Catalina Quevedo-Gonzalez for the lab support and for enduring my heavy-metal music and quirkiness. Thank you Dr Lewis Spurgin, Tom Finch, Catriona Morrison and Ben Ward for letting me pick your brain over statistics. I would like to thank Owen Howison and Dr Hannah Dugdale for being gurus on all things database-related. I thank Prof Jan Komdeur and Prof Terry Burke for their additional support as the 'grandfathers' of the project and for injecting years of knowledge and wisdom into my project at the bi-annual warbler meetings.

Personal thanks go to my fantastic friends here at UEA that have helped me through some exceptionally rough patches: Jake Gearty, Kris Sales, Jessie Gardener, Gen Labram and Jenny Donnelan. I will also never be thankful enough for the friendship and academic support I received from Dr Karl Phillips, Dr David Collins and Will Nash; three outstanding gentlemen. I have been blessed to be a part of a dynamic, outstanding department of great people so I must thank you all (too many to name)! Outside of UEA I have had some great support from Dr Darren White, Oliver Reville and my twin Bethan Kinder.

The thesis write-up is always a challenging time, particularly when presented with medical blips, so I could not have completed it without my amazing and loving family who are always behind all that I do. I would like to think Zac Hinchcliffe for proof-reading absolutely everything and for being such a loving and supportive partner. Finally, the most personal and special of thanks must go to two people. Thanks to my nan Jutta Jacob for watering me, feeding me and generally looking after me and hugging me when I needed it most ('I'll keep you safe'); and thanks to the most inspiring and wonderful woman I know, Anita Gilroy. My mum never lets me give up and is my absolute rock and best friend. This thesis is for mother and daughter. Now, where is my glass of wine?

Contents

Chapter contributions

At the time of submission, three data chapters presented in this thesis are submitted for publication. Below, I provide a citation for each data chapter, highlight authorship and specify my contributions.

Chapter 2: Gilroy DL, van Oosterhout C, Komdeur JK, Burke TA & Richardson DS (in press: Conservation Genetics).

- DLG role in preparing museum samples, fieldwork, lab work and drafting manuscript (75%)

Chapter 3: Gilroy, DL, van Oosterhout C, Komdeur JK & Richardson DS (in press: Journal of Immunogenetics).

- DLG role in fieldwork, lab work and drafting manuscript (75%)

Chapter 4: Gilroy, DL, Komdeur JK, Richardson DS & van Oosterhout, C (in press: Journal of Molecular Ecology).

- DLG co-designed simulations with CVO and drafting manuscript (65%)

Chapter 1: General Introduction

1.1 Molecular Ecology

Molecular ecology, in its broadest sense, is the application of molecular methods to ecological problems. It delves into the fields of population and evolutionary genetics, behavioural ecology and into conservation biology, which has developed into its own discipline since its emergence in the mid-1980s and it continues to grow (Beebee & Rowe 2004). One of its main areas of focus is the understanding of evolutionary change in wild populations, and how it is differentially determined by evolutionary forces. It is fundamental that we elucidate the underlying mechanisms that influence genetic variation, particularly in fragmented or bottlenecked populations of conservation interest, if we want to conserve the evolvability or adaptive potential of a species in an unpredictable future (Frankel 1974). Genetic variation is fundamental to the long-term viability of a population. Molecular methods are proving increasingly useful in the field of conservation biology, particularly given the ever increasing rate of loss of global biodiversity, because they provide powerful methods and measures that can inform conservation practice (Rodriguez de Cara *et al.*; Hedrick 2001; Sommer 2005).

Genetic characteristics vary considerably within and among populations. The field of population genetics investigates a number of components of this variation, including genetic diversity, genetic differentiation and effective population size (*Ne*). Large populations tend to support higher levels of genetic diversity compared to smaller populations, because they are less prone to the stochastic loss of genetic variation due to genetic drift (Wright 1930; Frankham 1996). With more genetic variation, natural selection has a richer substrate to select from, which means that there is greater potential for the population to adaptively evolve (Fisher 1930). This is predicted by the early population genetic work of Fisher, and it is known as Fisher's fundamental theorem. However, studies have shown that this is not always observed, and that small inbred populations can show a high adaptive potential (Franklin and Frankham 1998; Frankham *et al.* 1999; Lanfear *et al.* 2013). Therefore, we need to understand to what extent genetic drift and selection shape genetic variation within and among populations.

A number of methods have been developed in this field to characterise variation in populations, which often require combined data on sequence divergence between species

and polymorphism within species (Morin *et al. 2*004). We can then use direct evidence (relating to the sites which are targets of selection) and indirect evidence (from nearby regions) that selection is shaping variation within a population. Neutral markers are useful to estimate genetic diversity within a population, which can infer the evolutionary potential of that population / species. However, these estimates can differ from those gained when using functional markers. Measures of adaptive variation should be combined with those of neutral variation in order to truly understand the evolutionary potential of that gene pool.

Genetic differences among individuals are generated by a number of evolutionary forces, in particular mutation and recombination-like processes (which include gene duplication and gene conversion). The Neutral theory (Kimura 1968) states that drift and mutation are the main forces that explain genetic variation (Lande 1976). Mutations are direct changes in the nucleotide sequence of DNA and there are many different types. Point mutations involve one nucleotide replacing another, while others can involve the insertion or deletion of a number of nucleotides. Some mutations can be the result of DNA replication slippage or by the movement of transposable elements within a sequence. These all act to increase individual variation and population differentiation. Mutation rates are highly variable across genes, taxa and developmental stages and subject to different types and strengths of selection (Kimura & Ohta 1969). Very few mutations are actually beneficial (ca. 1-2%) and the rest are neutral i.e. synonymous substitutions where there is no change to the translation of the protein (Kumar & Subramanian 2002)**.** Recombination, on the other hand, results in a restructuring of part of the genome, for example by the exchange of segments of homologous chromosomes during meiosis (Watterson 1975). It is still a type of mutation that can generate novel genotypes and it is an important process to consider when characterising variation (for examples, see Padidam *et al.* 1999; Schaschl *et al. 2*006; Cizkova *et al.* 2011).

1.1.1 Island models

Island-endemic species have long been important to evolutionary research since Darwin's HMS Beagle voyage around the Galapagos Islands eventually led to the publication of selection theory in the *Origin of Species* (1859). This is because oceanic islands make ideal

systems in which to investigate functional variation, due to a number of key features (Whittaker & Fernandez-palacios 1998). Firstly, islands create naturally fragmented study systems with discrete boundaries and so the island is readily-quantifiable. This is advantageous over continental systems in that they are more tractable (Emerson 2002). Secondly, islands are, to different degrees, isolated and consequently migration and gene flow between the populations that exist upon them is reduced. This can have consequences on the effective population size, thus leading to elevating levels of inbreeding and depleting genetic variation (Franklin & Frankham 1998) and because island populations tend to be small, genetic drift plays a particularly important role. Both demographic stochasticity (random variation among individuals in their survival and reproduction) and environmental stochasticity (containing a diversity of habitats despite their small geographical size, promoting local adaptation), will have much larger roles in shaping the genetic variation in island populations than they do on in large mainland populations (Emerson 2002). The attributes of islands combine to provide unusual research opportunities and the implications of these can stretch far beyond islands (Warren *et al.* 2015).

1.2 Genetic variation

It has long been debated about the role and importance of genetic variation in the drive, maintenance and long-term viability of populations (Lande 1988; Spielman *et al. 2*004; Frankham 2005; Pertoldi *et al. 2*007). While this debate has indeed been largely resolved and the importance of genetic variation widely established (Saccheri *et al.* 1998; Westemeier 1998; Reed & Frankham 2003; O'Grady *et al. 2*006), it is still proving difficult to gain a holistic understanding of the interaction between genetic, phenotypic, demographic and ecological factors in natural populations. The combination of these factors will lead to different genetic characteristics within and among populations. We can deduce the relative roles of the different evolutionary forces in natural populations using: i) population genetics, looking at the changes in allele, haplotype and genotype frequencies; ii) quantitative genetics, quantifying changes in fitness, behaviour or phenotype and iii) phylogenetics / macro-evolution that involves looking for footprints in the genome. While all evolutionary forces influence genetic variation, only natural selection and sexual selection act in a non-random manner and is responsible for a species or population being able to adapt to

environmental factors (Darwin 1859). However, other evolutionary forces can override this and can be hard to disentangle from one another and consequently, promote the loss of (Sutton *et al.* 2011).

Genetic drift is the predominant force responsible for the loss of genetic variation in small populations (Lacy 1987). This phenomenon is the random change in allele frequencies in a finite population with each generation, due to the random sampling of parental alleles under the laws of Mendelian inheritance (Wright 1930). In smaller populations, these random changes are greater and so smaller populations endure more genetic drift, which ultimately leads to the loss of alleles from the gene pool. Any process that reduces the effective size of a population, such as inbreeding, population fluctuations and bottlenecks, will lead to heightened levels of drift (Masatoshi *et al.* 1975; Nei & Tajima 1981). In turn, the loss of genetic variation will result in increased homozygosity, resulting in the increased expression of deleterious recessive alleles (inbreeding depression) (Crow 1980). Although the vast majority of evolutionary changes at the molecular level are caused by drift, most mutants are selectively-neutral and so do not affect fitness (Kimura & Ohta 1969; Kimura 1986). However, there are exceptions when drift can act so strongly. For example in small populations, positive mutations can be eliminated or mildly-negative mutations can reach fixation (for review, see Charlesworth 2009).

Gene flow among populations opposes genetic drift by increasing genetic variation. It is different to mutation and genetic drift in that it is not a random process because it can be phenotype or sex-dependant (Takahata & Palumbi 1985; Chesser 1991). The successful reproduction of individuals between two populations, allows the mixing of gene pools and the new individuals gain novel genotypes. By gaining variation, local adaptation is promoted and inbreeding is reduced. However, gene flow reduces coalescence time at the meta-population level i.e. it quickens the time it takes for two spatially-distinct populations to interact and merge together (Slatkin 1987). Therefore, on a global scale, gene flow reduces genetic variation and can act as a constraining force (Mayr 1996). However, it does importantly introduce novel genetic variation available to selection but unlike mutation, it is a non-random process that can be phenotype or sex dependent (Slatkin 1987; Chesser 1991).

Natural selection is the main driver of adaptive evolutionary change (Darwin 1859). It acts on heritable genetic variation that confers a fitness advantage, thus allowing an organism adapt to their environment (Fisher 1930). In a constant environment, natural selection will keep a population stable, but if a new variation which is advantageous to the individual it will increase in frequency within that population through successful transfer to offspring. In contrast, less successful genetic variants will decrease in frequency as natural selection acts to remove them from the gene pool (Lande 1976b; Mousseau & Roff 1987). Directional selection occurs when natural selection favours one extreme of continuous variation, resulting in the opposing extreme becoming rare or even lost from the gene pool (Vousif & Skibinski 1982), and stabilising selection is when natural selection favours the intermediate states of continuous variation and so the extremes become lost (Barnes 1968; Gibson & Bradley 1974). An alternative mode of natural selection is disruptive or diversifying selection, which is when both extremes of continuous variation are favoured within a population. The intermediates are thus reduced or lost, and in extreme cases, this can lead to two new species (Wolstenholme & Thoday 1959; Thoday 1972). Balancing selection refers to a variety of selection regimes that act to maintain genetic variation within populations that are advantageous to the individual in promoting fitness (Hedrick 2006; Mitchell-Olds *et al. 2*007).

By these different modes of natural selection, populations are able to adapt and persist to a heterogeneous environment, and there are many examples of wild populations that have rapidly responded to novel (often anthropogenic) challenges through evolutionary change. Examples include: reproductive methods in plants (Morran *et al. 2*009), herbivorous insects rapidly responding to invasive plant species (Siemann *et al. 2*006) in addition to insects adapting their host associations in response to anthropogenic change (Singer *et al.* 1993). Pink salmon populations have altered their life cycles (Waples *et al. 2*009), pocket mice can rapidly adapt their coat colours for camouflage (Nachman *et al. 2*003) and passerine birds rapidly evolve their singing in urban areas (Patricelli & Blickley 2006). It is the relative roles of adaptation versus non-adaptive forces in shaping the diversity of life within and between species that has become a key question which lies at the heart of biology.

1.3 Pathogens as evolutionary drivers

Being able to elucidate what mechanisms are responsible for maintaining genetic variation in natural populations has received much attention in evolutionary biology because, in the words of Dobzhansky (1951 p.109), the '*absolute equality of adaptive values of two biological forms is….highly unlikely*.' Essentially, one form will replace the other eventually and variation is lost. Balancing selection can be mediated by strong selective agents and a particularly strong driver of demographic and evolutionary change in natural populations are pathogens (Jeffery & Bangham 2000; Ford 2002; Bernatchez & Landry 2003). Pathogens exploit other organisms for their own growth and survival, and thus have detrimental effects on the intrinsic growth rates of their host at both an individual and population level (Anderson & May 1978). They encompass a vast variety of groups of organisms including protozoa, viruses, bacteria, fungi, flatworms, nematodes and arthropods (for review, see Noble *et al.* 1989). It is their intimate relationship with the host that is responsible for a continuous and cyclic co-evolutionary arms race. This concept is outlined in the Red Queen Hypothesis, which states that organisms need to constantly adapt evolve and proliferate against opposing organisms (Peters & Lively 1999). Therefore, pathogens can effectively mediate balancing selection through influencing their hosts ability to adapt and survive (Sorci & Moller 1997; Merino *et al. 2*000; Sol *et al. 2*003; Moller & Saino 2004; Worley *et al. 2*010; la Puente *et al. 2*010).

Mortality caused by pathogens has been shown to drive the demographic structure of populations (Hudson 1986; Redpath *et al. 2*006; Deter *et al. 2*007; Pedersen & Greives 2008; Llaurens *et al. 2*012). They also affect other factors like reproductive success (Brouwer *et al. 2*010; Knowles *et al. 2*011; Eizaguirre *et al. 2*012; Radwan *et al. 2*012), secondary sexual features and behavioural traits (for review, see Piertney & Oliver 2006). This makes pathogen-mediated selection (PMS) ideal to investigate a number of different balancing-selection mechanisms in order to understand how polymorphisms are maintained. Previously, studies have treated these mechanisms as if they were mutually exclusive when in fact, they can act in concert. However, it is difficult to disentangle the effects of one mechanism from another (Spurgin & Richardson 2010).

While a number of mechanisms have been put forward to explain pathogen-mediated balancing selection (for reviews, see Potts & Slev 1995; Hedrick 2002; Garcia de Leaniz *et al. 2*007), there are three main mechanisms proposed: heterozygote advantage, rare allele advantage and fluctuating selection (Doherty & Zinkernagel 1975; Hill *et al.* 1991; Slade & McCallum 1992, respectively). Heterozygote advantage is arguably the simplest model of balancing selection, and often referred to as 'overdominance' since its initial proposal by plant geneticists (East 1908; Shull 1908) to explain observations of hybridisation and inbreeding depression (Darwin 1876; Crow 1948). Dobzhansky outlined overdominance as a key explanation for balanced polymorphism in populations based on an 'adaptive superiority' of heterozygotes (Dobzhansky 1951 p.132). The heterozygote advantage was further developed to apply to the extra-ordinarily high levels of polymorphism at the Major Histocompatibility Complex (MHC) in that selection would favour heterozygous individuals because they could recognise more different antigens and have better immune defence compared to homozygotes (Doherty & Zinkernagel, 1975). There are two principal forms of heterozygote advantage: over-dominance, and simple dominance. The MHC has been used to demonstrate over-dominance in that there is a superior fitness of heterozygous genotypes over homozygous genotypes at a single MHC locus (Shull 1908; Doherty & Zinkernagel 1975). Simple dominance involves the cancelling of deleterious or inferior recessive alleles inherited from a parent, by advantageous or superior dominant alleles contributed by another parent at different loci (Bruce 1910; Jones 1917). The two forms are under much examination and as it stands, there is still no consensus on the genetic basis underlying heterozygote advantage.

Frequency-dependant selection, also called rare allele advantage (Edwards & Hedrick 1998) was first proposed when arguing the battle of the sexes, in that the total reproductive success of each sex is equal (Fisher 1930). For PMS, rare allele advantage occurs when common parasites evolve resistance to common host genotypes and thus the host with rare alleles have a selective advantage (Slade & McCallum 1992). This predicts that parasite and host genotypes would constantly evolve in cycles in relation to each other thus retaining polymorphisms (Jeffery & Bangham 2000; Hedrick *et al. 2*001). Fluctuating selection suggests that spatiotemporal variation in the pathogen fauna challenging a host, and thus the associated selection pressure contributes to increased immune-gene diversity (Hill *et al.*

1991). Geographical and temporal variation in pathogen type and prevalence within populations can cause differences in selection in space and time. The key points to this particular model is that (i) selection is directional rather than cyclical (like the rare allele advantage model) and that (ii) pathogen fluctuations are determined by external biotic and/ or abiotic factors, chance dispersal and extinction events (for review, see Botero & Rubenstein 2012)**.** Theoretically, it has been shown that fluctuating selection could maintain diversity at the MHC, even in the absence of heterozygote and rare-allele advantage (Hedrick 2002). However, there is still little empirical work identifying balancing selection via fluctuating selection pressures.

1.3.1 Avian-Malaria models

A good host-pathogen system is needed in order to fully examine and understand how the mechanisms of balancing selection operate. Malaria is widely-studied because its parasites are responsible for some of the highest-impact diseases in humans, livestock and wildlife (Garnham 1980). The genus *Plasmodium* alone infects over a third of the world's human population with 90% of cases originating in Africa (Snow *et al. 2*005) with an estimated 584,000 deaths in 2013 alone (World Health Organisation 2015). They have a relatively complex life cycle involving indirect transmission by blood sucking insects (of the Dipteran order), for example mosquitoes, in which stages of development occur in both tissues and circulating red blood cells (Atkinson & van Riper 1991) (Fig 1). Despite being studied for over a century, malaria parasites have resisted all efforts of eradication. Studies on malaria parasite resistance have elucidated many sophisticated examples of evolution such as adaptive manipulation by the parasite of host behaviour and host sex (for excellent examples, see Lafferty & Kimo Morris 1996; Hurst *et al.* 1999, respectively).

Avian malaria is an excellent host-pathogen study system and the presence of malaria blood parasites has been used specifically in birds to look at immune-competence (Marzal *et al. 2*005; Lee *et al. 2*006; Mendes *et al. 2*006; Hale & Briskie 2009). Field studies of avian malaria parasite-host systems commonly use two traits in hypotheses testing: (i) prevalence at the population level, and (ii) parasitaemia, the density of parasites within the infected host (for review, see Knowles *et al. 2*011). We can test specific predictions about the infection intensity and fitness parameters (Friedl & Groscurth 2012). For example, there

are some passerine studies that show significant fitness consequences of malaria infections in wild birds including survival (e.g. Atkinson *et al.* 1995; Bensch *et al. 2*007; Lachish *et al. 2*011; Ferrer *et al. 2*014; Marzal *et al. 2*015) and more recently consequences of infection relating to the degradation of telomeres, the protective caps at the end of chromosomes (Asghar *et al.* 2015; Watson *et al.* 2015). Additionally, there has been a demonstrated trade-off between reproduction and defence against *Plasmodium* infections (Bonneaud *et al.* 2006; Podmokła *et al.* 2014, 2015; Staley & Bonneaud 2015). We must compare and contrast different immune responses of birds to novel pathogens and see how they associate with clinical symptoms, pathogen load and mortality or on the contrary, how they are linked to increased pathology and reduced host survival.

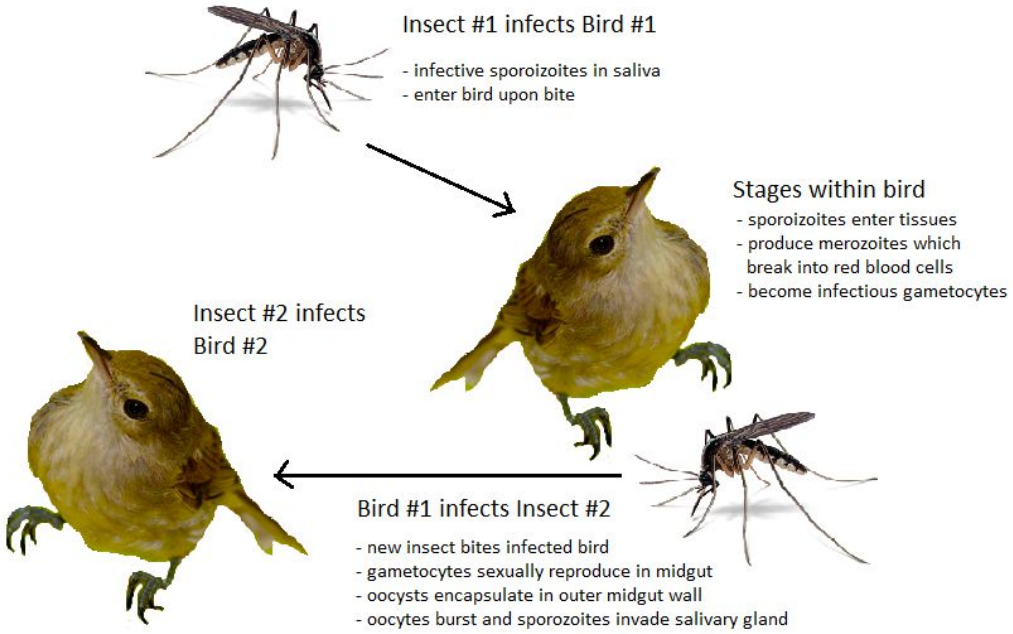


Figure 1. Indirect transmission of malarial parasites via a Dipteran vector (mosquito) to its definitive host (Seychelles warbler).

1.4 Candidate gene approach

When assessing variation in finite natural populations, it is fundamental that the effects of drift and selection (including PMS) are combined. Drift can reduce the effectivity of natural selection because its random changes can override the effects that selection has on fitness (Lande & Terms 1976; Lacy 1987). In order to best assess the effects of both evolutionary forces, the genetic composition of natural populations needs to be characterised with a

focus on adaptive variation. This is a better correlate of mean individual fitness when wanting to understand the adaptive potential of a population and so molecular measures have shifted towards the focus on specific genes that are likely candidates to be under strong selective pressures. It is necessary to characterise variation at these 'critical loci' to measure variability at ecologically important traits, particularly in endangered species. However, there is no best method for detecting genetic variation in natural populations and it is often advised to combine these with neutral markers.

The 'bottom-up' candidate gene approach (CGA) in population genetics is one way that molecular ecologists can examine functional variation within and among populations. This approach involves identifying a gene(s) based on existing knowledge of its function from previous work and / or other species, ideally from model organisms and proceeding to investigate how they function in terms of the phenotype(s) expressed in your own model population (Fitzpatrick *et al. 2*005; Amos *et al. 2*011). However, this approach does present some challenges. For example, the relationship between the genetic variant and behaviour is not necessarily deterministic and there can be considerable phenotypic plasticity (Woltereck 1928). The success of CGA depends on a number of things including the individuals' chosen, whether a gene has multiple effects (pleiotropy) or whether gene-gene interactions can affect the overall phenotype (epistasis) (for review, see Piertney & Webster 2010). Nonetheless, this approach has many advantages in helping us to understand the function of variation at a specific locus. It allows the quantification of genetic diversity among populations in order to identify depauperate populations, and is in fact, more applicable to natural populations where pedigrees may be unavailable than the 'top-down' approach (Fitzpatrick *et al. 2*005).

Top-down approaches, such as examining quantitative trait loci, genome wide association studies (GWAS) and linkage disequilibrium, reverse the order of investigation by starting with a phenotype of interest and using genetic analysis to identify candidate genes. Both approaches mean that CGA can identify both the strength and mode of selection acting on specific genes being targeted. It also can shed light on direct mechanistic links between allelic richness, allelic variation at specific loci at single genes, and variation in individual fitness (Amos *et al. 2*011). Therefore, it is an ideal approach to take when investigating

functional variation in natural populations and considering PMS as the major explanatory force in question.

Immune-genes make ideal candidates for this approach, particularly since the association between health against infectious diseases and evolutionary fitness are well-documented across a range of taxa (May & Anderson 1983; Ohlberger *et al. 2*011; McTaggart *et al. 2*012). This pre-requisite helps avoid the risk that comes with CGA that the candidate gene(s) in question may not be functional or indeed important in the study population. The major candidate gene group tested for evidence of pathogen-driven allelic variation is the most polymorphic vertebrate gene cluster, the Major Histocompatibility Complex (MHC). It has a pivotal role in recognising self from non-self molecules by binding to peptides and presenting them to T-cells. If the T-cells fail to recognise the peptide then the MHC triggers an appropriate immune response (Snell 1978; Klein 1986). There is exceptional evidence of the relationship between MHC variations and pathogen resistance (for some examples, see Aguilar *et al. 2*004; Bonneaud *et al. 2*006; Schwensow *et al. 2*007; Westerdahl *et al. 2*010; Eimes *et al. 2*011). Studies have shown that balancing selection can maintain variation at MHC loci, particularly by pathogen-mediated selection (for review, see Bernatchez & Landry 2003). This variation, in turn, affects many other key biological traits such as mate choice (Landry *et al. 2*001; Reusch *et al. 2*001; Richardson *et al. 2*005) kin recognition (Manning *et al.* 1992; Olsén *et al.* 1998; Zelano & Edwards 2002) and autoimmune disease (Akilesh *et al. 2*004; Fernando *et al. 2*008).

Whilst the MHC has major roles in adaptive immunity and a plethora of studies on this exist, its research has been weighted with problems given the MHC's complex evolutionary history involving multiple duplications resulting in difficulties phasing MHC alleles (for review, see Garrigan & Hedrick 2003). There is also a large bias in the focus on MHC II B molecules, as highlighted in a meta-analysis by Sutton *et al.* (2011) where they found that 94% of bottleneck studies were based on MHC II polymorphisms. As it stands, there are many non-MHC immune genes which are just as important, if not more, to immune defence (Acevedo-Whitehouse & Cunningham 2006)**.** It is thought that the focus should now shift towards innate immunity since it is our first line of defence (Kaiser 2007, 2010). Furthermore, it would allow us to better understand the interplay between both the innate and adaptive arms of the immune system.

1.4.1 Defensins

Defensin genes encode for antimicrobial peptides (AMPs (Table 1)), which have long been established as an important component of innate immune defence (Ganz *et al.* 1985; Selsted *et al.* 1985)**.** These peptides kill a broad range of bacterial strains, as first shown with *Escherichia coli* in mammalian hosts (Lehrer *et al.* 1989)**.** The AMPs can do this directly by physically attacking and disrupting the pathogen's membrane, rendering the organisms inviable. This is done via their cationic and amphipathic properties (Hancock & Sahl 2006) and was well-shown in a population of great tits *Parus major*, where much of the allelic variation observed had an effect on amino acid composition and altered the net charge and hydrophilicity of the peptide produced (Hellgren 2015). As a result, this changed the properties associated with efficiency of being able to bind to and rupture pathogens.

Table 1. Comparison of vertebrate defensin sub-families, modified from Sugiarto & Yu (2004).

Name	Structure	Size (kDa)	Residues	Source
α-defensins	Beta-sheet dimer	3.5-4	29-35	Human, rabbit, rat, guinea pig, mouse
β-defensins	Beta-sheet dimer	4-6	38-42	Human, turkey, ostrich, chicken, king, penguin, pig, bovine
γ-defensins	Cyclic	2	18	Rhesus monkey

Table 2. Primers and corresponding annealing temperatures for amplification of nine different avian β-defensins from passerine species. Modified from Hellgren & Sheldon (2011).

Gene	Primer	Anneal temp, (°C)	Fragment length (bp)	Number of polymorphic sites
AvBD2	F2mat, R2mat	51	147	16
AvBD4	F1, R1	55	120	21
AvBD7	F2mat, R2	57	170-176	51
AvBD8	F2, R1	50	153	31
AvBD9	F1, R1	51	142	18
AvBD10	F1, R1	52	133	32
AvBD 11	F1mat, R1	60	157-209	53
AvBD12	F2, R1	53	162-190	38
AvBD13	F1mat, R1mat	60	137-140	21

Alternatively, they can carry out innate defence indirectly through cytokine production (Hancock & Scott 2000) and liaising with other immune defence components. For example, a recent study on the pigeon virus *Paramyxovirus type 1 3* (PPMV-1), showed a correlation between the expression of different avian β-defensins and different toll-like receptors (Li *et al. 2*015).

AMPs are characterised by having six cysteine residues defensin motifs and their evolutionary history is only just coming to light as more genomic data for different avian lineages becomes available for avian defensin genes. Chen *et al.* (2015) have just released genomic data for the golden pheasant *Chrysolophus pictus* (a Galliformes species) and the hwamei *Garrulax canorus* (a Passeriformes species) and found that by combining them with the model species of chicken and zebra finch, an evolutionary history of duplications and deletions have been found to give rise to the clearly different genomic structures (Chen *et al. 2*015). They further found that transposable elements were agents of their evolution, causing direct and indirect copy number variations in β-defensins via these duplication events. Different taxonomic groups have different classes and numbers of defensins in their immune repertoire (Selsted & Ouellette 2005). For example, birds have only β-defensins, of which 14 different loci have been identified in the domestic chicken *Gallus gallus domesticus* (Lynn *et al. 2*004; Xiao *et al. 2*004), whereas mammals have both α and β-defensins (Yang *et al. 2*002). The β-defensin number in a species has been shown to be highly relevant to the ever-changing microbial challenges from the environment in which the host inhabits (Tu *et al. 2*015). There is little information on how much natural genetic variation defensin genes exhibit in wild vertebrate populations, but the influence of allelic variation at these genes on infection outcome has been shown in a range of vertebrate hosts (Meredith *et al. 2*008; Mukherjee *et al. 2*009; Hellgren *et al. 2*010; Chow *et al. 2*012).

Avian β-defensins (AvBDs) are ideal candidates for functional variation study, since *in vitro* tests have already showed that small nucleotide variations in sequence encoding the AMPs can change the peptide's physical properties. Consequently, this alters efficiency (or effectiveness) in preventing microbial growth (Meade *et al. 2*008)**.** Another *in vitro test* looked at the differences in the anti-microbial properties of the synthesised products of two alleles of *avian* β-*defensin 7* (*AvBD7*) (Hellgren *et al. 2*010). Both alleles occur at high frequency in natural populations of great tits *Parus major* and were found to strongly inhibit

the growth of *Escherechia coli* among other closely-related gram-negative bacteria infections. This was the first demonstration of functional allelic variation in natural defensin genes having different effects on pathogens.

Antimicrobial defensins are strong candidates for examining pathogen-mediated balancing selection, particularly in passerines, since a locus-specific protocol has now been set-up to amplify and investigate inter- and intra-specific genetic variation within AvBD loci in passerines (Hellgren & Sheldon 2011; Table 2)**.** New sequence blocks are selected and amplified by aligning genomic sequences from the domestic chicken and zebra finch *Taeniopygia guttata*, and polymorphisms at critical loci (for AvBD genes) can be confirmed by 454-transcriptome sequencing (Fig 2). By directly comparing AvBD genes among the chicken and zebra finch genomes, it was found that whilst the galliformes-passeriformes split ~10 mya gave rise to 12 novel AvBD genes, there are still 10 genes which are highly conserved and orthologous out of the 22 investigated (Hellgren & Ekblom 2010)**.** Furthermore, we can consider the findings from an analysis of immune genes in the zebra finch genome (n = 144) where several candidate gene groups including AvBDs had elevated ratios of non-synonymous substitutions to synonymous substitutions (Ekblom *et al. 2*010)**.** This is indicative of positive selection acting at these genes, which in combination with being conserved across avian lineages, makes them ideal candidates for this research.

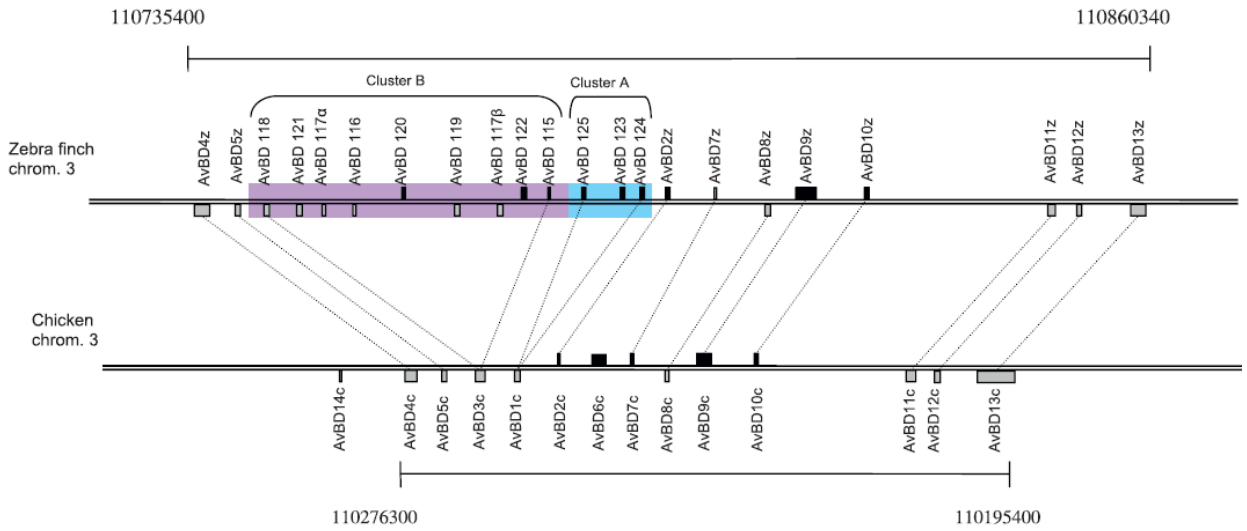


Figure 2. Location of beta-defensin genes on chromosome 3 of genomic models, chicken *Gallus gallus domesticus* and zebra finch *Taeniopygia guttata.* Directly taken from Hellgren & Ekblom, 2010.

1.4.2 Toll-like receptors (TLRs)

Toll-like receptors (TLRs) are membrane-bound sensors that play a key role in recognising distinctive molecular features of invading microbes, acting as part of the innate immune system (for review, see Jin & Lee 2008). They bind to pathogen-associated molecular patterns (PAMPs), thus triggering an intracellular signal cascade to activate an appropriate immune response (Belvin & Anderson 1996; Takeda & Akira 2005). They have an extra-cellular domain that is characterised by varying numbers of leucine-rich repeats (LRRs) which form a 'horse-shoe' structure to interact with nucleic acids and proteinaceous ligands (for review, see Skevaki *et al.* 2015) (Fig 3). Variants in the toll gene were first identified in *Drosophila melanogaster* and since then, 13 mammalian toll genes have been identified (Anderson *et al.* 1985). TLRs are divided into six families based on the types of PAMPs they bind to (Roach *et al. 2*005). These include TLRs which bind to bacterial lipoproteins, lipopolysaccharides or DNA motifs (Takeuchi *et al.* 2002; Bihl *et al. 2*003; Keestra *et al. 2*010). TLRs link the innate immune system with the adaptive immune system in vertebrates, in that they identify the infectious agent as a first line of defence. Furthermore, by recognising these specific PAMPs, they effectively inform other components of the immune repertoire (Schnare *et al. 2*001; Roach *et al. 2*005).

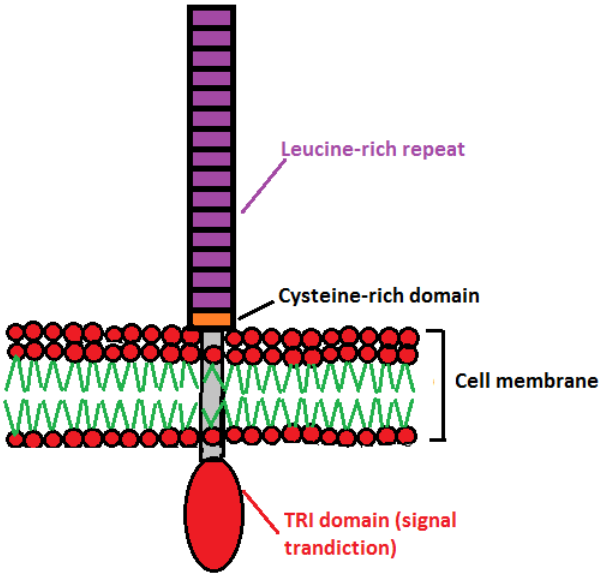


Figure 3. Toll-like receptor molecule structure.

TLRs have proved to be good candidates already for investigating functional variation. A recent study looked at ten TLR genes in the Tasmanian devil *Sarcophilus harrisii*, a mammal of conservation interest where previous studies have revealed low genetic diversity at microsatellite and MHC loci, found diversity was also low at TLR loci (Cui *et al.* 2015). By assessing their 'insurance' population that safeguards the species from extinction, they managed to show that they had captured all known TLR alleles in that species. The same ten TLR genes were screened in seven phylogenetically-diverse avian species and several alleles that appeared to confer low individual fitness, decreased in frequency across the different avian species examined (Alcaide & Edwards 2011). Slow rates of non-synonymous substitution were also observed, which would indeed help to preserve their immunological function (Nei & Gojobori 1986; Ohta & Ina 1995). The same study then focuses on TLR polymorphism in wild populations of lesser kestrel *Falco naumanni* and house finch *Carpodacus mexicanus.* Results showed low to moderate levels of polymorphism and an excess of synonymous substitutions, indicative of negative (purifying) selection. This is surprising, given their similar structure and function to the MHC, which is the model candidate for positive (balancing) selection studies. A recent study supported this by investigating TLRs in the grey partridge *Perdix perdix* and despite finding non-synonymous polymorphisms, they found the variation to have minor functional impact and assume that either negative selection or a bottleneck may have reduced TLR population viability in this species (Vinkler *et al.* 2015).

Direct associations between polymorphisms within TLR loci and pathogen resistance and susceptibility have been established (see: Creagh & O'Neill 2006; Vinkler *et al. 2*009; Franklin *et al. 2*011), however, it is still unclear how TLR polymorphisms will compare among other immune genes in avian species. Regardless, they are excellent candidates for investigating the role of PMS in maintaining variation within this group and this is shown by the number of studies that have already been carried out in fish (Palti 2011), mammals (Nakajima *et al. 2*008; Areal *et al. 2*011; Tschirren *et al. 2*013) and in birds (Downing *et al. 2*010; Grueber *et al. 2*013, 2014). Primers are readily available (Table 3) and by targeting conserved coding regions, specific roles in pathogen recognition and antimicrobial defence have been identified (Table 4). It has already been suggested that patterns of genetic variation at TLR loci would be particularly interesting to study in bottlenecked, fragmented

and decimated populations (Acevedo-Whitehouse & Cunningham 2006), in order to understand the evolutionary dynamics of TLR genes in organisms able to colonise new habitats. It can then be established that the genetic differentiation is indeed adaptive and related to survival and / or disease resistance. Furthermore, a recent study that looked at the relationship between microsatellite and TLR heterozygosity in a bottlenecked avian population, showed that the lack of a relationship is evidence that the predictive power of microsatellites in evaluating functional diversity is poor (Grueber & Carolyn 2015). This highlights the importance of adding data from putatively functional genomic regions, such as TLRs, in the study of genetic variation of endangered or threatened species.

Table 3. Polymorphism statistics at ten TLR genes in house finches *Carpodacus mexicanus*. Modified from Alcaide & Edwards (2011).

Gene	Size (bp)	Tajimas D	SNPs (dS: dN)	GenBank Accession Numbers
TLR1LA	1,161	-0.93	44 (27:17)	GU904709-70
TLR1LB	951	-0.37	25 (19:6)	GU904771-90
TLR2A	560	-1.27	13 (8:5)	GU904791-98
TLR2B	513	0.11	11 (7:4)	GU904799-803
TLR3	952	-0.51	11 (5:6)	GU904804-812
TLR4	789	-0.95	16 (8:8)	GU904813-826
TLR5	951	n/a	2 (n/a)	GU904827
TLR7	982	-0.35	27 (15:12)	GU904828-42
TLR15	1,300	-0.17	37 (19:18)	GU904843-58
TLR21	831	n/a	2 (1:1)	GU904859-60

Table 4. Avian toll-like receptors. Modified from Brownlie & Allan, 2011.

TLR	2nd name	Agonist	Pathogen
TLR1LA	*TLR1.1*		
TLR1LB	*TLR1.2*	Lipoprotein	Mycoplasma
TLR2A	*TLR2.1*	Peptidoglycan	G+ bacteria
TLR2B	*TLR2.2*		
TLR3		dsRNA	Viruses
TLR4		LPS	G- bacteria
TLR5		Flagellin	G- bacteria
TLR7		Imiquimod, ssRNA	Viruses
TLR15		Unknown	
TLR21		CpG motifs, chromosomal DNA	Bacteria and viruses

1.5 Conservation genetics

Using the principles and tools of molecular ecology, we can seek to understand the causes and consequences of genetic variation within populations and, consequently, get an understanding of the genetic vulnerabilities of any wild populations and species (Grueber & Carolyn 2015). We need to understand what evolutionary forces have shaped this variation and it is particularly important to characterise variation in bottlenecked populations to assess whether they have become genetically depauperate as a consequence. Conservation genetics itself did not establish until ca 1980 when three consecutive books were published outlining the key principles as a branching off from molecular ecology (Soulé & Wilcox 1980; Frankel & Soulé 1981; Schonewald-Cox *et al.* 1983). It has become particularly important, when combining the study of selection and drift in natural populations as these forces interact and it makes natural selection more difficult to assess and predict. Therefore, a number of factors must be taken into account and no doubt the field of conservation biology will undoubtedly expand in the future (Allendorf *et al. 2*010 Fig 4; Avise 2010; Frankham 2010).

Pathogens are being increasingly cited as major threats in conservation (Tompkins & Poulin 2006). Naïve hosts are often susceptible to the introduction of exotic reservoir species that cause disease and the infection spreads rapidly throughout the host population (for examples, see Cunningham *et al. 2*003; Anderson *et al. 2*004). A well-known case is the infection by the malaria species, *Plasmodium relictum*, carried by exotic avian hosts that were introduced to endemic Hawaiian land birds, notably Hawaiian honey-creeper species. It had catastrophic consequences for the island population following the establishment of its mosquito vector, *Culex quinquefasciatus* (Atkinson & van Riper III 1991; Atkinson *et al.* 1995). Whilst to date there is no empirical evidence that any global host extinction has been due to disease as a direct causation factor (De Castro & Bolker 2005), the loss of individuals through parasite infections can accelerate genetic drift and result in a so-called extinction vortex (Shaffer 1981; Gilpin & Soulé 1986; Fagan & Holmes 2006).

When developing conservation plans and management for species and populations, the maintenance of pathogens is not often considered, despite their roles in maintaining overall biodiversity (Hall 1999). Pathogens can indeed have severe consequences in naïve

populations, but endemic pathogens play a greater role in maintaining genetic diversity than previously anticipated. If PMS is a sufficient form of balancing selection that can maintain diversity at immune genes, then a paucity of pathogens could have important consequences on the long-term genetic viability of a host population. This is exacerbated in translocated or populations which undergo a series of bottleneck events and already suffer from reduced genetic variability (Frankham 1995). If further variation is lost at their immune loci, they will be more vulnerable to infectious diseases in the long-term future (O'Brian & Evermann, 1988). Therefore, it would be of value to maintain pathogen diversity in such populations and to always consider the overall biodiversity.

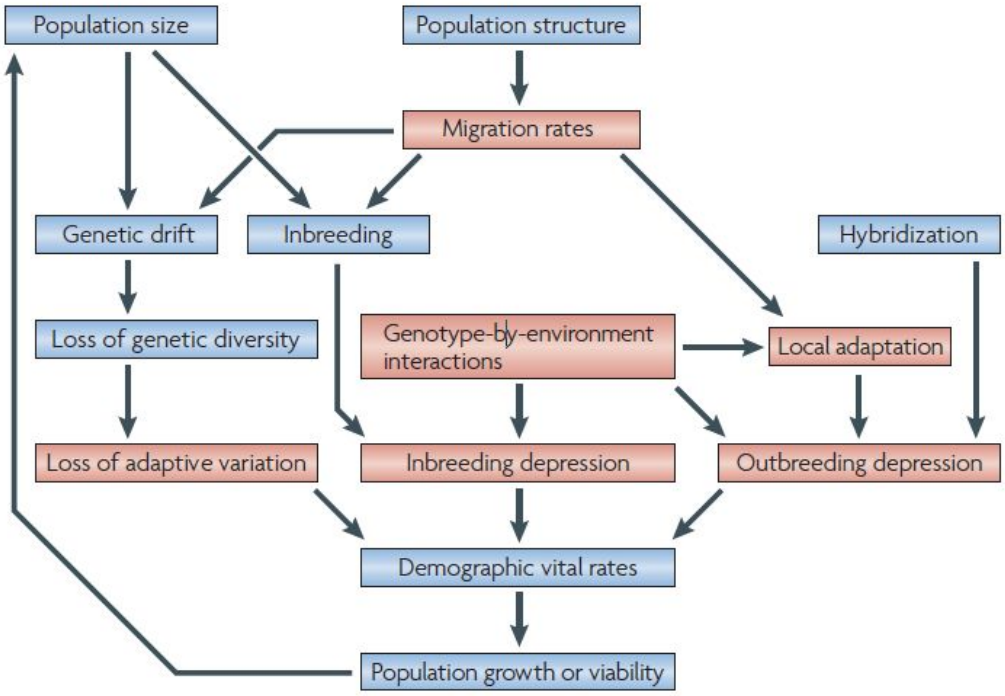


Figure 4. Schematic diagram of interacting factors in conservation of natural populations taken directly from Allendorf *et al.* (2010). Traditional conservation genetics, using neutral markers, provides direct estimates of some interacting factors (blue). Conservation genomics can address a wider range of factors (red).

The world is facing a biodiversity crisis at the hands of humans, where we are predicted to lose one quarter of all vertebrate species within the next century (Baillie *et al. 2*010) as part of a current global mass extinction (Diamond 1989; Barnosky *et al. 2*011). This predicament is most evident on oceanic islands where native and endemic species are

under threat due to the effects of human colonisation (Butchart *et al. 2*010). Island populations have been shown to have a much higher risk of extinction than mainland populations for a number of different reasons in addition to human activity (Wilcox & Murphy 1985; Pimm *et al.* 1988; Case & Bolger 1991; Smith *et al.* 1993; Tilman *et al.* 1994). Whilst island species represent a minority of total species in all animal and plant groups, there are still a substantial proportion of extinctions that are island species (Frankham 1997). An example is that even though only 20% of all bird species are on islands, 90% of bird species driven to extinction historically have been island species (Myers 1979). Human activity is the primary cause of island species becoming extinct over the last 50 000 years (Olson 1989), principally through over-exploitation, habitat loss / fragmentation and introducing species. These factors can cause population bottlenecks in wild populations, in addition to other underlying causes such as founder effects, disease, starvation, environmental change and other catastrophes (Wayne *et al.* 1991; Leakey & Lewin 1995; Frankham 1998).

The loss of genetic variation during a population bottleneck can reduce the population's ability to adapt and evolve. It has been estimated theoretically that small populations with Ne <1000 are more likely to go extinct due to environmental change than larger populations (Burger & Lynch 1995). This emphasises the importance of population size, changes in both its duration and magnitude such as occurs with bottlenecks, that influence the extent to which population-level processes shape genetic variation (Fisher 1930; Ellegren *et al.* 1993; Garza & Williamson 2001; Williamson-Natesan 2005). Pre- and post-bottleneck studies can directly assess this extent. For example, the northern elephant seal *Mirounga angostirostris,* was heavily exploited (over-hunted) during the nineteenth century and reduced to a bottleneck population size estimated to be 10–30 individuals (Bonnell & Selander 1974; Hoelzel *et al.* 1993). A comparison of genetic diversity in pre-bottleneck and post- bottleneck samples shows a 50% reduction in mitochondrial DNA-haplotype diversity (Hoelzel *et al. 2*002). The reduction in heterozygosity at microsatellite loci, however, was less pronounced but still observed. Other studies have also demonstrated the direct genetic consequences of bottlenecks and its relationship with population size, and used this information to inform conservation practice (Eldridge *et al.* 1999; Hedrick *et al. 2*001; Taylor *et al. 2*005; Spurgin *et al. 2*014). These studies have shown

that some species manage to persist and recover from their small numbers, thus making them ideal model systems for molecular ecology and on a broader scale, evolutionary study.

1.6 The Seychelles warbler

The Seychelles warbler *Acrocephalus sechellensis* (Fig 5) is a small (ca 12-15 g) insectivorous passerine endemic to the Seychelles archipelago (Safford & Hawkins 2013; Fig 4). It is currently listed as vulnerable on the IUCN red list since 2004, having been downgraded from critically endangered (IUCN 2015). Historically, it is thought that the Seychelles warbler existed on a number of islands within the archipelago, but the population distribution from when the islands were first settled in the 1770s, remains unclear (Komdeur 1991). However, human colonisation brought the removal of the native forest habitat in favour of coconut *Cocos nucifera* plantations. This had disastrous effects on the Seychelles warbler population (Crook 1960). The species' global population was reduced with censuses reporting as few as 26 individuals remaining on the single small island of Cousin ($4^{o}20'S$, $55^{o}40'E$, 0.29 km^2) by the 1960s (Collar & Stuart 1985) (Fig 5).

Figure 5. Adult Seychelles warbler (*Acrocephalus sechellensis*).

The crisis was recognised by the International Council for Bird Preservation (now established as BirdLife International) and a consortium was led for the island's successful purchase in 1968. As a result of this intervention and the implementation of an intensive program of habitat restoration and conservation, the population recovered and reached saturation by 1982 (Komdeur 1992), and has been relatively stable at ca 320 adults ever since (Brouwer *et al. 2*009; Wright *et al. 2*014). Four translocations have been undertaken from the source population on Cousin (Komdeur 1994; Richardson *et al. 2*006; Wright *et al. 2*014a) as part of the conservation programme managed by Nature Seychelles (Richardson 2001). A total of 29 birds was translocated to both Aride island (0.68 km^2) in 1988 and to Cousine island (0.25 km^2) in 1990 (Komdeur 1994). A further 58 birds were translocated to Denis island (1.42 km^2) in 2004 (Richardson *et al. 2*006) and 59 birds to Frégate island (2.19 km^2) in 2011 (Wright *et al. 2*014) (Fig 6). The global population now stands at ca 3500 and continues to rise (Fig 7). There is practically no inter-island dispersal (0.1%), primarily thought to be because the species has evolved an aversion to crossing open bodies of water despite being physiologically capable (Komdeur *et al. 2*004).

The Seychelles warbler has a complex cooperative breeding system where a male and female form long-term pair bonds and defend a well-defined territory all year-round (Komdeur 1992). If breeding opportunities are scarce, offspring of either sex from previous years can delay their own breeding and become subordinates either within their natal territory or even in new territories (Komdeur 1991). As subordinates these individuals often, although not always, assist as helpers to the dominant breeding pair in the construction of nests, incubation of eggs (females only) or food provision for chicks in the nest (Komdeur 1991; Richardson *et al. 2*001; Richardson *et al. 2*002). Parentage analysis has shown that in any given breeding season, 44% of female subordinates gain maternity by laying an egg in the dominant female's nest (Richardson *et al. 2*001). Male subordinates rarely gain parentage (%) in spite of high levels of extra-pair paternity (EPP) in the system, with 40% of offspring being fathered by a male other than the dominant breeding male. The extra-pair males are nearly always from outside of that territory (Richardson *et al. 2*001; Hadfield *et al. 2*006). Grandparental help also occurs within this breeding system: of the 14% of breeding females that were displaced before dying, 68% remained within the group as subordinate 'grandparent' helpers (Richardson *et al. 2*007). The primary breeding season for the

Figure 6. Map of the inner granitic Seychelles islands (main) and their position with respect to Africa (inset). Arrows indicate the islands now containing Seychelles warbler populations (Cousin, Aride, Cousine, Denis and Frégate). Historic evidence shows a past population existing on Marianne (Oustalet 1878).

Seychelles warbler is between June-September, with a secondary smaller season between November to March (Komdeur & Daan 2005); although they are known to breed all year round. The Seychelles warblers species' dynamic breeding system has been the focus of considerable study investigating the evolution of cooperative breeding, mate choice and other reproductive behaviours (Richardson *et al. 2*001; Richardson *et al. 2*002; Richardson *et al. 2*003; Komdeur 2003; Komdeur & Richardson 2007; Komdeur *et al. 2*014).

Since 1997, >96% of the Cousin population has been caught, blood-sampled and marked with a unique combination of UV-resistant colour rings and a metal British Trust for Ornithology ring (Richardson *et al. 2*002). Blood-samples and census and reproductive data are collected at least once a year during the birds' main summer breeding season, in addition to population and territory surveys. There are no natural predators for adult Seychelles warblers on Cousin Island, although a number of other species, including Seychelles fodies *Foudia sechallarum*, skinks (*Mabuya* spp.) and crabs (*Ocypode* spp.), have been known to prey on eggs and even nestlings (Veen *et al. 2*000)**.** Given that there is no inter-island dispersal, if an individual is not seen for two consecutive years it is assumed to be dead (Komdeur 1994). This means that we have access to data over the entire lifetime over the majority of birds in the population and this survival data is not confounded by dispersal. Using the blood samples, we are able to use genetic techniques to identify individual genotypes, assign parentage and determine sex (Richardson *et al. 2*001).

The Seychelles warbler makes an ideal evolutionary model because it is not confounded by gene flow and has undergone a recent severe bottleneck. Microsatellite analyses show that the Seychelles warbler has low genetic diversity as a result of the bottleneck, where the effective population size was reduced from ca 7000 in the early 1800s (as inferred from the genetic analysis of samples taken from museum specimens taken in 1877-1905) to less than 50 in the contemporary population (Spurgin *et al. 2*014). This means that the Seychelles warbler has a simpler more tractable genome of which to conduct 'bottom up' approach studies focusing on specific genes of interest. The patterns of neutral variation across individuals have been compared to that observed in functional markers i.e. MHC genes of the immune system. There is evidence that MHC class I genes have historically been under balancing selection in this species (Richardson & Westerdahl 2003). Furthermore, there is evidence of MHC-dependent extra-pair fertilisation (EPF) with females more likely to gain

EPF when their social mate had low MHC diversity. Therefore, the female would choose an extra-pair mate that had significantly higher MHC diversity than that of her social mate (Richardson *et al. 2*005). Direct associations between a specific MHC variant (*Ase-ua4*) and individual survival has also been shown (Brouwer *et al. 2*010). These significant interactions between MHC variation and fitness (mate choice and survival) give promise to further study into similar and parallel interactions of innate immune gene variation with survival (chapter 6).

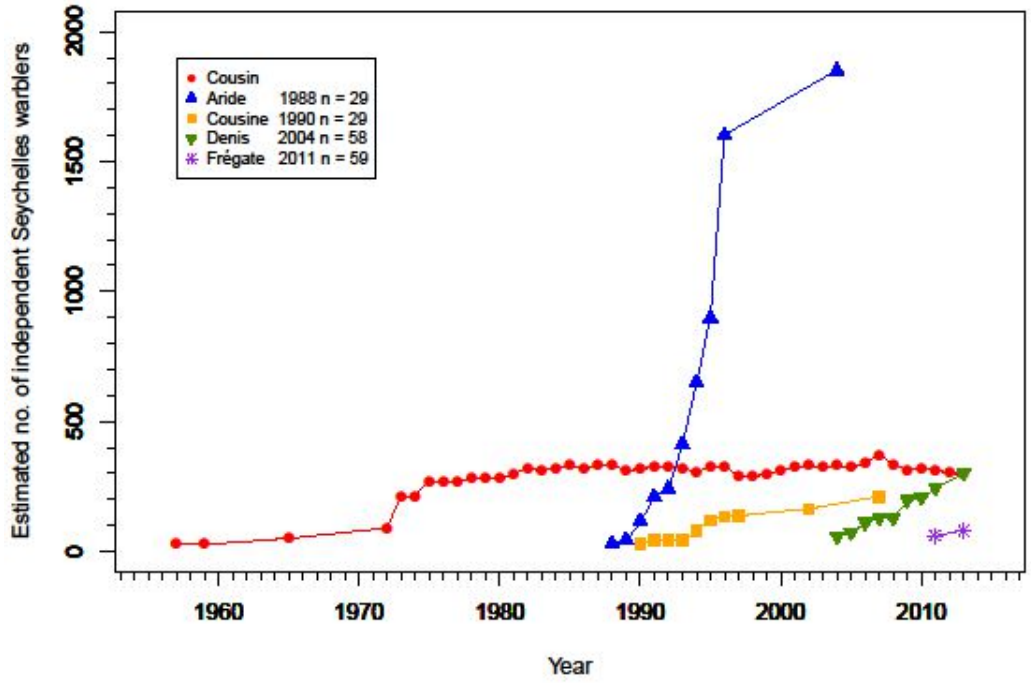


Figure 7. Seychelles warbler population trends over time on the islands of Cousin, Aride, Cousine, Denis and Frégate. Figure in R (R Core Team 2014) by Dr David Wright and Prof David S Richardson.

The Seychelles warbler is also an ideal host-parasite model for evolutionary study because it only has one parasite identified to date in its system, which is a malaria strain of *Haemoproteus* 'GRW1' (Hutchings 2009). All individuals from 1997-2014 have been screened for *Haemoproteus* and *Plasmodium* malaria parasites, in addition to *Leucocytozoan* parasites. GRW1 prevalence has been found to be significantly higher in juveniles (75%) than in adults (37%) (Hutchings 2009). No other parasite has been identified in the circulatory system and there are no gastro-intestinal parasites to our knowledge, therefore we do not have the issue of mixed or co-infections and host-parasite complexity is more tractable. Therefore, we have a simplified avian-pathogen model for which we can investigate pathogen-mediated balancing selection, which we have already shown to have

maintained variation at specific functional genes (i.e. the MHC) despite the recent bottleneck. By understanding the relative roles of neutral and selective processes, both historic and contemporary, we are able to predict the long-term persistence of the species in terms of their evolutionary potential, in the face of new challenges in the future.

1.7 Thesis outline

In this thesis, I investigate the causes of functional variation at innate immune loci in a small bottlenecked population of the Seychelles warbler. In chapter 2, I characterise variation at avian beta-defensins (AvBDs) in the contemporary Seychelles warbler population and compare this to variation at the same loci in other *Acrocephalus* species with different demographic histories. Furthermore, I focus on a specific AvBD locus in the Seychelles warbler to make a pre- and post-bottleneck comparison and assess the relative roles of drift and selection in shaping variation at this locus across the bottleneck. In chapter 3, I characterise variation at toll-like receptors (TLRs) in both the Seychelles warbler and in other *Acrocephalus* species, to carry out a detailed population genetic analysis of the evolution of this multigene family using traditional statistical methods for single-locus sequence data to detect any signatures of selection. In chapter 4, I overcome the limitations imposed by the methods used in chapter 3 by taking a forward-in-time simulation strategy to delineate the effects of demography from selection when looking at TLR variation in the Seychelles warbler. I use microsatellite diversity measures from a previous study on museum-sourced samples of this species to simulate the ancestral population of Seychelles warblers. I then define a specific demographic scenario and several selection regimes in order to determine the most likely series of events to explain the TLR variation characterised in chapter 3. In chapter 5, I investigate if there are long-term population consequences of variation at a specific TLR locus identified as potentially being under selection in chapters 3 and 4, by testing for an association between specific TLR alleles and individual survival and malaria resistance. This analysis is extended by also considering ecological factors that may influence malaria infection within the natural population. Finally, in chapter 6 I discuss my overall findings, their significance to evolutionary biology and conservation, and ideas for further research. As this thesis has been written in the style of a series of manuscripts for publication, there is some repetition, e.g. in methodology, between chapters.

1.8 References

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

Aguilar A, Roemer G, Debenham S *et al.* (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**, 3490–4.

Akilesh S, Petkova S, Sproule TJ *et al.* (2004) The MHC class I – like Fc receptor promotes humorally mediated autoimmune disease. *Journal of Clinical Investigation*, **113**, 1328–1333.

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

Allendorf FW, Hohenlohe P a, Luikart G (2010) Genomics and the future of conservation genetics. *Nature reviews. Genetics*, **11**, 697–709.

Amos W, Driscoll E, Hoffman JI (2011) Candidate genes versus genome-wide associations: which are better for detecting genetic susceptibility to infectious disease? *Proceedings. Biological sciences / The Royal Society*, **278**, 1183–1188.

Anderson PK, Cunningham A a., Patel NG *et al.* (2004) Emerging infectious diseases of plants: Pathogen pollution, climate change and agrotechnology drivers. *Trends in Ecology and Evolution*, **19**, 535–544.

Anderson K V, Jürgens G, Nüsslein-Volhard C (1985) Establishment of dorsal-ventral polarity in the Drosophila embryo: genetic studies on the role of the Toll gene product. *Cell*, **42**, 779–89.

Anderson RM, May RM (1978) Regulation and Stability of Host-Parasite Population Interactions. *Journal of Animal Ecology*, **47**, 219–247.

Areal H, Abrantes J, Esteves PJ (2011) Signatures of positive selection in Toll-like receptor (TLR) genes in mammals. *BMC evolutionary biology*, **11**, 368.

Asghar M, Hasselquist D, Hansson B *et al.* (2015) Hidden costs of infection: Chronic malaria accelerates telomere degradation and senescence in wild birds. *Science*, **347**, 436–438.

Atkinson CT, van Riper III C (1991) *Pathogenicity and epizootiology of avian haematozoa: Plasmodium, Leucocytozoon, and Haemoproteus.*

Atkinson CT, Woods KL, Dusek RJ, Sileo LS, Iko WM (1995) Wildlife disease and conservation in Hawaii: Pathogenicity of avian malaria (Plasmodium relictum) in experimentally infected Iiwi (Vestiaria coccinea). *Parasitology*, **111**, S59.

Avise JC (2010) Perspective: Conservation genetics enters the genomics era. *Conservation Genetics*, **11**, 665–669.

Baillie JEM, Griffiths J, Turvey ST, Loh J, Collen B (2010) *Evolution Lost: status and trends of the world's vertebrates*.

Barnes BW (1968) Stablising selection in Drosophila melanogaster. *Heredity*, **23**, 433–442.

Barnosky AD, Matzke N, Tomiya S *et al.* (2011) Has the Earth's sixth mass extinction already arrived? *Nature*, **471**, 51–57.

Beebee TJC, Rowe G (2004) *An Introduction to Molecular Ecology*. Oxford University Press.

Belvin MP, Anderson K V (1996) A conserved signaling pathway: the Drosophila toll-dorsal pathway. *Annual Review of Cell and Developmental Biology*, **12**, 393–416.

Bensch S, Waldenström J, Jonzén N *et al.* (2007) Temporal dynamics and diversity of avian malaria parasites in a single host species. *Journal of Animal Ecology*, **76**, 112–122.

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

Bihl F, Salez L, Beaubier M *et al.* (2003) Overexpression of Toll-like receptor 4 amplifies the host response to lipopolysaccharide and provides a survival advantage in transgenic mice. *Journal of Immunology*, **170**, 6141–50.

Bonneaud C, Perez-Tris J, Federici P, Chastel O, Sorci G (2006) Major Histocompatibility Alleles Associated with Local Resistance to Malaria in a Passerine. *Evolution*, **60**, 383–389.

Bonnell ML, Selander RK (1974) Elephant seals: genetic variation and near extinction. *Science (New York, N.Y.)*, **184**, 908–909.

Botero C a., Rubenstein DR (2012) Fluctuating environments, sexual selection and the evolution of flexible mate choice in birds. *PLoS ONE*, **7**.

Brouwer L, Barr I, van de Pol M *et al.* (2010) MHC-dependent survival in a wild population: evidence for hidden genetic benefits gained through extra-pair fertilizations. *Molecular Ecology*, **19**, 3444–3455.

Brouwer L, Tinbergen JM, Both C *et al.* (2009) Experimental evidence for density-dependent reproduction in a cooperatively breeding passerine. *Ecology*, **90**, 729–741.

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

Bruce AB (1910) The Mendelian Theory of Heredity and the Augmentation of Vigor. *Science*, **32**, 627–628.

Burger R, Lynch M (1995) Evolution and extinction in a changing environment: a quantitative-genetic analysis. *Evolution*, **49**, 151–163.

Butchart SHM, Walpole M, Collen B *et al.* (2010) Global Biodiversity: Indicators of Recent Declines. , **328**, 1164–1169.

Case TJ, Bolger DT (1991) The role of introduced species in shaping the distribution and abundance of island reptiles. *Evolutionary Ecology*, **5**, 272–290.

De Castro F, Bolker B (2005) Mechanisms of disease-induced extinction. *Ecology Letters*, **8**, 117–126.

Charlesworth B (2009) Fundamental concepts in genetics: effective population size and patterns of molecular evolution and variation. *Nature Reviews Genetics*, **10**, 195–205.

Chen H, Ma M-Y, Sun L, Fang S-G, Wan Q-H (2015) Genomic structure and evolution of beta-defensin genes in the golden pheasant and hwamei. *Science Bulletin*, **60**, 679–690.

Chesser RK (1991) Influence of gene flow and breeding tactics on gene diversity within populations. *Genetics*, **129**, 573–583.

Chow BT, Soto M, Lo BL, Crosby DC, Camerini D (2012) Antibacterial Activity of Four Human Beta-Defensins: HBD-19, HBD-23, HBD-27, and HBD-29. *Polymers*, **4**, 747–758.

Cizkova D, Gouy de Bellocq J, Baird SJ, Pialek J, Bryja J (2011) Genetic structure and contrasting selection pattern at two major histocompatibility complex genes in wild house mouse populations. *Heredity (Edinb)*, **106**, 727–740.

Collar NJ, Stuart SN (1985) *Threatened birds of Africa and related islands.* tional Council for Bird Preservation, Cambridge.

Creagh EM, O'Neill L a J (2006) TLRs, NLRs and RLRs: a trinity of pathogen sensors that co-operate in innate immunity. *Trends in Immunology*, **27**, 352–7.

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

Crow JF (1948) Alternative hypotheses of hybrid vigor. *Genetics*, **33**, 477–487.

Crow JF (1980) The estimation of inbreeding from isonymy. *Human Biology*, **61**, 935–948.

Cui J, Cheng Y, Belov K (2015) Diversity in the Toll-like receptor genes of the Tasmanian devil (Sarcophilus harrisii). *Immunogenetics*, **67**, 195–201.

Cunningham AA, Daszak P, Rodriguez JP (2003) Pathogen pollution: defining a parasitological threat to biodiversity conservation. *Journal of Parasitology*, **89**, S78–S83.

Darwin C (1859) *On the origin of species by means of natural selection.* John Murray.

Darwin C (1876) *The effects of cross and self-fertilisation in the vegetable kingdom.* J. Murray, Oxford University.

Deter J, Charbonnel N, Cosson J-F, Morand S (2007) Regulation of vole populations by the nematode Trichuris arvicolae: insights from modelling. *European Journal of Wildlife Research*, **54**, 60–70.

Diamond JM (1989) The present, past and future of human-caused extinctions. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **325**, 469–477.

Dobzhansky T (1951) *Genetics and the origin of species.* Columbia University Press, New York.

Doherty PC, Zinkernagel RM (1975) Enhanced immunological surveillance in mice heterozygous at the H-2 gene complex. *Nature*, **256**, 50–52.

Downing T, Lloyd AT, O'Farrelly C, Bradley DG (2010) The differential evolutionary dynamics of avian cytokine and TLR gene classes. *Journal of Immunology*, **184**, 6993–7000.

East EM (1908) East 1908 inbreeding in corn.pdf. *Report of the Connecticut Agricultural Experiment Station*, 419–428.

Edwards S V., Hedrick PW (1998) Evolution and ecology of MHC molecules: From genomics to sexual selection. *Trends in Ecology and Evolution*, **13**, 305–311.

Eimes J a, Bollmer JL, Whittingham L a *et al.* (2011) Rapid loss of MHC class II variation in a bottlenecked population is explained by drift and loss of copy number variation. *Journal of Evolutionary Biology*, **24**, 1847–56.

Eizaguirre C, Lenz TL, Kalbe M, Milinski M (2012) Divergent selection on locally adapted major histocompatibility complex immune genes experimentally proven in the field. *Ecology Letters*, **15**, 723–731.

Ekblom R, French L, Slate J, Burke T (2010) Evolutionary analysis and expression profiling of zebra finch immune genes. *Genome Biology & Evolution*, **2**, 781–790.

Eldridge MDB, King JM, Loupis a K *et al.* (1999) Unprecedented low levels of genetic variation and inbreeding depression in an island population of the black-footed rock-wallabi. *Conservation Biology*, **13**, 531–541.

Ellegren H, Hartman G, Johansson M, Andersson L (1993) Major histocompatibility complex monomorphism and low levels of DNA fingerprinting variability in a reintroduced and rapidly expanding population of beavers. *Proceedings of the National Academy of Sciences of the United States of America*, **90**, 8150–8153.

Emerson BC (2002) Evolution on oceanic islands: molecular phylogenetic. *Molecular Ecology*, **11**, 951–966.

Fagan WF, Holmes EE (2006) Quantifying the extinction vortex. *Ecology Letters*, **9**, 51–60.

Fernando MM a, Stevens CR, Walsh EC *et al.* (2008) Defining the role of the MHC in autoimmunity: A review and pooled analysis. *PLoS Genetics*, **4**.

Ferrer ES, García-Navas V, Sanz JJ, Ortego J (2014) Individual genetic diversity and probability of infection by avian malaria parasites in blue tits (Cyanistes caeruleus). *Journal of evolutionary biology*, 1–15.

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

Fitzpatrick MJ, Ben-Shahar Y, Smid HM *et al.* (2005) Candidate genes for behavioural ecology. *Trends in Ecology & Evolution*, **20**, 96–104.

Ford MJ (2002) Applications of selective neutrality tests to molecular ecology. *Molecular Ecology*, **11**, 1245–62.

Frankel OH (1974) Genetic conservation: our evolutionary responsibility. *Genetics*, **78**, 53–65.

Frankel OH, Soulé ME (1981) *Conservation and Evolution.* Cambridge University Press.

Frankham R (1996) Relationship of genetic variation to population size in wildlife. *Conservation Biology*, **10**, 1500–1508.

Frankham R (1997) Do island populations have less genetic variation than mainland populations? *Heredity*, **78 (Pt 3)**, 311–327.

Frankham R (1998) Inbreeding and extinction: Island populations. *Conservation Biology*, **12**, 665–675.

Frankham R (2005) Genetics and extinction. *Biological Conservation*, **126**, 131–140.

Frankham R (2010) Challenges and opportunities of genetic approaches to biological conservation. *Biological Conservation*, **143**, 1919–1927.

Frankham R, Lees K, Montgomery ME *et al.* (1999) Do population size bottlenecks reduce evolutionary potential ? *Animal Conservation*, **2**, 255–260.

Franklin IR, Frankham R (1998) How large must populations be to retain evolutionary potential? *Animal Conservation*, **1**, 69–70.

Franklin BS, Ishizaka ST, Lamphier M *et al.* (2011) Therapeutical targeting of nucleic acid-sensing Toll-like receptors prevents experimental cerebral malaria. *Proceedings of the National Academy of Sciences USA*, **108**, 3689–3694.

Friedl TWP, Groscurth E (2012) A real-time PCR protocol for simple and fast quantification of blood parasite infections in evolutionary and ecological studies and some data on intensities of blood parasite infections in a subtropical weaverbird. *Journal of Ornithology*, **153**, 239–247.

Ganz T, Selsted ME, Szklarek D *et al.* (1985) Defensins. Natural peptide antibiotics of human neutrophils. *Journal of Clinical Investigation*, **76**, 1427–1435.

Garcia de Leaniz C, Fleming I a, Einum S *et al.* (2007) A critical review of adaptive genetic variation in Atlantic salmon: implications for conservation. *Biological reviews of the Cambridge Philosophical Society*, **82**, 173–211.

Garnham PCC (1980) *Malaria in its various vertebrate hosts. In "Malaria, Epidemiology, Chemotherapy, Morphology, and Metabolism".* (JP Kreier, Ed,). Academic Press, New York.

Garrigan D, Hedrick PW (2003) Perspective: detecting adaptive molecular polymorphism: lessons from the MHC. *Evolution*, **57**, 1707–22.

Garza JC, Williamson EG (2001) Detection of reduction in population size using data from microsatellite loci. *Molecular Ecology*, **10**, 305–18.

Gibson JB, Bradley BP (1974) Stabilising selection in constant and fluctuating environments. *Heredity*, **33**, 293–302.

Gilpin ME, Soulé ME (1986) *Minimum viable populations: processes of extinction. In: Conservation Biology: The Science of Scarcity and Diversity (ed. Soule´, M.E.).* Sinauer Associates, Sunderland, MA.

Grueber CE, Carolyn J (2015) Impacts of early viability selection on management of inbreeding and genetic diversity in conservation. *Molecular Ecology*, **24**, 1645–1653.

Grueber CE, Wallis GP, Jamieson IG (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**, 4470–4482.

Grueber CE, Wallis GP, Jamieson IG (2014) Episodic positive selection in the evolution of avian toll-like receptor innate immunity genes. *PloS one*, **9**, e89632.

Hadfield JD, Richardson DS, Burke T (2006) Towards unbiased parentage assignment: Combining genetic, behavioural and spatial data in a Bayesian framework. *Molecular Ecology*, **15**, 3715–3730.

Hale K a., Briskie J V. (2009) Rapid recovery of an island population of the threatened South Island Saddleback Philesturnus c. carunculatus after a pathogen outbreak. *Bird Conservation International*, **19**, 239.

Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium*, **41**, 95–98.

Hancock REW, Sahl H-G (2006) Antimicrobial and host-defense peptides as new anti-infective therapeutic strategies. *Nature Biotechnology*, **24**, 1551–7.

Hancock RE, Scott MG (2000) The role of antimicrobial peptides in animal defenses. *Proceedings of the National Academy of Sciences of the United States of America*, **97**, 8856–8861.

Hedrick PW (2001) Conservation genetics: where are we now? *Trends in Ecology & Evolution*, **16**, 629–636.

Hedrick PW (2002) Pathogen resistance and genetic variation at MHC loci. *Evolution*, **56**, 1902–1908.

Hedrick PW (2006) Genetic Polymorphism in Heterogeneous Environments: The Age of Genomics. *Annual Review of Ecology, Evolution, and Systematics*, **37**, 67–93.

Hedrick PW, Gutierrez-Espeleta G a, Lee RN (2001) Founder effect in an island population of bighorn sheep. *Molecular Ecology*, **10**, 851–7.

Hellgren O (2015) Allelic variation at innate immune genes (avian β-defensins), within a natural population of great tits. *Journal of Avian Biology*, **46**, 113–118.

Hellgren O, Ekblom R (2010) Evolution of a cluster of innate immune genes (beta-defensins) along the ancestral lines of chicken and zebra finch. *Immunome Research*, **6**, 3.

Hellgren O, Sheldon BC (2011) Locus-specific protocol for nine different innate immune genes (antimicrobial peptides: beta-defensins) across passerine bird species reveals within-species

coding variation and a case of trans-species polymorphisms. *Molecular Ecology Resources*, **11**, 686–692.

Hellgren O, Sheldon BC, Buckling A (2010) In vitro tests of natural allelic variation of innate immune genes (avian β-defensins) reveal functional differences in microbial inhibition. *Journal of Evolutionary Biology*, **23**, 2726–2730.

Hill AVS, Allsopp CEM, Kwiatkowski D *et al.* (1991) Common West African HLA antigens are associated with protection from severe malaria. *Nature*, **352**, 595–560.

Hoelzel AR, Fleischer RC, Campagna C, le Boeuf BJ, Alvard G (2002) Impact of a population bottleck on symmetry and genetic diversity in the norhern elephant seal. *Journal of Evolutionary Biology*, **15**, 567–575.

Hoelzel a R, Halley J, O'Brien SJ *et al.* (1993) Elephant seal genetic variation and the use of simulation models to investigate historical population bottlenecks. *Journal of Heredity*, **84**, 443–9.

Hudson PJ (1986) The Effect of a Parasitic Nematode on the Breeding Production of Red Grouse. *Journal of Animal Ecology*, **55**, 85–92.

Hurst GDD, Jiggins FM, Schulenburg JHG Von Der *et al.* (1999) Male-killing Wolbachia in two species of insect. *Proceedings of the Royal Society B: Biological Sciences*, **266**, 735.

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

Jeffery KJM, Bangham CRM (2000) Do infectious diseases drive MHC diversity? *Microbes and Infection*, **2**, 1335–1341.

Jin MS, Kim SE, Heo JY *et al.* (2007) Crystal structure of the TLR1-TLR2 heterodimer induced by binding of a tri-acylated lipopeptide. *Cell*, **130**, 1071–1082.

Jin MS, Lee JO (2008) Structures of the toll-like receptor family and its ligand complexes. *Immunity*, **29**, 182–191.

Jones DF (1917) Dominance of Linked Factors as a Means of Accounting for Heterosis. *Genetics*, **2**, 466–479.

Kaiser P (2007) The avian immune genome--a glass half-full or half-empty? *Cytogenetic Genome Research*, **117**, 221–230.

Kaiser P (2010) Advances in avian immunology--prospects for disease control: a review. *Avian Pathology*, **39**, 309–324.

Keestra M, de Zoete MR, Bouwman LI, van Putten JPM (2010) Chicken TLR21 is an innate CpG DNA receptor distinct from mammalian TLR9. *Journal of Immunology*, **185**, 460–7.

Kim HM, Park BS, Kim JI *et al.* (2007) Crystal structure of the TLR4-MD-2 complex with bound endotoxin antagonist Eritoran. *Cell*, **130**, 906–917.

Kimura M (1968) Evolutionary rate at the molecular level. *Nature*, **217**, 624–626.

Kimura M (1986) DNA and the Neutral Theory. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **312**, 343–354.

Kimura M, Ohta T (1969) The Number of Heterozygous Nucleotide Sites Maintained in a Finite Population Due to Steady Flux of Mutations. *Genetics*, **61**, 763–771.

Klein J (1986) *Natural History of the Major Histocompatibility Complex.* John Wiley & Sons, New York.

Knowles SC, Wood MJ, Alves R *et al.* (2011) Molecular epidemiology of malaria prevalence and parasitaemia in a wild bird population. *Molecular Ecology*, **20**, 1062–1076.

Komdeur J (1991) *Cooperative breeding in the Seychelles warbler*. University of Cambridge.

Komdeur J (1992) Importance of habitat saturation and territory quality for evolution of cooperative breeding in the Seychelles warbler. *Nature*, **358**, 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**, 143–152.

Komdeur J (2003) Daughters on request: about helpers and egg sexes in the Seychelles warbler. *Proceedings. Biological sciences / The Royal Society*, **270**, 3–11.

Komdeur J, Burke T, Dudgale H, Richardson DS (2014) *Seychelles warblers: the complexities of the helping paradox.* (JL Koenig, WD Dickinson, Eds,). Cambridge University Press.

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

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

Komdeur J, Richardson DS (2007) Molecular Ecology Reveals the Hidden Complexities of the Seychelles Warbler. *Advances in the Study of Behaviour*, **37**, 147–187.

Kumar S, Subramanian S (2002) Mutation rates in mammalian genomes. *Proceedings of the National Academy of Sciences of the United States of America*, **99**, 803–8.

Lachish S, Knowles SC, Alves R, Wood MJ, Sheldon BC (2011) Fitness effects of endemic malaria infections in a wild bird population: the importance of ecological structure. *Journal of Animal Ecology*, **80**, 1196–1206.

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

Lafferty KD, Kimo Morris a. (1996) Altered behavior of parasitized killifish increases susceptibility to predation by bird final hosts. *Ecology*, **77**, 1390–1397.

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

Lande R (1976b) The maintenance of genetic variability by mutation in a polygenic character with linked loci. *Genetical research*, **89**, 373–387.

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

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

Landry C, Garant D, Duchesne P, Bernatchez L (2001) "Good genes as heterozygosity": the major histocompatibility complex and mate choice in Atlantic salmon (Salmo salar). *Proceedings. Biological sciences / The Royal Society*, **268**, 1279–1285.

Lanfear R, Kokko H, Eyre-Walker A (2013) Population size and the rate of evolution. *Trends in Ecology & Evolution*, **29**, 33–41.

Leakey R, Lewin R (1995) *The sixth extinction.* Weidenfeld and Nicolson, London.

Lee K a., Martin LB, Hasselquist D, Ricklefs RE, Wikelski M (2006) Contrasting adaptive immune defenses and blood parasite prevalence in closely related Passer sparrows. *Oecologia*, **150**, 383–392.

Lehrer RI, Barton a., Daher K a. *et al.* (1989) Interaction of human defensins with Escherichia coli. Mechanism of bactericidal activity. *Journal of Clinical Investigation*, **84**, 553–561.

Li Y, Xu Q, Zhang T *et al.* (2015) Host avian beta-defensin and toll-like receptor responses of pigeons following infection with pigeon Paramyxovirus type 1 virus. *Applied and Environmental Microbiology*, AEM.01413–15.

Llaurens V, McMullan M, van Oosterhout C (2012) Cryptic MHC polymorphism revealed but not explained by selection on the class IIb peptide-binding region. *Molecular Biology and Evolution*, **29**, 1631–1644.

Lynn DJ, Higgs R, Gaines S *et al.* (2004) Bioinformatic discovery and initial characterisation of nine novel antimicrobial peptide genes in the chicken. *Immunogenetics*, **56**, 170–7.

Manning CJ, Wakeland EK, Potts WK (1992) Communal nesting patterns in mice implicate MHC genes in kin recognition. *Nature*, **356**, 133–135.

Marzal A, García-Longoria L, Cárdenas Callirgos JM, Sehgal RN (2015) Invasive avian malaria as an emerging parasitic disease in native birds of Peru. *Biological Invasions*, **17**, 39–45.

Marzal A, De Lope F, Navarro C, Møller AP (2005) Malarial parasites decrease reproductive success: An experimental study in a passerine bird. *Oecologia*, **142**, 541–545.

Masatoshi N, Takeo M, Chakraborty R (1975) The Bottleneck Effect and Genetic Variability in Populations. *Evolution*, **29**, 1–10.

May RM, Anderson RM (1983) Epidemiology and genetics in the coevolution of parasites and hosts. *Proceedings of the Royal Society B: Biological Sciences*, **219**, 281–313.

Mayr E (1996) What is a species, and what is not? *Philosophy of Science*, **63**, 262–277.

McTaggart SJ, Wilson PJ, Little TJ (2012) Daphnia magna shows reduced infection upon secondary exposure to a pathogen. *Biology Letters*, 972–975.

Meade KG, Cahalane S, Narciandi F *et al.* (2008) Directed alteration of a novel bovine β-defensin to improve antimicrobial efficacy against methicillin-resistant Staphylococcus aureus (MRSA). *International Journal of Antimicrobial Agents*, **32**, 392–397.

Mendes L, Piersma T, Hasselquist D, Matson KD, Ricklefs RE (2006) Variation in the innate and acquired arms of the immune system among five shorebird species. *The Journal of experimental biology*, **209**, 284–291.

Meredith JM, Hurd H, Lehane MJ, Eggleston P (2008) The malaria vector mosquito Anopheles gambiae expresses a suite of larval-specific defensin genes. *Insect Molecular Biology*, **17**, 103–112.

Merino S, Moreno J, Sanz JJ, Arriero E (2000) Are avian blood parasites pathogenic in the wild? A medication experiment in blue tits (Parus caeruleus). *Proceedings of the Royal Society B*, **267**, 2507–10.

Mitchell-Olds T, Willis JH, Goldstein DB (2007) Which evolutionary processes influence natural genetic variation for phenotypic traits? *Nature Reviews Genetics*, **8**, 845–856.

Moller AP, Saino N (2004) Immune response and survival. *Oikos*, **104**, 299–304.

Morin PA, Luikart G, Wayne RK, The SNP workshop group (2004) SNPs in ecology, evolution and conservation. *Trends in Ecology & Evolution*, **19**, 208–216.

Morran LT, Parmenter MD, Phillips PC (2009) Mutation load and rapid adaptation favour outcrossing over self-fertilization. *Nature*, **462**, 350–352.

Mousseau T a, Roff D a (1987) Natural selection and the heritability of fitness components. *Heredity*, **59**, 181–197.

Mukherjee S, Sarkar-roy N, Wagener DK, Majumder PP (2009) Signatures of natural selection are not uniform across genes of innate immune system , but purifying. *PNAS*, **106**, 7073–7078.

Myers N (1979) *The sinking ark.* Pergamon Press, Oxford.

Nachman MW, Hoekstra HE, D'Agostino SL (2003) The genetic basis of adaptive melanism in pocket mice. *Proceedings of the National Academy of Sciences of the United States of America*, **100**, 5268–5273.

Nakajima T, Ohtani H, Satta Y *et al.* (2008) Natural selection in the TLR-related genes in the course of primate evolution. *Immunogenetics*, **60**, 727–35.

Nei M, Gojobori T (1986) Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. *Molecular Biology and Evolutionology and evolution*, **3**, 418–26.

Nei M, Tajima F (1981) Genetic drift and estimation of effective population size. *Genetics*, **98**, 625–640.

Noble ER, Noble GA, Schad GA, McInnes AJ (1989) *Parasitology: The biology of animal parasites.* Lea and Febiger, London.

O'Grady JJ, Brook BW, Reed DH *et al.* (2006) Realistic levels of inbreeding depression strongly affect extinction risk in wild populations. *Biological Conservation*, **133**, 42–51.

Ohlberger J, Langangen Ø, Edeline E *et al.* (2011) Stage-specific biomass overcompensation by juveniles in response to increased adult mortality in a wild fish population. *Ecology*, **92**, 2175–2182.

Ohta T, Ina Y (1995) Variation in Synonymous Substitution Rates Among Mammalian Genes and the Correlation Between Synonymous and Nonsynonymous Divergences. *Journal of Molecular Evolution*, **41**, 717–720.

Olsén K, Grahn M, Lohm J, Langefors Å (1998) MHC and kin discrimination in juvenile Arctic charr, Salvelinus alpinus (L.). *Animal behaviour*, **56**, 319–327.

Olson SL (1989) *Extinction on islands: man as a catastrophe.*

Oustalet E (1878) Etude sur la faune ornithologique des iles Seychelles. *Bulletin de La Société Philomáthique*, **2**, 161–206.

Padidam M, Sawyer S, Fauquet CM (1999) Possible emergence of new geminiviruses by frequent recombination. *Virology*, **265**, 218–25.

Palti Y (2011) Toll-like receptors in bony fish: from genomics to function. *Developmental and Comparative Immunology*, **35**, 1263–72.

Patricelli G, Blickley JJL (2006) Avian communication in urban noise: causes and consequences of vocal adjustment. *The Auk*, **123**, 639–649.

Pedersen AB, Greives TJ (2008) The interaction of parasites and resources cause crashes in a wild mouse population. *Journal of Animal Ecology*, **77**, 370–377.

Pertoldi C, Bijlsma R, Loeschcke V (2007) Conservation genetics in a globally changing environment: Present problems, paradoxes and future challenges. *Biodiversity and Conservation*, **16**, 4147–4163.

Peters AD, Lively CM (1999) The Red Queen and Fluctuating Epistasis: A Population Genetic Analysis of Antagonistic Coevolution. *The American Naturalist*, **154**, 393–405.

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

Piertney SB, Webster LM (2010) Characterising functionally important and ecologically meaningful genetic diversity using a candidate gene approach. *Genetica*, **138**, 419–432.

Pimm SL, Jones L, Diamond J (1988) On the risk of extinction. *The American Naturalist*, **132**, 757–785.

Podmokła E, Dubiec A, Arct A *et al.* (2015) Malaria infection status predicts extra-pair paternity in the blue tit. *Journal of Avian Biology*, **46**, 303–306.

Podmokła E, Dubiec A, Drobniak SM *et al.* (2014) Avian malaria is associated with increased reproductive investment in the blue tit. *Journal of Avian Biology*, **45**, 001–006.

Potts WK, Slev PR (1995) Pathogen-based models favoring MHC genetic diversity. *Immunological Reviews*, **143**, 181–97.

La Puente JM, Merino S, Tomás G *et al.* (2010) The blood parasite Haemoproteus reduces survival in a wild bird: a medication experiment. *Biology Letters*, **6**, 663–5.

Radwan J, Zagalska-Neubauer M, Cichon M *et al.* (2012) MHC diversity, malaria and lifetime reproductive success in collared flycatchers. *Molecular Ecology*, **21**, 2469–2479.

Redpath SM, Mougeot F, Leckie FM, Elston D a, Hudson PJ (2006) Testing the role of parasites in driving the cyclic population dynamics of a gamebird. *Ecology Letters*, **9**, 410–418.

Reed DH, Frankham R (2003) Correlation between Fitness and Genetic Diversity. *Conservation Biology*, **17**, 230–237.

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

Richardson DS (2001) *Species Conservation Assessment and Action Plan: Seychelles warbler (Timerl Dezil).*

Richardson DS, Bristol R, Shah NJ (2006) Translocation of the Seychelles warbler Acrocephalus sechellensis to establish a new population on Denis Island, Seychelles. *Conservation Evidence*, **3**, 54–57.

Richardson DS, Burke T, Komdeur J (2002) Direct benefits and the evolution of female-biased cooperative breeding in Seychelles warblers. *Evolution*, **56**, 2313–21.

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

Richardson DS, Burke T, Komdeur J (2007) Grandparent helpers: The adaptive significance of older, postdominant helpers in the Seychelles warbler. *Evolution*, **61**, 2790–2800.

Richardson DS, Jury FL, Blaakmeer K, Komdeur J, Burke T (2001) Parentage assignment and extra-group paternity in a cooperative breeder: the Seychelles warbler (Acrocephalus sechellensis). *Molecular Ecology*, **10**, 2263–73.

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

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

Roach JC, Glusman G, Rowen L *et al.* (2005) The evolution of vertebrate Toll-like receptors. *Proceedings of the National Academy of Sciences USA*, **102**, 9577–9582.

Rodriguez de Cara MA, Villanueva B, Toro MA, Fernandez J Using genomic tools to maintain diversity and fitness in conservation programmes ´.

Saccheri I, Kuussaari M, Kankare M *et al.* (1998) Inbreeding and exctinction in a butterfly metapopulation. *Nature*, **392**, 491–494.

Safford R, Hawkins F (2013) *The Birds of Africa, Volume 8: The Malagasy Region*.

Schaschl H, Wandeler P, Suchentrunk F, Obexer-Ruff G, Goodman SJ (2006) Selection and recombination drive the evolution of MHC class II DRB diversity in ungulates. *Heredity (Edinb)*, **97**, 427–437.

Schnare M, Barton GM, Holt a C *et al.* (2001) Toll-like receptors control activation of adaptive immune responses. *Nature Immunology*, **2**, 947–50.

Schonewald-Cox CM, Chambers SM, MacBryde B, Thomas WL (1983) *Genetics and Conservation: a reference for managing wild plant and animal populations.* Benjamin/Cummings, Menlo Park, California.

Schwensow N, Fietz J, Dausmann KH, Sommer S (2007) Neutral versus adaptive genetic variation in parasite resistance: importance of major histocompatibility complex supertypes in a free-ranging primate. *Heredity*, **99**, 265–77.

Selsted ME, Harwig SS, Ganz T, Schilling JW, Lehrer RI (1985) Primary structures of three human neutrophil defensins. *Journal of Clinical Investigation*, **76**, 1436–1439.

Selsted ME, Ouellette AJ (2005) Mammalian defensins in the antimicrobial immune response. *Nature Immunology*, **6**, 551–7.

Shaffer ML (1981) Minimum Population Sizes for Species Conservation. *BioScience*, **31**, 131–134.

Shull GH (1908) The composition of a field of maize. *Journal of Heredity*, 296–301.

Siemann E, Rogers WE, Dewalt SJ (2006) Rapid adaptation of insect herbivores to an invasive plant. *Proceedings. Biological sciences / The Royal Society*, **273**, 2763–2769.

Singer MC, Thomas CD, Parmesan C (1993) Rapid human-induced evolution of insect–host associations. *Nature*, **366**, 681–683.

Skevaki C, Pararas M, Kostelidou K, Tsakris A, Routsias JG (2015) Single nucleotide polymorphisms of Toll-like receptors and susceptibility to infectious disease. *Clinical and Experimental Immunology*, **180**, 165–177.

Slade RW, McCallum HI (1992) Overdominant vs. Frequency-Dependent Selection at MHC Loci. *Genetics*, **132**, 861–862.

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

Smith FD, May RM, Pellew R, Johnson TH, Walter KR (1993) How much do we know about the current extinction rate? *Trends in Ecology & Evolution*, **8**, 375–378.

Snell GD (1978) T cells, T cell recognition structures, and the Major Histocompatibility Complex. *Immunological Reviews*, **38**, 3–69.

Snow RW, Guerra CA, Abdisalan MN, Myint HY, Hay SL (2005) The global distribution of clinical episodes of Plasmodium falciparum malaria. *Nature*, **434**, 214–217.

Sol D, Jovani R, Torres J (2003) Parasite mediated mortality and host immune response explain age-related differences in blood parasitism in birds. *Oecologia*, **135**, 542–7.

Sommer S (2005) The importance of immune gene variability (MHC) in evolutionary ecology and conservation. *Frontiers in zoology*, **2**, 16.

Sorci G, Moller a. P (1997) Comparative evidence for a positive correlation between haematozoan prevalence and mortality in waterfowl. *Journal of Evolutionary Biology*, **10**, 731–741.

Soulé ME, Wilcox BA (1980) *Conservation biology. An evolutionary-ecological perspective.* Sinauer Associates Inc.

Spielman D, Brook BW, Frankham R (2004) Most species are not driven to extinction before genetic factors impact them. *Proceedings of the National Academy of Sciences of the United States of America*, **101**, 15261–15264.

Spurgin LG, Richardson DS (2010) How pathogens drive genetic diversity: MHC, mechanisms and misunderstandings. *Proceedings of the Royal Society B*, **277**, 979–988.

Spurgin LG, Wright DJ, van der Velde M *et al.* (2014) Museum DNA reveals the demographic history of the endangered Seychelles warbler. *Evolutionary Applications*, 1–10.

Staley M, Bonneaud C (2015) Immune responses of wild birds to emerging infectious diseases. *Parasite Immunology*, **37**, 242–254.

Sugiarto H, Yu PL (2004) Avian antimicrobial peptides: the defense role of beta-defensins. *Biochemical & Biophysical Research Communications*, **323**, 721–727.

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

Takahata N, Palumbi SR (1985) Extranuclear differentiation and gene flow in the finite island model. *Genetics*, **109**, 441–457.

Takeda K, Akira S (2005) Toll-like receptors in innate immunity. *International Immunology*, **17**, 1–14.

Takeuchi O, Sato S, Horiuchi T *et al.* (2002) Cutting Edge: Role of Toll-Like Receptor 1 in Mediating Immune Response to Microbial Lipoproteins. *The Journal of Immunology*, **169**, 10–14.

Taylor SS, Jamieson IG, Armstrong DP (2005) Successful island reintroductions of New Zealand robins and saddlebacks with small numbers of founders. *Animal Conservation*, **8**, 415–420.

Thoday JM (1972) Disruptive selection. *Proceedings of the Royal Society B*, 109143.

Tilman D, May RM, Lehman CL, Nowak M a. (1994) Habitat destruction and the extinction debt. *Nature*, **371**, 65–66.

Tompkins DM, Poulin R (2006) Parasites and biological invasions. *Biological Invasions in New Zealand Ecological Studies*, **186**, 67–84.

Tschirren B, Andersson M, Scherman K *et al.* (2013) Polymorphisms at the innate immune receptor TLR2 are associated with Borrelia infection in a wild rodent population Polymorphisms at the innate immune receptor TLR2 are associated with Borrelia infection in a wild rodent population.

Tu J, Li D, Li Q *et al.* (2015) Molecular Evolutionary Analysis of β -Defensin Peptides in Vertebrates. *Evolutionary Bioinformatics*, 105–114.

Veen T, Richardson DS, Blaakmeer K, Komdeur J (2000) Experimental evidence for innate predator recognition in the Seychelles warbler. *Proceedings. Biological sciences / The Royal Society*, **267**, 2253–2258.

Vinkler M, Bainová H, Bryjová A *et al.* (2015) Characterisation of Toll-like receptors 4, 5 and 7 and their genetic variation in the grey partridge. *Genetica*, **143**, 101–112.

Vinkler M, Bryjova A, Albrecht T, Bryja J (2009) Identification of the first Toll-like receptor gene in passerine birds: TLR4 orthologue in zebra finch (Taeniopygia guttata). *Tissue Antigens*, **74**, 32–41.

Vousif ME, Skibinski DOF (1982) Directional-disruptive selection in Drosophila melanogaster. *Heredity*, **49**, 71–79.

Waples R, Beechie T, Pess GR (2009) Evolutionary History, Habitat Disturbance Regimes, and Anthropogenic Changes: What Do These Mean for Resilience of Pacific Salmon Populations? *Ecology and Society*, **14**, 3.

Warren BH, Simberloff D, Ricklefs RE *et al.* (2015) Islands as model systems in ecology and evolution : prospects fifty years after MacArthur-Wilson. *Ecology Letters*, **18**, 200–217.

Watson H, Bolton M, Monaghan P (2015) Variation in early-life telomere dynamics in a long-lived bird: links to environmental conditions and survival. *Journal of Experimental Biology*, **218**, 668–674.

Watterson GA (1975) On the number of segregating sites in genetical models without recombination. *Theoretical Population Biology*, **7**, 256–276.

Wayne RK, Lehman N, Girman D *et al.* (1991) Conservation genetics of isle royale gray wolf. *Conservation Biology*, **5**, 41–51.

Westemeier RL (1998) Tracking the Long-Term Decline and Recovery of an Isolated Population. *Science*, **282**, 1695–1698.

Westerdahl H, Ashgar M, Hasselquist D, Bensch S (2010) To resist or suppress, a theoretical framework & a case study on MHC-based disease resistance. *Proceedings of the Royal Society B*.

Whittaker BRJ, Fernandez-palacios JM (1998) *Island Biogeography: Ecology, Evolution, and Conservation*. Oxford University Press.

Wilcox BA, Murphy DD (1985) Conservation strategy: the effects of fragmentation on extinction. *The American Naturalist*, **125**, 879–887.

Williamson-Natesan EG (2005) Comparison of methods for detecting bottlenecks from microsatellite loci. *Conservation Genetics*, **6**, 551–562.

Wolstenholme DR, Thoday JM (1959) Effects of disruptive selection. *Heredity*, 413–431.

Woltereck R (1928) Uber die population Frederiksborges Schloss-see Von Daphnia cucullata und einige daraus neuentstadene Erbrassen, besonders diejenige des Nemisees. *International Review of Hydrobiology*, **19**, 172–203.

Worley K, Collet J, Spurgin LG *et al.* (2010) MHC heterozygosity and survival in red junglefowl. *Molecular Ecology*, **19**, 3064–3075.

Wright S (1930) Evolution in Mendelian populations. *Bulletin of Mathematical Biology*, **52**, 241–295.

Wright DJ, Shah NJ, Richardson DS (2014a) Translocation of the Seychelles warbler Acrocephalus sechellensis to establish a new population on Frégate Island , Seychelles. *Conservation Evidence*, **11**, 20–24.

Wright DJ, Spurgin LG, Collar NJ *et al.* (2014b) The impact of translocations on neutral and functional genetic diversity within and among populations of the Seychelles warbler. *Molecular Ecology*, **23**, 2165–2177.

Xiao Y, Hughes AL, Ando J *et al.* (2004) A genome-wide screen identifies a single beta-defensin gene cluster in the chicken: implications for the origin and evolution of mammalian defensins. *BMC Genomics*, **5**, 56.

Yang D, Biragyn A, Kwak LW, Oppenheim JJ (2002) Mammalian defensins in immunity: more than just microbicidal. *Trends in Immunology*, **23**, 291–6.

Zelano B, Edwards S V (2002) An MHC component to kin recognition and mate choice in birds: predictions, progress, and prospects. *The American Naturalist*, **160 Suppl** , S225–S237.

Chapter 2: Characterisation of Avian Beta-Defensins (AvBDs) in the Seychelles warbler

Abstract

β-defensins are important components of the vertebrate innate immune system responsible for encoding a variety of anti-microbial peptides. Therefore, balancing selection is thought to act on these genes and maintain allelic variation in the face of other evolutionary forces. The Seychelles warbler, *Acrocephalus sechellensis,* is an endemic passerine that underwent a recent bottleneck in its last remaining population, resulting in a considerable reduction in genome-wide variation. We use contemporary and museum samples to investigate whether variation at avian β-defensin (AvBD) genes through this bottleneck. For comparison we examine AvBD variation across four other *Acrocephalus* species with varying demographic histories. No variation was detected at four of the six AvBD loci screened in the post-bottleneck population of Seychelles warbler, but two silent nucleotide variations were identified at *AvBD8* and one potentially functional amino-acid variation was observed at *AvBD11.* This level of variation was significantly lower than in the mainland migratory congeneric species investigated but similar to that found in other bottlenecked species. We were able to screen variation at one locus, *AvBD7,* in 15 museum specimens of Seychelles warblers taken prior to the bottleneck (1877-1905). Two alleles were observed in the pre-bottleneck population compared to the single allele in the contemporary population. Overall, the results suggest that little AvBD variation remains in the Seychelles warbler as a result the genetic drift associated with its past demographic history. While this limited immunogenetic variation may not be a problem in the face of the limited pathogen fauna that the isolated Seychelles warbler currently faces, it might be detrimental to the species' long-term persistence if new pathogens reach the population in the future.

Introduction

Genetic drift is the predominant evolutionary force shaping the genetic variation in small populations (Hedrick *et al.* 2001; Miller & Lambert 2004; Jensen *et al.* 2013), and its effects on genetic variation normally outweigh the influence of selection (Miller & Lambert 2004; Alcaide 2010; Grueber *et al.* 2014). Nevertheless, various studies have shown that within small natural populations, variation at specific key loci can be elevated above that of the genome-wide average, and be maintained across bottleneck events as a result of balancing selection (Aguilar *et al.* 2004; Tompkins 2007; van Oosterhout *et al.* 2006). Given that a loss of genetic variation within a population impacts on both inbreeding depression and adaptive potential (for review, see Garrigan & Hedrick 2003), the maintenance of polymorphisms at key loci will be important to a populations' long-term viability (Meyers & Bull 2002; Ellegren & Sheldon 2008; Zhu *et al.* 2013).

Genes that contribute to immune function are ideal candidates in which to assess the roles of drift and selection in maintaining functional genetic diversity within natural populations (for review, see Acevedo-Whitehouse & Cunningham 2006)**.** Many such studies have focused on the highly polymorphic genes of the Major Histocompatibility Complex (MHC), which play a central role in the acquired immune system (Doherty & Zinkernagel 1975; Klein 1986; Piertney & Oliver 2006). However, there are complex interacting evolutionary forces acting upon the MHC, including the effects of epistasis and selection against the so-called 'sheltered load', which is the accumulation of recessive deleterious mutations next to genes under selection (van Oosterhout 2009). Additionally, frequent gene duplication (Eimes *et al.* 2011) and recombination-like processes i.e. gene conversion (Ohta 1995; Spurgin *et al.* 2011) confound the interpretation of the population genetic mechanisms maintaining variation at these genes. In contrast, studies of variation within natural populations in genes that play a role in the innate immune system are relatively scarce (Sutton *et al.* 2011), despite the fact that these genes are often simpler in form and function than the MHC (for review, see Kaiser 2007)**.** Variation in such genes may be crucial, given that the innate immune response is the first line of defence against pathogens. Moreover, a number of innate immune gene families, including the toll-like receptors (TLRs) and cytokines, have been shown to be targets of balancing selection (for examples, see Schlenke & Begun 2003; Ferrer-admetlla *et al.* 2008; Mukherjee *et al.* 2009).

Anti-microbial peptides (AMPs) are effector molecules involved in the innate immune system and they include defensins and cathelecidins. AMPs directly kill invading pathogens via the disruption of membranes through cationic attack mechanisms (Hancock & Sahl 2006). All defensin molecules have six cysteine residues but are sorted into three classes based on their physical structure (Yang *et al.* 2002). Both α-defensins and β-defensins form beta-she*et* dimers but they have different residue lengths and pairing of cysteine linkages, whereas γ-defensins have a cyclic structure (Sugiarto & Yu 2004). Different taxonomic groups have different classes and numbers of defensins in their immune repertoire (Selsted & Ouellette 2005). For example, birds have only β-defensins, of which 14 different loci have been identified in the domestic chicken, *Gallus gallus domesticus* (Lynn *et al.* 2004; Xiao *et al.* 2004), whereas mammals have both α and β-defensins (Yang *et al.* 2002). The number of β-defensins in a species has been shown to highly relevant to the ever-changing microbial challenges of the environment in which that species' inhabits (Tu *et al.* 2015). The allelic variation which exists at these genes in a range of vertebrate hosts has been shown to have a large effect on pathogen infection *in* vitro (Meredith *et al.* 2008; Mukherjee *et al.* 2009; Hellgren *et al.* 2010; Chow *et al.* 2012). These studies show that the greater the variety of AMPs encoded, the greater the ability to combat a range of bacteria invasions. These studies therefore suggest that there could be an advantage to individuals (and populations) which are heterozygous at these loci.

Birds provide excellent systems in which to study the causes and consequences of innate immune gene variation, such as that observed within defensin genes, under natural conditions. Functional variation has been shown to exist within and among species (for review, see van Dijk *et al.* 2008) and now locus-specific protocols have been developed to screen for avian β-defensins (AvBDs) in passerines (Hellgren & Sheldon 2011). Importantly, variation within these loci has been shown to influence anti-microbial properties *in vitro* (Hellgren & Ekblom 2010; Hellgren *et al.* 2010). Specific defensin alleles have also been shown to be associated with different pathogen outcomes across various avian species (Higgs *et al.* 2007; Ma *et al.* 2012; Ramasamy *et al.* 2012), although assessing whether individual heterozygosity has a direct advantage has yet to be shown.

The Seychelles warbler, *Acrocephalus sechellensis*, is an ideal species in which to study the influence of different evolutionary forces on AvBD genes. As a result of

anthropogenic factors- this population experienced a bottleneck- during the last century when it was on the verge of extinction with ca 26 individuals remaining on a single island (Collar & Stuart 1985). As a result, considerable variation has been lost across the genome (Spurgin *et al.* 2014), although diversity appears to have been maintained at MHC class I loci (Richardson & Westerdahl 2003; Hansson & Richardson 2005) due to a combination of both natural and sexual selection (Richardson *et al.* 2005; Brouwer *et al.* 2010). Given these patterns, we hypothesise that genetic variation could also have been maintained at other immune loci. If our candidate loci have maintained variation, we can carry out association analysis between this immunogenetic variation and individual fitness parameters using data collected over the last two decades. This would allow the follow up of any initial evidence of important functional variation from this study at both individual and population levels.

Here, we screened six AvBD loci in the contemporary bottlenecked population of the Seychelles warbler. For one AvBD locus identified to be polymorphic in most other passerine species (Hellgren & Ekblom 2010), we used museum samples of the Seychelles warbler dating from 1877-1940 to assess variation that existed at this locus prior to the population bottleneck. This enabled us to compare the variation in pre- and post-bottleneck populations, at least at this locus. We also screened AvBD variation for a small sample of individuals in four other *Acrocephalus* species to gain a comparison for the levels of AvBD variation observed in the Seychelles warbler, and to test for signatures of selection within the sequences of these genes across the genus. Therefore, this study assesses levels of AvBD variation in a wild population both pre- and post-bottleneck, and across other closely-related congeneric species with different demographic histories in order to assess genetic variation across this genus.

Materials and Methods

Study species and sampling

The Seychelles warbler is a small (ca 12-15 g) insectivorous passerine endemic to the Seychelles archipelago (Safford & Hawkins 2013). As a result of anthropogenic factors, the species' global population was dramatically reduced to an estimated low of 26 individuals on the single small island of Cousin in the 1960s (Collar & Stuart 1985). This reduced the

species effective population size from 2600 - 9700 in the early 1800s to less than 50 in the contemporary population (Spurgin *et al.* 2014). After conservation intervention, the population on Cousin recovered and reached saturation by 1982 (Komdeur 1992) remaining relatively stable at ca 320 adults ever since (Brouwer *et al.* 2009; Wright *et al.* 2014). Four translocations have been undertaken from the original population on Cousin as part of a conservation programme. A total of 29 birds were translocated to both Aride in 1988 and to Cousine island in 1990 (Komdeur 1994). A further 58 birds were translocated to Denis in 2004 (Richardson *et al.* 2006) and 59 to Frégate in 2011 (Wright *et al.* 2014). This species has since been intensively studied as a model system for evolutionary, ecological and conservation questions (Komdeur 1992; Richardson *et al.* 2003; van de Crommenacker *et al.* 2011; Barrett *et al.* 2013). Since 1997, > 96% of the Cousin population has been caught, blood-sampled and marked with a unique combination of colour rings and a metal British Trust for Ornithology (BTO) ring (Richardson *et al.* 2002).

The great reed warbler, *A. arundinaceus*, and Eurasian reed warbler, *A. scirpaceus*, are two mainland migratory species classified as 'under least concern' with estimated populations (N_c) in Europe of 950,000 and 3.1 million respectively (after Hagemeijer & Blair 1997; BirdLife International 2015). In contrast, the Cape Verde warbler, *A. brevipennis*, and Henderson's Island warbler, *A. taiti*, are two island species with restricted but stable populations of an N_c estimated at 1000-1500 (Schulze-Hagen & Leisler 2011) and ca 7000 individuals (Brooke & Hartley 1995; Birdlife International 2015) respectively. The Cape Verde warbler is endemic to the Cape Verde islands and until recently, was thought to be confined to just Santiago island until small populations were discovered in São Nicolau and Fogo in 1998 and 2004, respectively (BirdLife International, 2015). All samples used in this study are from the Santiago population. The population of Henderson's Island warbler appears to have remained stable despite the observed severe population bottlenecks in other endemic species during a human invasion of the Henderson Islands in the early 1900s (Brooke 2010).

Estimates of effective population sizes (N_e) are available for the great reed warbler at ca 20 000 (Bensch & Hasselquist 1999). However, for the other warbler species with only a census population size (N_c) known, we can only estimate that the N_e will be ca 10% or less of the population size (Frankham 1995). Samples were taken from all Seychelles warbler

museum specimens known to exist (n = 26) (Spurgin *et al.* 2014) including 19 from Cousin Island and seven from Marianne Island, all collected between 1876 and 1940 (Table S1). A small (ca 1.5 x 1.5 x 3.0 mm) piece of skin was cut from the ventral surface of the foot and stored at room temperature in a sterile microfuge tube. All other *Acrocephalus* samples were from unrelated adults (> 1 year old) from single populations with details as follow: 23 individuals were sampled for the Seychelles warbler between 2000 and 2008 from the Cousin Island population (ca 320 adults, 0.3 km^2, Wright *et al.* 2014). The Cape Verde warbler samples (n = 5) were sourced from the Santiago Island population (ca 500 adults, 991 km^2, Batahla unpublished) and the Henderson's Island warbler were from Henderson Island (n = 5) (ca 7200 adults, 41 km^2, BirdLife International 2015) (Brooke & Hartley 1995). The two migratory *Acrocephalus* species *A. scirpaceus* (n = 5) and *A. arundinaceus* (n = 6), were both sampled from breeding areas in central Sweden and Belgium, respectively and used in previous studies (Hansson & Richardson 2005; Hansson *et al.* 2006).

Molecular methods

Genomic DNA was extracted from the Seychelles warbler blood samples using a salt extraction method (Richardson *et al.* 2001). The same procedure had been used for the Cape Verde warbler blood samples (provided by Juan-Carlos Illera) and the Eurasian reed warbler and the great reed warbler DNA samples (provided by Andrew Dixon and Bengt Hansson, respectively). The Henderson's Island warbler DNA samples were provided by Mike Brooke and extracted by Ian Hartley using a phenol-chloroform protocol (Brooke & Hartley 1995). We extracted DNA from Seychelles warbler museum samples (Table S1) using a Qiagen DNeasy tissue kit (Qiagen, Crawley, UK) under the manufacturer's instructions with the following changes: (i) each sample was finely chopped in a small volume of ATL buffer prior to digestion with proteinase K, (ii) 20 uL 1 $_M$ DTT (Dithiothreitol, Sigma-Aldrich, UK) was added at incubation; and (iii) 1 uL (Qiagen, final concentration = 20 pg/ml) was added during the precipitation phase. All extractions and PCRs based on historical DNA were carried out in a laminar flow cabinet in a 'clean room' isolated from the main laboratory with no record of passerine DNA use in that facility with contemporary sample controls (see Spurgin *et al.* 2014, for further details).

Locus-specific primers (Hellgren & Sheldon 2011) were used to screen six AvBD genes*: AvBD4, AvBD7, AvBD8, AvBD9, AvBD11* and *AvBD13.* These loci were chosen based on their successful amplification in congeneric species (Table S2) (Hellgren & Sheldon 2011). *AvBD7* was chosen for amplification in the museum samples because (i) it was polymorphic in most bird species examined (Hellgren *et al.* 2010), and (ii) the available primer set produced the shortest amplicon length and thus could be amplified in the degraded DNA we obtained from the museum samples, for which we have been unable to amplify fragments > 200 bp (Spurgin *et al.* 2014).

For each locus, PCRs were carried out in volumes of 10 µl with genomic DNA at a concentration of 5-10 ng / µl. Taq PCR Master Mix was used (Qiagen, UK), which included: *Taq*-DNA Polymerase, QIAGEN PCR Buffer, $MgCl_2$, and ultrapure dNTPs at optimised concentrations. PCRs were carried out using the following conditions: 30 s at 94^{o}C, 30 s at the locus-specific annealing temperature of 55^{o}C (*AvBD4, AvBD7, AvBD8, AvBD9*) and 60^{o}C (*AvBD11, AvBD13*), 45 s at 72^{o}C, all repeated for 39 cycles. All PCRs started with an incubation step of 3 min at 94^{o}C and finished with an incubation step of 10 min at 72^{o}C. PCR products were electrophoresed on a 2% agarose gel containing ethidium bromide to confirm successful amplification of the expected size fragment. Successful samples were submitted to the Genome Analysis Centre, Norwich, for Sanger-sequencing. All sequence variants were confirmed by sequencing in both the forward and reverse direction. All sequences were aligned against targ*et* sequences of the given loci from other passerine species available in the basic local alignment search tool (BLAST) nucleotide database (NCBI) using BioEdit (Hall 1999) via ClustalW codon alignment. Each chromatogram was examined by eye to identify single-nucleotide polymorphisms (SNPs) and haplotypes were phased in DnaSP (Librado & Rozas 2009). Amino acid sequences were translated in BioEdit (Hall 1999).

Phylogenetic trees were constructed in Mega v6 (Tamura *et al.* 2013) using the Maximum-Likelihood method based on the general time reversible model (Nei & Kumar 2000), to infer evolutionary history both within and between AvBD loci across the *Acrocephalus* genus. The trees with the highest log likelihood are presented, based on nucleotide variation given the short sequence sizes of < 150bp. All *Acrocephalus* sequences used originate from this study. Outgroup non-*Acrocephalus* passerine species sequences were obtained using the NCBI BLAST database and included: Eurasian blackcap, *Sylvia*

atricapilla, house sparrow, *Passer domesticus,* icterine warbler, *Hippolais icterina,* lesser redpoll, *Carduelis cabaret* and zebra finch, *Taeniopygia guttata* (Table S3). These trees were constructed to examine allelic richness at each locus for the Seychelles warbler, with further insight into AvBD loci evolution across the *Acrocephalus* genus. An overall tree was constructed to encompass all AvBD loci with the single most common allele at each locus used for each *Acrocephalus* species.

Analyses

Population size for each species was obtained from BirdLife International (2015) and their relationships with AvBD haplotype diversity was analysed using a simple linear regression in Sigmaplot from Systat Software Inc., San Jose California USA. Haplotype diversity is a measure of the uniqueness of a given haplotype in a given subset / population of individuals where its formula includes a measure of the relative haplotype frequency (x_i) in the sample of individuals and differences in sample size (N) (Nei 1987). Tests for linkage disequilibrium and deviation from the Hardy-Weinberg equilibrium (HWE) were carried out using GenePop (Raymond & Rousset 1995) and tests were based on (i) heterozygote excess and (ii) heterozygote deficiency. Polymorphism statistics and tests for neutrality were carried out in the Seychelles warbler, including: Tajima's D statistic (Tajima 1989), Fu and Li's D (Fu & Li 1993) and Fu and Li's F statistics (Fu 1996) in the program DnaSP (Librado & Rozas 2009). Z-tests of selection based on phylogenetically-variable dN / dS rates were carried out for each locus across the *Acrocephalus* genus to identify selection based on dN / dS at the haplotype-level. Site-specific dN/dS tests were then carried out using two different models (i) MEME and (ii) FUBAR to identify any individual codons under putative selection. MEME is a mixed effects model of evolution where the significance level of 0.1 is used to classify a site as positively or negatively selected as this method tends to be more conservative than empirical Bayesian approaches (Murrell *et al.* 2012). FUBAR is a fast unconstrained Bayesian approximation model using a Markov chain Monte Carlo routine which has a Bayes Factor / posterior probability set at 0.9 as a minimum value for inclusion in the inferred Bayesian graph (Murrell *et al.* 2013). Both models come highly recommended as part of the HyPhy package available for detecting individual sites under episodic diversifying selection using the DataMonkey web application (Delport *et al.* 2010).

Results

Four out of six AvBD loci were found to be monomorphic in the contemporary Seychelles warbler population (Table 1). In the two that were variable- *AvBD8* and *AvBD11*- we identified two synonymous single-nucleotide polymorphisms (SNPs) within *AvBD8* and one non-synonymous SNP within *AvBD11* (from 20 screened individuals) (Fig S1). Of the 26 museum DNA samples screened, only 15 successfully amplified the *AvBD7* locus. From these, two alleles were identified, but one allele was found in just one individual (Table S1). This novel allele, just one non-synonymous nucleotide different from the common allele, was confirmed by an independent PCR. Given the low levels of variation identified, no meaningful statistical analysis of the difference in *AvBD7* variation between the pre- and post-bottleneck populations, or the intra-specific variation at *AvBD8* and *AvBD11,* were possible. There was no evidence of selection at *AvBD8* or *AvBD11* based on the tests of neutrality or results from the Z-tests of selection based on dN/ dS (Table S4). There was no evidence found of linkage disequilibrium between all pairwise combinations of polymorphic loci. Furthermore, there was no evidence of significant deviation from Hardy-Weinberg equilibrium based on the observed allele frequencies ($P > 0.1$).

Across the *Acrocephalus* genus there was considerable variation at the AvBD loci. Five out of six loci screened were polymorphic (Table 1; Fig S1) and only *AvBD9* was monomorphic across all five *Acrocephalus* species screened. However, one of these polymorphic loci *AvBD4,* only had one SNP (and additional allele) in the Eurasian reed warbler and there was no other variation across the other species. In the Seychelles warbler, there was no evidence for selection within any of the six AvBD loci using Tajima's D, Fu and Li's F and D statistical tests ($P > 0.1$). However *AvBD8*, the most polymorphic locus observed, had significant evidence of purifying selection in the Z-test of selection looking across the *Acrocephalus* genus ($Z = 1.72$, $P = 0.04$) (Table S3).

Site-specific dN / dS based tests were carried out on *AvBD7, AvBD8* and *AvBD11* as a minimum of three unique haplotype sequences are needed. The MEME model failed to detect any sites under episodic diversifying selection across the *Acrocephalus* genus, but the FUBAR model which focuses on putative selection detected one site under diversifying selection at the *AvBD8* locus (posterior probability dN > dS = 0.90, dN - dS = 1.19). It also

detected two sites under purifying selection at the same locus (posterior probability dN < dS = 0.90 and 0.91, dN - dS = -2.89 and -0.86 respectively), in addition to one site each at *AvBD7* (posterior probability dN < dS = 0.98, dN – dS = -4.04) and *AvBD11* (posterior probability dN < dS = 0.98, dN – dS = -4.18).

The mainland migratory species, *A. arundinaceus* and *A. scirpaceus*, had significantly more nucleotide variation observed across the AvBD gene family in comparison to the island endemic species, *A. taiti, A. brevipennis* and *A. sechellensis* (t = 2.90, df = 6, *P* = 0.027) (Table 3). The mean number of alleles observed per locus for mainland migratory *Acrocephalus* species was 2 or 3, whereas the mean number of alleles was 1 or 2 for the island *Acrocephalus* species. Notably, *A. sechellensis* had the most variation observed across AvBD loci compared to the other island species. However, when only considering amino acid variation (only dN substitutions) the difference between mainland and island species loses its statistical significance (t = 2.35, df = 6, *P* = 0.057). When looking at the relationship between census population size and mean AvBD haplotype diversity, it was almost statistically significant for all nucleotide variation (t = 2.94, *P* = 0.06) (Fig 1a) but became less correlated for amino acid variation only (t = 2.14, *P* = 0.12) (Fig 1b).

The neighbour-joining trees show the levels of functional polymorphism that occur within and between the *Acrocephalus* species for each locus (Fig 2). The other outgroup passerine species consistently cluster separately from the *Acrocephalus* species for each AvBD locus. The tree for all AvBD loci combined using the single most common haplotype for each *Acrocephalus* species and the reference sequences for outgroup species, shows definite segregation by locus and confirm the independent locus-specific evolution of these immune genes (Fig 3).

Discussion

We characterised variation within the AvBD gene group in the Seychelles warbler. Four out of the six AvBD loci examined were monomorphic in the contemporary post-bottleneck population while two loci had low levels of polymorphism, with only a single nucleotide polymorphism causing a change in the protein translated at one locus (*AvBD11*). In the historical samples, we detected only two alleles, diverging by a single nucleotide substitution, in the usually highly polymorphic *AvBD7* locus (Hellgren & Ekblom 2010).

Table 1. Polymorphism indices for AvBD genes across five *Acrocephalus* species with different demographic histories. Table 1a is for the Seychelles warbler, Table 1b is for all other *Acrocephalus* species. Abbreviations include: N (number of individuals), SNP (single-nucleotide polymorphism), H (number of unique haplotypes), Hd (haplotype diversity), Pi (nucleotide diversity), dN (non-synonymous SNPs) and dS (synonymous SNPs). Standard deviation is provided in brackets.

Table 1a.

Locus	N	Size (bp)	SNPs	H	Hd (Sd)	Pi (Sd)	dN	dS
AvBD4	22	42	0	1	0.00 (0.00)	0.000 (0.000)	0	0
AvBD7	20	102	0	1	0.00 (0.00)	0.000 (0.000)	0	0
AvBD8	22	96	2	3	0.17 (0.07)	0.002 (0.001)	0	2
AvBD9	20	66	0	1	0.00 (0.00)	0.000 (0.000)	0	0
AvBD11	24	117	1	2	0.04 (0.04)	0.000 (0.000)	1	0
AvBD13	18	78	0	1	0.00 (0.00)	0.000 (0.000)	0	0

1b.

Locus	N	Species	SNPs	H	Hd (Sd)	Pi (Sd)	dN	dS
AvBD4	4	*A. arundinaceus*	0	1	0	0	0	0
	4	*A. brevipennis*	0	1	0	0	0	0
	4	*A. scirpaceus*	1	2	0.25 (0.18)	0.005 (0.0033)	1	0
	5	*A. taiti*	0	1	0	0	0	0
	22	*A. sechellensis*	0	1	0	0	0	0
AvBD7	4	*A. arundinaceus*	4	3	0.61 (0.16)	0.016 (0.0043)	3	1
	4	*A. brevipennis*	3	3	0.71 (0.12)	0.014 (0.0035)	1	2
	4	*A. scirpaceus*	1	2	0.43 (0.17)	0.004 (0.0012)	1	0
	4	*A. taiti*	0	1	0	0	0	0
	20	*A. sechellensis*	0	1	0	0	0	0
AvBD8	4	*A. arundinaceus*	1	2	0.25 (0.18)	0.0025 (0.0018)	1	1
	4	*A. brevipennis*	0	1	0	0	0	0
	4	*A. scirpaceus*	6	7	0.96 (0.08)	0.019 (0.0032)	2	4
	5	*A. taiti*	0	1	0	0	0	0
	22	*A. sechellensis*	2	3	0.17 (0.07)	0.0022 (0.001)	0	2
AvBD9	4	*A. arundinaceus*	0	1	0	0	0	0
	3	*A. brevipennis*	0	1	0	0	0	0
	4	*A. scirpaceus*	0	1	0	0	0	0
	4	*A. taiti*	0	1	0	0	0	0
	20	*A. sechellensis*	0	1	0	0	0	0
AvBD11	4	*A. arundinaceus*	2	3	0.63 (0.07)	0.0063 (0.0011)	1	1
	4	*A. scirpaceus*	1	2	0.40 (0.11)	0.0034 (0.0010)	0	1
	4	*A. taiti*	0	1	0	0	0	0
	25	*A. sechellensis*	1	2	0.04 (0.03)	0.0004 (0.0002)	1	0
AvBD13	3	*A. arundinaceus*	0	1	0	0	0	0
	4	*A. brevipennis*	0	1	0	0	0	0
	4	*A. scirpaceus*	2	3	0.61 (0.16)	0.011 (0.0037)	1	1
	1	*A. taiti*	0	1	0	0	0	0
	18	*A. sechellensis*	0	1	0	0	0	0

Figure 1. Mean AvBD haplotype diversity (Hd) verses population size (N_c) (log-transformed) in five *Acrocephalus* species; A) All nucleotide variants, B) only haplotype variants resulting in different amino-acids (putatively functional variants). Standard errors are shown. The dashed lines represent the regression and adjusted R-squared values are given.

Figure 1a.

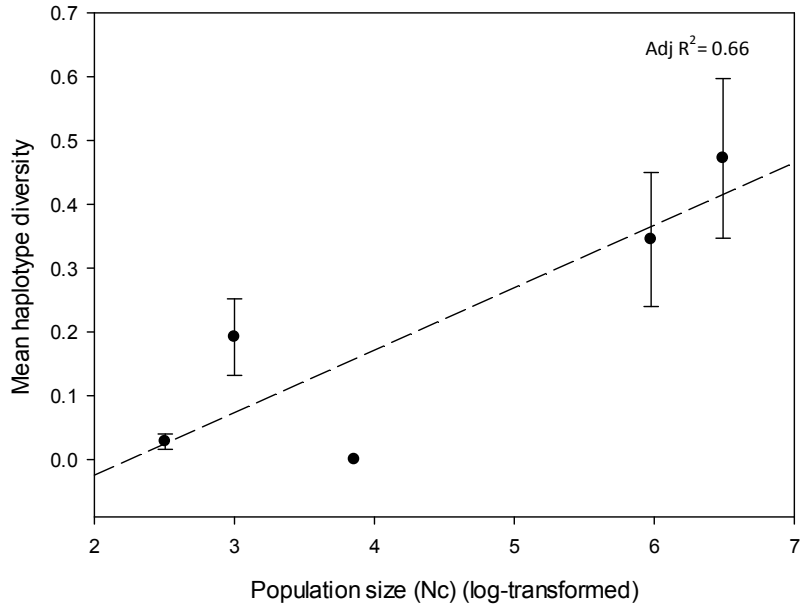


Figure 1b.

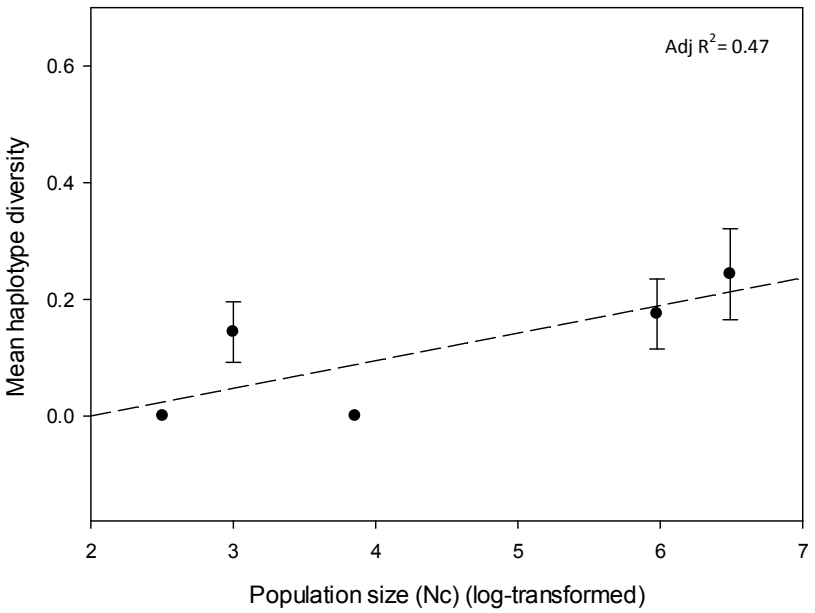
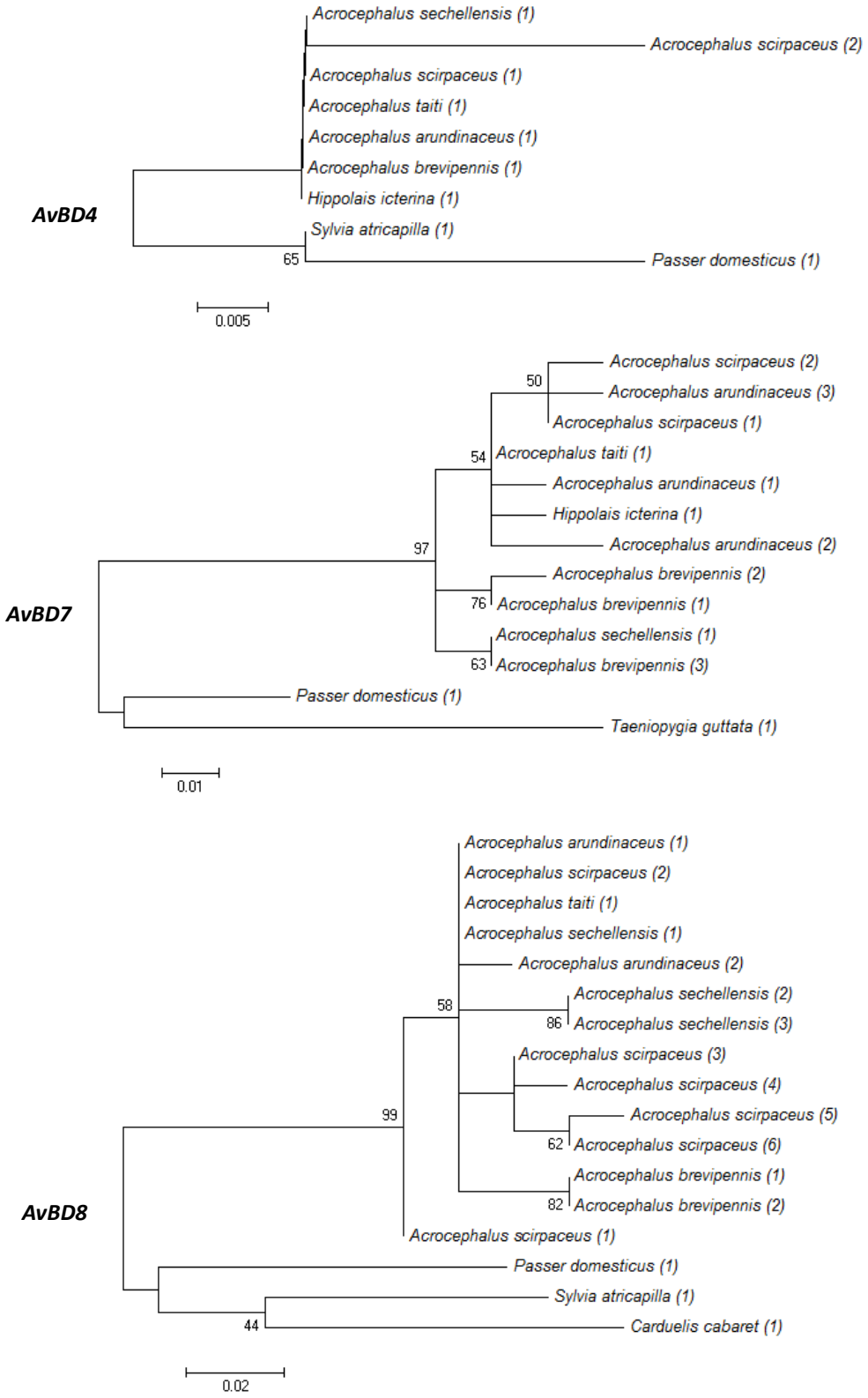
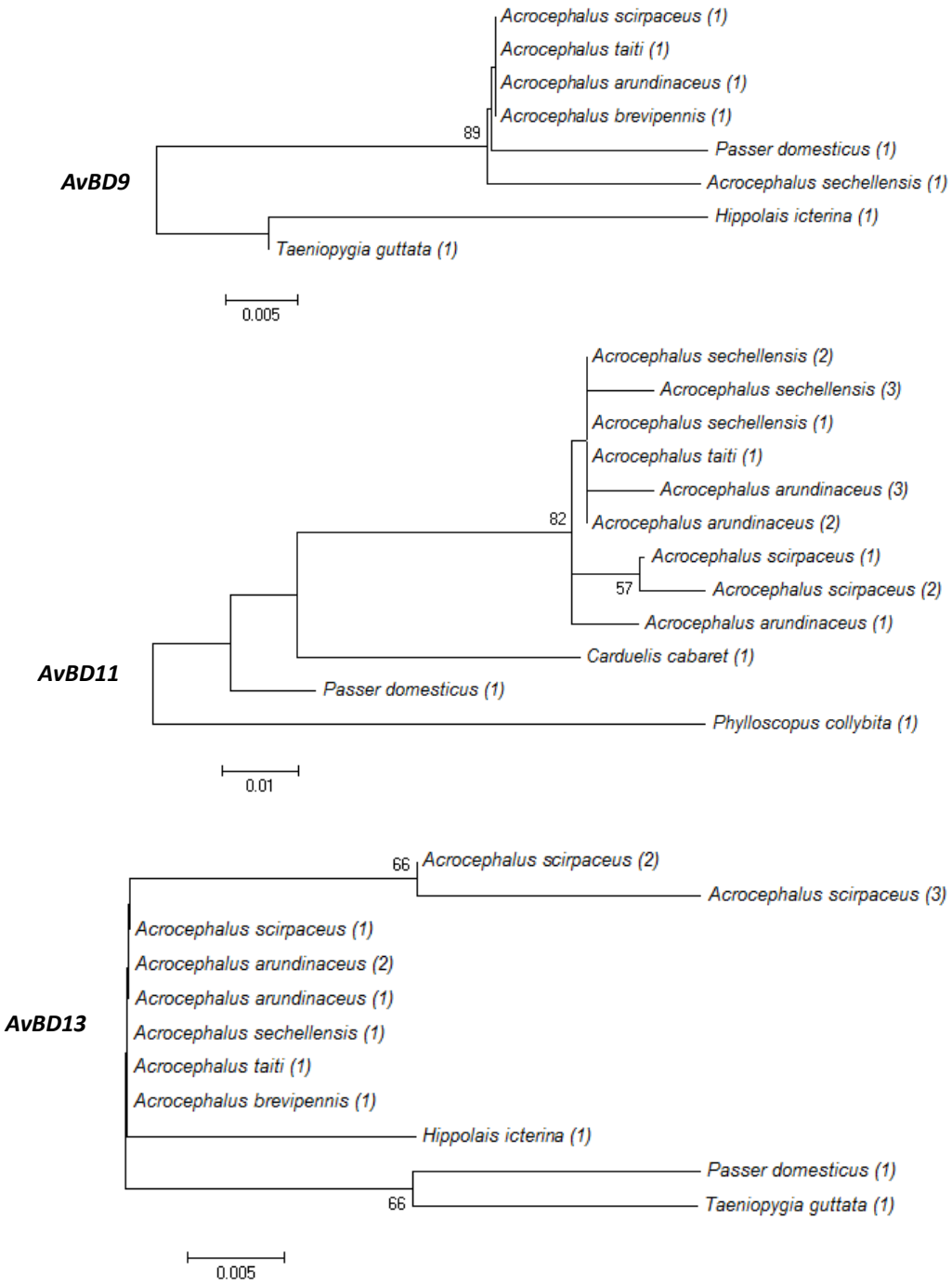


Figure 2. Trees inferring intra-locus evolutionary history of AvBD genes across five *Acrocephalus* species, inferred by using the Maximum Likelihood method based on the General Time Reversible model with applied bootstrapping of 1000 repetitions. Bootstrap values are cut-off at a threshold of 50%. Non-*Acrocephalus* passerine species are included as outgroups (see Methods). Trees are drawn to scale with haplotype number given in brackets and branch lengths measured in the number of substitutions per site.



(Figure 2 continued)



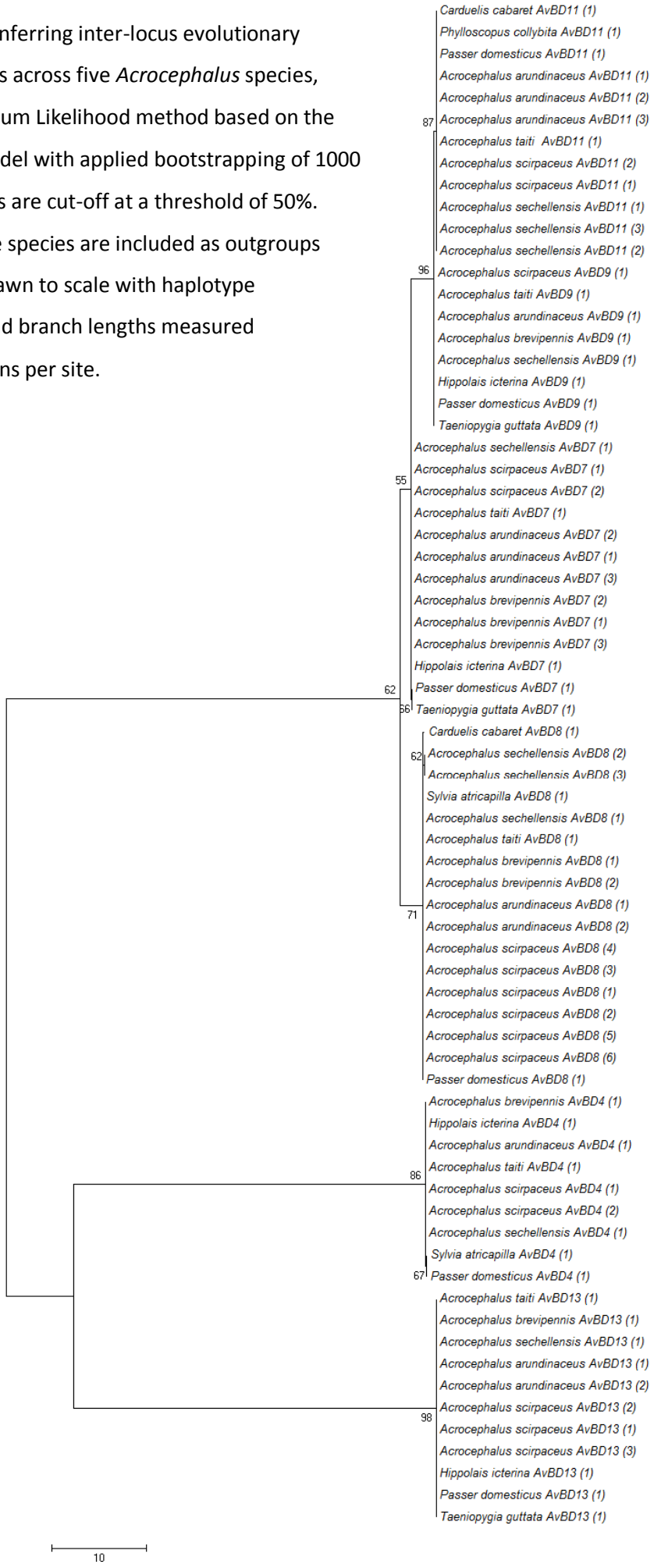


Figure 3. Phylogenetic tree inferring inter-locus evolutionary history across all AvBD genes across five *Acrocephalus* species, inferred by using the Maximum Likelihood method based on the General Time Reversible model with applied bootstrapping of 1000 repetitions. Bootstrap values are cut-off at a threshold of 50%. Non-*Acrocephalus* passerine species are included as outgroups (see Methods). Trees are drawn to scale with haplotype number given in brackets and branch lengths measured in the number of substitutions per site.

These low levels of polymorphism meant we were unable to detect signatures of selection using traditional population genetic tests. In order to increase power, we characterised variation within the AvBD gene group in a small number (3-5) individuals from four other *Acrocephalus* species' populations and looked at the same loci across the genus. One locus, *AvBD8*, was inferred to be under purifying selection, given its high ratio of synonymous substitutions (dS) compared to non-synonymous (dN) substitution across the haplotype. Looking at specific sites within the haplotype sequence, we identified one site to be under putative diversifying selection when all other loci failed to identify any sites under episodic or putative positive selection. Overall, the lack of variation at these loci in the Seychelles warbler (and other island species) indicates that balancing selection has not maintained AvBD variation in this bottlenecked population.

Considerable AvBD variation was observed in the two outbred migratory species, the great reed warbler and Eurasian reed warbler, in contrast to the three island species i.e. the Seychelles warbler, Cape Verde warbler and Henderson's Island warbler where there was little or no variation. The significant difference in nucleotide variation was almost the same for amino acid variation given there was only some difference in haplotype diversity for the mainland migratory species when only considering non-synonymous polymorphisms. This strong difference could not be explained by census population size alone and perhaps involves other demographic variables and the statistical tests may have been limited in power from the number of species included. It is clear though that being an island endemic species does reduce the ability to maintain variation. Interestingly, the recently bottlenecked Seychelles warbler has more variation at AvBD loci than the Henderson's Island warbler, despite the fact that the former species now exists within a smaller population than the latter. Henderson's Island is, however, an uplifted coral atoll at the end of a chain of small volcanic islands very isolated in the middle of the Pacific Ocean. Consequently it is highly likely that Henderson's island warbler has undergone at least one bottleneck, if not multiple sequential bottleneck events, in colonising this island, resulting in the low levels of genetic variation observed in our study of AvBDs, and at neutral genetic markers (Brooke & Hartley 1995). In contrast, until recently the Seychelles warbler existed in a larger population across multiple islands (Spurgin *et al.* 2014) and only lost ca 25% of its variation in the recent bottleneck.

Our results showing that almost no functional variation exists at the AvBD loci in the Seychelles warbler refute our *a priori* hypothesis that pathogen-mediated selection would maintain variation at these immunologically important loci. Similar losses in diversity in immune defence genes associated with bottleneck events have been reported in other endangered vertebrates (Eimes *et al.* 2011; Jamieson 2011; Basu *et al.*. 2012; Zhu *et al.* 2013)**.** The majority of polymorphic Seychelles warbler individuals are heterozygous for the rare variants observed, which suggests there may be a selective advantage with heterozygosity. However, a lack of any deviation from Hardy-Weinberg proportions suggests that this is not the case, thus it is likely that the alternate variants are merely in the heterozygous form because they are rare (and so it is unlikely that both parents possess the same rare variant to pass onto offspring). Our results do not, therefore, confirm those from an outbred population of the blue tit, *Parus major*, where all but one of 40 individuals screened showed functional heterozygosity within the exon coding for the mature defensin peptide of *AvBD2, 4, 7, 9, 10* and *12,* thus supporting a heterozygote advantage (Hellgren 2015).

Given the near-absence of variation found in both pre- and post-bottleneck populations of this species, it is impossible to statistically assess the roles that drift and selection may have played in shaping AvBD variation through this particular bottleneck. The *AvBD7* locus shows considerable intra-specific variation in other species with many nucleotide substitutions among the *Acrocephalus* genus and entire codon insertions between different families in the Passeriformes (Hellgren *et al.* 2010). At this locus n the Seychelles warbler we only detected two alleles in the population prior to the bottleneck and one thereafter. Given the low frequency of the additional allele in the historical sample (1/15 individuals) a large sample would need to be screened to confirm its absence in the contemporary population. Here we screened 20 individuals, so if the allele is present it is probably at a frequency < 0.05.

Pathogen-mediated selection (PMS) has been shown to be an important force in maintaining variation at immune genes such as the MHC and innate immune components like cytokines (Potts & Slev 1995; Jeffery & Bangham 2000; Spurgin & Richardson 2010; Turner *et al.* 2012). However, while a number of studies on β-defensins have been carried out on laboratory populations and in humans (Hollox & Armour 2008; Lazzaro 2008; Ardia *et*

al. 2011), to our knowledge there is as yet no information on PMS acting on β-defensins in wild populations. Furthermore, remote isolated populations often have fewer pathogens, as shown recently in a study of haematozoans, bacteria and viruses in avian populations (Vögeli *et al.* 2011). Indeed the diversity of pathogens in the Seychelles warbler population is very low; despite extensive screening efforts, no gastro-intestinal parasites or signs of virus infection have been detected, and only one strain of avian malaria (GRW1) has ever been observed (Hutchings 2009). This shows that stochastic processes which prevail with small island populations, not only erode immunogenetic variation (i.e. due to drift), but can reduce pathogen biodiversity (Vögeli *et al.* 2011). The combination of increased drift and reduced pathogen-mediated selection may therefore explain why variation at the AvBD genes is lost in bottlenecked island populations, such as the Seychelles warbler. In addition, if the parasite biodiversity is reduced such that only one (or a few) parasite strains are retained, the effects of pathogen-mediated selection on immunogenetic variation might be reversed. For example, the AvBD alleles observed at each locus may have become fixed in the Seychelles warbler because they provided adequate defence against the limited pathogens remaining in the environment. In such a situation, directional selection may have acted in concert with neutral effects to eliminate variation. Several studies have found that immunogenetic variation eroded faster than (neutral) microsatellite variation in small isolated populations (Bollmer *et al.* 2011; Eimes *et al.* 2011; Sutton *et al.* 2011).

In conclusion, our results show that the low levels of AvBD variation observed in the Seychelles warbler are in line with the low levels observed in other small island populations of *Acrocephalus*, and contrast to the higher levels found in mainland migratory congeneric populations. This suggests that drift may be the main force driving the patterns of variation seen these bottlenecked species. Nevertheless, it does not rule out the possibility that balancing selection may have attenuated the loss of variation caused by a reduction in population size. However in the Seychelles warbler the effect must be very limited as we only found one functional variant at just one of the five AvBD loci. It is important to report observations of invariant genes within natural populations, such as observed here in this bottlenecked species. Firstly, it prevents a publication-bias towards studies that outline where and when genes are polymorphic, potentially leading to erroneous conclusions. Secondly, studies that show depleted genetic variation at loci that are typically polymorphic

can be of conservation interest. This is because they may identify populations that are particularly vulnerable to future challenges such as pathogen infections (Frankel 1974; Hedrick 2001; Pertoldi *et al.* 2007) and can both inform and result in more effective management and prioritisation of populations and species (Schonewald-Cox *et al.* 1983; Soulé & Simberloff 1986; Frankham 2010).

Acknowledgments

Nature Seychelles kindly facilitate and support our long-term Seychelles warbler study on Cousin Island. The Seychelles Bureau of Standards and the Department of Environment gave permission for sampling and fieldwork. We thank Prof Terry Burke for the use of the NERC Biomolecular Analysis Facility at the University of Sheffield, and also would like to thank a number of collaborators for providing *Acrocephalus* DNA samples: Drs Deborah Dawson, Juan Carlos Illera, Andrew Dixon, Bengt Hansson, Michael Brooke and Ian Hartley.

Data Accession Statement

GenBank do not accept sequences which are < 200 bp, therefore, we have provided all sequences originating from this study in the supplementary material (Table S5) for easy and full access.

References

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

Aguilar A, Roemer G, Debenham S *et al.* (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**, 3490–4.

Alcaide M (2010) On the relative roles of selection and genetic drift in shaping MHC variation. *Molecular Ecology*, **19**, 3842–4.

Ardia DR, Parmentier HK, Vogel L (2011) The role of constraints and limitation in driving individual variation in immune response. *Functional Ecology*, **25**, 61–73.

Barrett ELB, Burke T a, Hammers M, Komdeur J, Richardson DS (2013) Telomere length and dynamics predict mortality in a wild longitudinal study. *Molecular Ecology*, **22**, 249–59.

Basu M, Das T, Ghosh A *et al.* (2012) Gene-gene interaction and functional impact of polymorphisms on innate immune genes in controlling Plasmodium falciparum blood infection level. *PLoS One*, **7**, e46441.

Bensch S, Hasselquist D (1999) Phylogeographic population structure of great reed warblers: an analysis of mtDNA control region sequences. *Biological Journal of the Linnean Society*, **66**, 171–185.

Bisconti R, Canestrelli D, Nascetti G (2013) Has living on islands been so simple? Insights from the insular endemic frog Discoglossus montalentii. *PLoS One*, **8**, e55735.

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

Brooke M (2010) Henderson Island. *British Birds*, **103**, 428–444.

Brooke MDL, Hartley IR (1995) Nesting Henderson Reed-Warblers (*Acrocephalus* vaughani *taiti*) Studied by DNA Fingerprinting : Unrelated Coalitions in a Stable Habitat? *The Auk*, **112**, 77–86.

Brouwer L, Barr I, van de Pol M *et al.* (2010) MHC-dependent survival in a wild population: evidence for hidden genetic benefits gained through extra-pair fertilizations. *Molecular Ecology*, **19**, 3444–3455.

Brouwer L, Tinbergen JM, Both C *et al.* (2009) Experimental evidence for density-dependent reproduction in a cooperatively breeding passerine. *Ecology*, **90**, 729–741.

Chow BT, Soto M, Lo BL, Crosby DC, Camerini D (2012) Antibacterial Activity of Four Human Beta-Defensins: HBD-19, HBD-23, HBD-27, and HBD-29. *Polymers*, **4**, 747–758.

Collar NJ, Stuart SN (1985) *Threatened birds of Africa and related islands.* tional Council for Bird Preservation, Cambridge.

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

Delport W, Poon AFY, Frost SDW, Kosakovsky Pond SL (2010) Datamonkey 2010: a suite of phylogenetic analysis tools for evolutionary biology. *Bioinformatics (Oxford, England)*, **26**, 2455–7.

Van Dijk A, Veldhuizen EJ, Haagsman HP (2008) Avian defensins. *Vetinary Immunology & Immunopathology*, **124**, 1–18.

Doherty PC, Zinkernagel RM (1975) Enhanced immunological surveillance in mice heterozygous at the H-2 gene complex. *Nature*, **256**, 50–52.

Eimes J a, Bollmer JL, Whittingham L a *et al.* (2011) Rapid loss of MHC class II variation in a bottlenecked population is explained by drift and loss of copy number variation. *Journal of Evolutionary Biology*, **24**, 1847–56.

Ellegren H, Sheldon BC (2008) Genetic basis of fitness differences in natural populations. *Nature*, **452**, 169–175.

Ferrer-admetlla A, Bosch E, Sikora M *et al.* (2008) Balancing Selection Is the Main Force Shaping the Evolution of Innate Immunity. *Journal of Immunology*, **181**, 1315–1322.

Frankel OH (1974) Genetic conservation: our evolutionary responsibility. *Genetics*, **78**, 53–65.

Frankham R (1995) Conservation genetics. *Annual review of Genetics*, **29**, 305–327.

Frankham R (2010) Challenges and opportunities of genetic approaches to biological conservation. *Biological Conservation*, **143**, 1919–1927.

Fu Y-X (1996) New statistical tests of neutrality for DNA samples from a population. *Genetics*, **143**, 557–70.

Fu Y-X, Li W-H (1993) Statistical tests of neutrality of mutations. *Genetics*, **133**, 693–709.

Garrigan D, Hedrick PW (2003) Perspective: detecting adaptive molecular polymorphism: lessons from the MHC. *Evolution*, **57**, 1707–22.

Grueber CE, Wallis GP, Jamieson IG (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**, 4470–4482.

Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium*, **41**, 95–98.

Hancock REW, Sahl H-G (2006) Antimicrobial and host-defense peptides as new anti-infective therapeutic strategies. *Nature Biotechnology*, **24**, 1551–7.

Hansson B, Jack L, Christians JK *et al.* (2006) No evidence for inbreeding avoidance in a great reed warbler population. *Behavioral Ecology*, **18**, 157–164.

Hansson B, Richardson DS (2005) Genetic variation in two endangered *Acrocephalus* species compared to a widespread congener: estimates based on functional and random loci. *Animal Conservation*, **8**, 83–90.

Hedrick PW (2001) Conservation genetics: where are we now? *Trends in Ecology & Evolution*, **16**, 629–636.

Hedrick PW, Gutierrez-Espeleta G a, Lee RN (2001) Founder effect in an island population of bighorn sheep. *Molecular Ecology*, **10**, 851–7.

Hellgren O (2015) Allelic variation at innate immune genes (avian β-defensins), within a natural population of great tits. *Journal of Avian Biology*, **46**, 113–118.

Hellgren O, Ekblom R (2010) Evolution of a cluster of innate immune genes (beta-defensins) along the ancestral lines of chicken and zebra finch. *Immunome Research*, **6**, 3.

Hellgren O, Sheldon BC (2011) Locus-specific protocol for nine different innate immune genes (antimicrobial peptides: beta-defensins) across passerine bird species reveals within-species coding variation and a case of trans-species polymorphisms. *Molecular Ecology Resources*, **11**, 686–692.

Hellgren O, Sheldon BC, Buckling A (2010) In vitro tests of natural allelic variation of innate immune genes (avian β-defensins) reveal functional differences in microbial inhibition. *Journal of Evolutionary Biology*, **23**, 2726–2730.

Higgs R, Lynn DJ, Cahalane S *et al.* (2007) Modification of chicken avian beta-defensin-8 at positively selected amino acid sites enhances specific antimicrobial activity. *Immunogenetics*, **59**, 573–580.

Hollox EJ, Armour JA (2008) Directional and balancing selection in human beta-defensins. *BMC Evolutionary Biology*, **8**, 113.

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

Jamieson IG (2011) Founder effects, inbreeding, and loss of genetic diversity in four avian reintroduction programs. *Conservation Biology*, **25**, 115–123.

Jeffery KJM, Bangham CRM (2000) Do infectious diseases drive MHC diversity? *Microbes and Infection*, **2**, 1335–1341.

Jensen H, Moe R, Hagen IJ *et al.* (2013) Genetic variation and structure of house sparrow populations: is there an island effect? *Molecular Ecology*, **22**, 1792–1805.

Kaiser P (2007) The avian immune genome--a glass half-full or half-empty? *Cytogenetic Genome Research*, **117**, 221–230.

Klein J (1986) *Natural History of the Major Histocompatibility Complex.* John Wiley & Sons, New York.

Komdeur J (1992) Importance of habitat saturation and territory quality for evolution of cooperative breeding in the Seychelles warbler. *Nature*, **358**, 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**, 143–152.

Lazzaro BP (2008) Natural selection on the Drosophila antimicrobial immune system. *Current Opinion in Microbiology*, **11**, 284–289.

Librado P, Rozas J (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics*, **25**, 1451–2.

Lynn DJ, Higgs R, Gaines S *et al.* (2004) Bioinformatic discovery and initial characterisation of nine novel antimicrobial peptide genes in the chicken. *Immunogenetics*, **56**, 170–7.

Ma D, Zhou C, Zhang M *et al.* (2012) Functional analysis and induction of four novel goose (Anser cygnoides) avian beta-defensins in response to salmonella enteritidis infection. *Comparative Immunology, Microbiology & Infectious Diseases*, **35**, 197–207.

Meredith JM, Hurd H, Lehane MJ, Eggleston P (2008) The malaria vector mosquito Anopheles gambiae expresses a suite of larval-specific defensin genes. *Insect Molecular Biology*, **17**, 103–112.

Meyers LA, Bull JJ (2002) Fighting change with change: Adaptive variation in an uncertain world. *Trends in Ecology and Evolution*, **17**, 551–557.

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

Mukherjee S, Sarkar-roy N, Wagener DK, Majumder PP (2009) Signatures of natural selection are not uniform across genes of innate immune system , but purifying. *PNAS*, **106**, 7073–7078.

Murrell B, Moola S, Mabona A *et al.* (2013) FUBAR: a fast, unconstrained bayesian approximation for inferring selection. *Molecular Biology and Evolutioniology and evolution*, **30**, 1196–205.

Murrell B, Wertheim JO, Moola S *et al.* (2012) Detecting individual sites subject to episodic diversifying selection. *PLoS genetics*, **8**, e1002764.

Nei M (1987) *Molecular evolutionary genetics.* Columbia University Press, New York.

Nei M, Kumar S (2000) *Molecular Evolution and Phylogenetics*. Oxford University Press, New York.

Ohta T (1995) Gene Conversion vs Point Mutation in Generating Variability at the Antigen Recognition Site of Major Histocompatibility Complex Loci. *Journal of Molecular Evolution*, **41**, 115–119.

Van Oosterhout C (2009) A new theory of MHC evolution: beyond selection on the immune genes. *Proceedings. Biological sciences / The Royal Society*, **276**, 657–65.

Van Oosterhout C, Joyce DA, Cummings SA *et al.* (2006) Balancing Selection, Random Genetic Drift, and Genetic Variation at the Major Histocompatibility Complex in Two Wild Populations of Guppies (Poecilia reticulata). *Evolution*, **60**, 2562–2574.

Pertoldi C, Bijlsma R, Loeschcke V (2007) Conservation genetics in a globally changing environment: Present problems, paradoxes and future challenges. *Biodiversity and Conservation*, **16**, 4147–4163.

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

Potts WK, Slev PR (1995) Pathogen-based models favoring MHC genetic diversity. *Immunological Reviews*, **143**, 181–97.

Ramasamy KT, Verma P, Reddy MR (2012) Differential gene expression of antimicrobial peptides beta defensins in the gastrointestinal tract of Salmonella serovar Pullorum infected broiler chickens. *Vetinary Research Communications*, **36**, 57–62.

Raymond M, Rousset F (1995) GENEPOP (Version 1.2): Population Genetics Software for Exact Tests and Ecumenicism. *Heredity*, **86**, 248–249.

Richardson DS, Bristol R, Shah NJ (2006) Translocation of the Seychelles warbler *Acrocephalus sechellensis* to establish a new population on Denis Island, Seychelles. *Conservation Evidence*, **3**, 54–57.

Richardson DS, Burke T, Komdeur J (2002) Direct benefits and the evolution of female-biased cooperative breeding in Seychelles warblers. *Evolution*, **56**, 2313–21.

Richardson DS, Jury FL, Blaakmeer K, Komdeur J, Burke T (2001) Parentage assignment and extra-group paternity in a cooperative breeder: the Seychelles warbler (*Acrocephalus* sechellensis). *Molecular Ecology*, **10**, 2263–73.

Richardson DS, Komdeur J, Burke T, Lusk C (2003) Altruism and infidelity among warblers. *Nature*, **422**, 580–581.

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

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

Safford R, Hawkins F (2013) *The Birds of Africa, Volume 8: The Malagasy Region*.

Schlenke TA, Begun DJ (2003) Natural Selection Drives Drosophila Immune System Evolution. *Genetics*, **164**, 1471–1480.

Schonewald-Cox CM, Chambers SM, MacBryde B, Thomas WL (1983) *Genetics and Conservation: a reference for managing wild plant and animal populations.* Benjamin/Cummings, Menlo Park, California.

Schulze-Hagen K, Leisler B (2011) *Reed Warblers: Diversity in a Uniform Bird Family.* KNNV Uitgeverij; 1 edition.

Selsted ME, Ouellette AJ (2005) Mammalian defensins in the antimicrobial immune response. *Nature Immunology*, **6**, 551–7.

Soulé ME, Simberloff D (1986) What do genetics and ecology tell us about the design of nature reserves? *Biological Conservation*, **35**, 19–40.

Spurgin LG, van Oosterhout C, Illera JC *et al.* (2011) Gene conversion rapidly generates major histocompatibility complex diversity in recently founded bird populations. *Molecular Ecology*, **20**, 5213–5225.

Spurgin LG, Richardson DS (2010) How pathogens drive genetic diversity: MHC, mechanisms and misunderstandings. *Proceedings of the Royal Society B*, **277**, 979–988.

Spurgin LG, Wright DJ, van der Velde M *et al.* (2014) Museum DNA reveals the demographic history of the endangered Seychelles warbler. *Evolutionary Applications*, 1–10.

Sugiarto H, Yu PL (2004) Avian antimicrobial peptides: the defense role of beta-defensins. *Biochemical & Biophysical Research Communications*, **323**, 721–727.

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

Tajima F (1989) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics*, **123**, 585–95.

Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Molecular Biology and Evolution*, **24**, 1596–9.

Tompkins DM (2007) Population bottlenecks and avian immunity: implications for conservation. *Animal Conservation*, **10**, 11–13.

Tu J, Li D, Li Q *et al.* (2015) Molecular Evolutionary Analysis of β -Defensin Peptides in Vertebrates. *Evolutionary Bioinformatics*, 105–114.

Turner AK, Begon M, Jackson JA, Paterson S (2012) Evidence for selection at cytokine loci in a natural population of field voles (Microtus agrestis). *Molecular Ecology*, **21**, 1632–1646.

Vögeli M, Lemus J a, Serrano D, Blanco G, Tella JL (2011) An island paradigm on the mainland: host population fragmentation impairs the community of avian pathogens. *Proceedings of the Royal Society B*, **278**, 2668–76.

Wright DJ, Shah NJ, Richardson DS (2014a) Translocation of the Seychelles warbler *Acrocephalus sechellensis* to establish a new population on Frégate Island , Seychelles. *Conservation Evidence*, **11**, 20–24.

Wright DJ, Spurgin LG, Collar NJ *et al.* (2014b) The impact of translocations on neutral and functional genetic diversity within and among populations of the Seychelles warbler. *Molecular Ecology*, **23**, 2165–2177.

Xiao Y, Hughes AL, Ando J *et al.* (2004) A genome-wide screen identifies a single beta-defensin gene cluster in the chicken: implications for the origin and evolution of mammalian defensins. *BMC Genomics*, **5**, 56.

Yang D, Biragyn A, Kwak LW, Oppenheim JJ (2002) Mammalian defensins in immunity: more than just microbicidal. *Trends in Immunology*, **23**, 291–6.

Young A, Boyle T, Brown T (1996) The population genetic consequences of habitat fragmentation for plants. *Trends in Ecology & Evolution*, **11**, 413–418.

Zhu L, Hu Y, Qi D *et al.* (2013) Genetic consequences of historical anthropogenic and ecological events on giant pandas. *Ecology*, **94**, 2346–57.

Supplementary material

Table S1. Details of the Seychelles warbler museum samples used to amplify the *AvBD7* gene (modified from Spurgin *et al.* 2014).

Year	Sample ID	Museum Reference	Sex	Island	Successful MHC screen	Successful *AvBD7* screen
1876	10	1876-377	-	Marianne	X	
1876	11	1876-574	-	Marianne		
1878	12	1878-552	Male	Marianne	X	X
1878	13	1878-553	Male	Marianne	X	X
1877	23	27/Syl/11/b/1	Male	Marianne	X	X
1877	25	27/Syl/11/b/3	Female	Marianne	X	X
1878	17	1878.7.30.3	Male	Marianne	X	X
1888	18	1927.12.18.391	Female	Cousin	X	X
1888	19	1927.12.18.395	Female	Cousin	X	X
1888	24	27/Syl/11/b/2	Male	Cousin	X	X
1890	1	USNM 119752	Male	Cousin	X	
1890	2	USNM 119753	Female	Cousin		
1904	3	SKIN 265502	Male	Cousin	X	
1904	4	SKIN 596991	Male	Cousin	X	
1904	5	SKIN 596992	Male	Cousin	X	X
1904	6	SKIN 596993	Female	Cousin	X	
1904	7	SKIN 596994	Female	Cousin	X	
1904	8	SKIN 596995	Male	Cousin	X	
1904	9	SKIN 596996	Male	Cousin	X	
1904	26	140287	Male	Cousin	X	
1905	14	CG1938-897	Male	Cousin	X	X
1905	15	CG1938-898	Male	Cousin	X	X
1905	16	CG1938-899	Male	Cousin	X	X*
1940	20	1946.75.23	Male	Cousin	X	X
1940	21	1946.75.24	Male	Cousin	X	X
1940	22	1946.75.25	Female	Cousin	X	X

* Polymorphism identified

Table S2. AvBD genes screened in other passerine species* (modified from Hellgren *et al.* 2011). Key: X = successful amplification. Two individuals were screened for each non-*Acrocephalus* species and so haplotypes cannot be phased from this information given the sample size.

Species	*AvBD4*	*AvBD7*	*AvBD8*	*AvBD9*	*AvBD11*	*AvBD13*
Blue tit (*Cyanistes caeruleus*)	X	X		X		
Great tit (*Parus major*)	X*	X*		X		
Eurasian reed warbler (*Acrocephalus scirpaceus*)	X	X		X		X
Great reed warbler (*Acrocephalus arundinaceus*)	X	X*	X	X	X*	X
Chiffchaff (*Phylloscopus collybita*)	X			X	X	
Willow warbler (*Phylloscopus trochilus*)	X		X	X	X*	X
Icterine warbler (*Hippolais icterina*)	X	X		X	X	X
Garden warbler (*Sylvia borin*)	X*	X		X		X
Blackcap (*Sylvia atricapilla*)	X	X	X	X		X
House sparrow (*Passer domesticus*)	X	X	X*	X	X	X
Blackbird (*Turdus merula*)	X		X	X		
Redwing (*Turdus iliacus*)	X	X*		X*		
Spotted flycatcher (*Muscicapa striata*)			X	X		
Bluethroat (*Luscinia svecica*)	X			X		
Redstart (*Phoenicurus phoenicurus*)		X	X	X		X
Common Redpoll (*Carduelis flammea*)	X*	X		X	X	X
Siskin (*Spinus spinus*)	X	X	X	X	X	
Zebrafinch (*Taeniopygia guttata*)		X		X	X	X
Rock firefinch (*Lagonisticta sanguinodorsalis*)	X	X	X	X		
Red-backed shrike (*Lanius collurio*)	X*		X	X	X	X
Total no. polymorphic sites	21	51	31	18	53	21
Total no. variable amino acid sites	8	25	13	6	20	11

* = non-synonymous SNPs that were found in the exon encoding for the anti-microbial peptide

Table S3. Details of additional sequences used for alignments from the NCBI BLAST database.

Locus	Species	Sequence	Accession ID
AvBD4	Icterine warbler (*Hippolais icterina*)	GGAAATGCCCTCGTGGCAACGATTACCTGGGGTCATGTCGTCCTG	GU551968.1
	House sparrow (*Passer domesticus*)	GGAAATGCCCTCGTGGCAACGATTACCTGGGGTCGTGTCGTTCTG	GU551985.1
	Eurasian blackcap (*Sylvia atricapilla*)	GGAAATGCCCTCGTGGCAACGATTACCTGGGGTCATGTCGTCCTG	GU551988.1
AvBD7	Icterine warbler (*Hippolais icterina*)	TTCTCCTTGTGCTGTGCAGGACAAGAAGTGTTTTCTAGGCTAGAATAATTCCTGTTTGATCCAAAACGGACGCTGCTTCCCAGGGATTTGTCGTCGCCCTTACTACTGGATTGGGGAGTGTAGCAATGGATATTCTTGCTGCAAAAGG	GU552005.1
	House sparrow (*Passer domesticus*)	TTCTCTTTGTGCTGTGCAGGACAAGTGTTTCCTAGGCTAGACAATTCTTGTTTTATCCAAAACGGACGCTGCTTCCCAGGGATTTGCCGTCGCCCTTATTACTGGATTGGAACATGTAGCAATGGATATTCTTGCTGCAAAAGG	GU552013.1
	Zebrafinch (*Taeniopygia guttata*)	TTCTCTTTGTGCTGTACAGCACAATTGTTTCCTAGGCTAAACAACCCTTGTTTGATGCAAAATGGACGCTGCTTCCGAGGGATTTGTCGCCGCCCTTATTACTGGATTGGAACGTGTAGCAATGGATATTCTTGCTGCAAAAGG	GU552011.1
AvBD8	Lesser redpoll (*Carduelis cabaret*)	CGTGCCCCCAGCACCGAGGTGCAGTGCAGACAAGCTGGGGGTGTCTGTTCCCACCACTGCCCCCTGCCCCACAGGAGACCCTTTGGAAGATGCCAGCAGGGAATTCCCTGCTGT	GU552025.1
	House sparrow (*Passer domesticus*)	CGTGCCCCCAACACCGAGGTGCAGTGCAGGAAGGCTGGGGGGGTCTGTTCCGACCGCTGCCCCCCGCCCCACTCGAGGCCCTTTGGGCGCTGCCAGCAGGGAATTCCCTGCTGT	GU552030.1
	Eurasian blackcap (*Sylvia atricapilla*)	CGTGCCCCCAACACCGAGGCACAGTGCAGCAAGGCTGGGGGGGTCTGCTCCCACCACTGCCCTCAGCCCCACACCAGACCCTTTGGACGCTGCCAGCAGGGAATTCCCTGCTGT	GU552031.1
AvBD9	Icterine warbler (*Hippolais icterina*)	GCTGACACCCTTGCATGCCGGCAGAACCGGGGCTCCTGCTCCTTCGTGCCCTGCTCTGCTCCTCTGGTTGACATCGGCACCTGCCGTGGAGGGAAGCTG	GU552044.1
	House sparrow (*Passer domesticus*)	GCTGACACCCTCGCCTGCCGGCAGAGCCGGGGCTCCTGCTCCTTCGTGCCCTGCTCTGCCCCTCTGGTTGACATCGGGACCTGCCGCGGTGGGAAGCTA	GU552054.1
	Zebrafinch (*Taeniopygia guttata*)	GCTGACACCCTCGCATGCCGGCAGAGCCGGGGCTCCTGCTCCTTCGTGCCCTGCTCTGCTCCTCTGGTTGACATCGGCACCTGCCGTGGTGGGAAGCTA	GU552052.1
AvBD11	Lesser redpoll (*Carduelis cabaret*)	GTCCAGGGACACCTCACGTTGTTTGGAATACCACGGCTACTGCTTCCACCTGAAATCCTGCCCGGAGCCCTTCGCTGCCTTTGGGACTTGCTATCGGCGCCGGAGGACCTGCTGCGT	GU552089.1
	House sparrow (*Passer domesticus*)	GCCCAGGGACACCTTGCGTTGTTTGGAATACCACGGATACTGCTTCCACCTGAAATCCTGCCCAGAGCCCTTTGCTGCCTTTGGGACTTGCTATCGGCGCCGGAGGACCTGCTGTGT	GU552093.1
	Zebrafinch (*Taeniopygia guttata*)	GCCCAGGGACACCTTGCGTTGTTTGGAATACCACGGCTACTGCTTCCACCTGAAATCCTGCCCAGAGCCGTTCGCTGCCTTTGGAACCTGCTATCGGCGCCGCAGGACCTGCTGCCT	GU552092.1
AvBD13	Icterine warbler (*Hippolais icterina*)	GCAGAAGCAACCGTGGGCACTGCCGGAGGCTCTGCTTCCACATGGAGCGCTGGGAGGGGAGCTGCAGCAACGGCCGCCTG	GU551975.1
	House sparrow (*Passer domesticus*)	GCAGAAGCAACCGTGGCCACTGCCGGAGGCTCTGCTTCCACATGGAGCGCTGGGAAGGGAGCTGCAGCAGCGGCCGCCTG	GU552137.1
	Zebrafinch (*Taeniopygia guttata*)	GCAGAAACAACCGTGGCCACTGCCGGAGGCTCTGCTTCCACATGGAGCGCTGGGAAGGGAGCTGCAGCAACGGACGCCTG	GU552136.1

Table S4. Tests of neutrality and the dN / dS Z-tests of selection on polymorphic AvBD loci sequences in five *Acrocephalus* species including the Seychelles warbler. Table 4a presents results for the Seychelles warbler and table 4b for all *Acrocephalus* warblers screened including the Cape Verde warbler, Great reed warbler, Henderson's Island warbler and Eurasian reed warbler.

Table S4a.

Locus	Tajima's D (*P*)	Fu and Li's F (*P*)	Fu and Li's F (*D*)	Z (dN = dS) (*P*)	Z (dN > dS) (*P*)	Z (dN < dS) (*P*)
AvBD8	-0.98 (> 0.1)	-0.86 (> 0.1)	-1.04 (> 0.1)	-1.22 (> 0.1)	-1.20 (> 0.1)	1.19 (> 0.1)
AvBD11	-1.11 (> 0.1)	-1.83 (> 0.1)	-1.87 (> 0.1)	1.01 (> 0.1)	1.04 (> 0.1)	-0.99 (> 0.1)

Table S4b.

Locus	Z (dN = dS) (*P*)	Z (dN > dS) (*P*)	Z (dN < dS) (*P*)
AvBD4	1.09 (> 0.1)	1.06 (> 0.1)	-1.05 (> 0.1)
AvBD7	-1.21 (> 0.1)	-1.30 (> 0.1)	1.35 (>0.05)
AvBD8	-1.84 (> 0.05)	-1.86 (> 0.1)	1.72 (0.04)*
AvBD11	0.31 (> 0.1)	0.31 (> 0.1)	-0.32 (> 0.1)
AvBD13	-0.17 (> 0.1)	-0.18 (> 0.1)	0.17 (> 0.1)

Significance: ** ($P < 0.01$) * ($P < 0.05$)

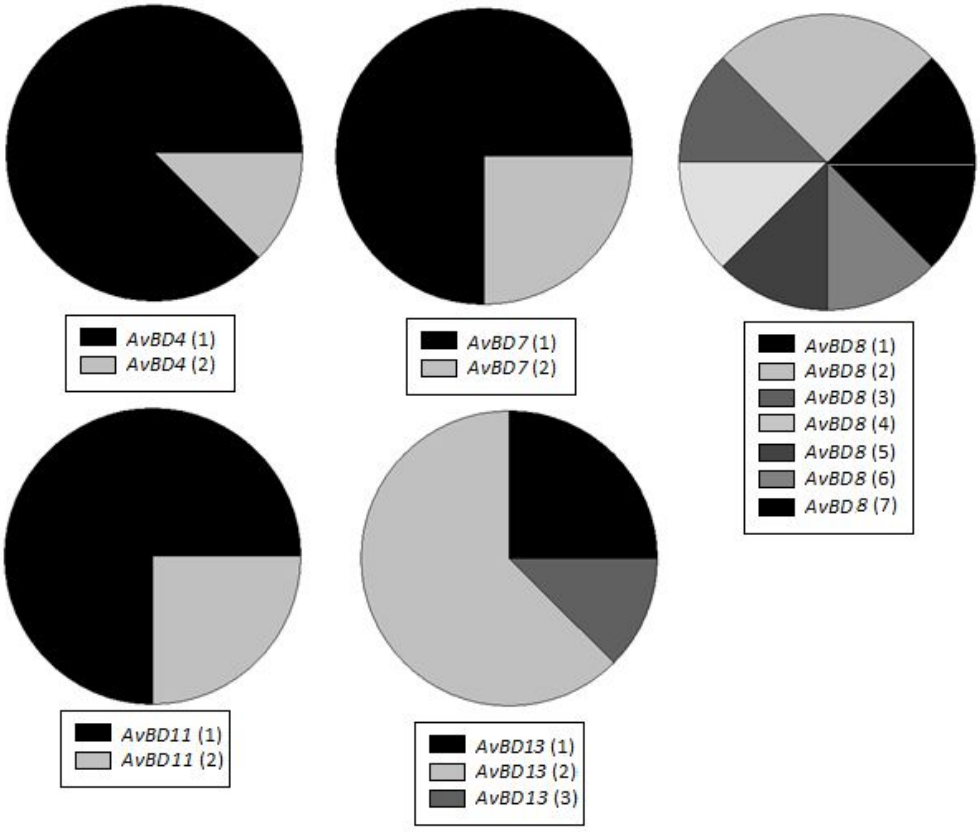
Table S5. Genomic DNA sequences for AvBD loci for all *Acrocephalus* species used in this study.

AvBD4	
Species	**Genomic sequence**
A. sechellensis	TGCCCTCGTGGCAACGATTACCTGGGGTCATGTCGTCCTGGGTACAGTTGCTGT
A. brevipennis	TGCCCTCGTGGCAACGATTACCTGGGGTCATGTCGTCCTGGGTACAGTTGCTGT
A. arundinaceus	TGCCCTCGTGGCAACGATTACCTGGGGTCATGTCGTCCTGGGTACAGTTGCTGT
A. taiti	TGCCCTCGTGGCAACGATTACCTGGGGTCATGTCGTCCTGGGTACAGTTGCTGT
A. scirpaceus (1)	TGCCCTCGTGGCAAGGATTACCTGGGGTCATGTCGTCCTGGGTACAGTTGCTGT
A. scirpaceus (2)	TGCCCTCGTGGCAACGATTACCTGGGGTCATGTCGTCCTGGGTACAGTTGCTGT
AvBD7	
A. sechellensis	GAAGTGTTTTCTAGGCTAGATAATTCCTGTTTGATCCAAAATGGACGCTGCTTCCCAGGGATTTGTC GTCGCCCTTATTACTGGATTGGAGAGTGTAGCAAT
A. brevipennis (1)	GAAGTGTTTTCTAGGCTAGATAATTCCTGTTTGATCCAAAACGGCCGCTGCCTCCCAGGGATTTGTC GTCGCCCTTATTACTGGATTGGAGAGTGTAGCAAT
A. brevipennis (2)	GAAGTGTTTTCTAGGCTAGATAATTCCTGTTTGATCCAAAACGGCCGCTGCTTCCCAGGGATTTGTC GTCGCCCTTATTACTGGATTGGAGAGTGTAGCAAT
A. brevipennis (3)	GAAGTGTTTTCTAGGCTAGATAATTCCTGTTTGATCCAAAATGGACGCTGCTTCCCAGGGATTTGTC GTCGCCCTTATTACTGGATTGGAGAGTGTAGCAAT
A. arundinaceus (1)	GAAGTGTTTTCTAGGCTAGATAATTCCTGTTTGATCCAAAATGGACGCTGCTTCCCAGGGATTTGTC GTCGCCCTTATTACTGGATTGGGGACTGTAGCAAT
A. arundinaceus (2)	GAAGTGTTTTCTAGGCTAGATAATTCCTGTTTGATCCAAAACGGACGCTGCTTCCCAGGGATTTGTC GTCGCCCTTATTACTGGATTGGGGAGTGTGGCAAT
A. arundinaceus (3)	GAAGTGTTTTCTAGGCTAGATAATTCCTGTTTGATCCAAAACGGACTCTGCTTCCCAGGGATTTGTC GTCGCCCTTATTACTGGATTGGGGACTGTAGCAAT
A. taiti	GAAGTGTTTTCTAGGCTAGATAATTCCTGTTTGATCCAAAACGGACGCTGCTTCCCAGGGATTTGTC GTCGCCCTTATTACTGGATTGGGGAGTGTAGCAAT
A. scirpaceus (1)	GAAGTGTTTTCTAGGCTAGATAATTCCTGTTTGATCCAAAACGGACTCTGCTTCCCAGGGATTTGTC GTCGCCCTTATTACTGGATTGGGGAGTGTAGCAAT
A. scirpaceus (2)	GAAGTGTTTTCTAGGCTAGATAATTCCTGTTTGATCCAAAACGGATTCTGCTTCCCAGGGATTTGTC GTCGCCCTTATTACTGGATTGGGGAGTGTAGCAAT
AvBD8	
A. sechellensis (1)	TGCAGACAGGCTGGAGGGGTCTGCTCCAGCGACCGCTGCCTCCTACGCCACATGAGACCCTTTGGA CGCTGCCAGCCGGGAATTCCCTGTTGTAGGACC
A. sechellensis (2)	TGCAGACAGGCTGGAGGGGTCTGCTCCAGCGACCGCTGCCTCCTACGCCACATGAGACCCTTTGGA CGCTGCCAGCCGGGAATTCCCTGTTGTAGGACC
A. sechellensis (3)	TGCAGACAGGCTGGGGGGGTCTGCTCCAGCGACCGCTGCCTCCTACGCCACATGAGACCCTTTGGA CGCTGCCAGCCGGGAATTCCCTGCTGTAGGACC
A. brevipennis (1)	TGCAGACATGCTGGGGGGGTCTGCTCCAGCGACCGCTGCCTCCTACGCCACATGAGACCCTTTGGA CGCTGCCAGCCAGGAATTCCCTGCTGTAGGACC
A. brevipennis (2)	TGCAGACATGCTGGGGGGGTCTGCTCCAGCGACCGCTGCCTCCTACGCCACATGAGACCCTTTGGA CGCTGCCAGCCAGGAATTCCCTGCTGTAGGACC
A. arundinaceus (1)	TGCAGACAGGCTGGGGGGGTCTGCTCCAGCGACCGCTGCCTCCTACGCCACATGAGACCCTTTGGA CGCTGCCAGCCGGGAATTCCCTGCTGTAGGACC
A. arundinaceus (2)	TGCAGACAGGCTGGGGGGGTCTGCTCCAGCGACCGCTGCCTCCTACGCCACATGAGACCCTTTGGA TGCTGCCAGCCGGGAATTCCCTGCTGTAGGACC
A. taiti	TGCAGACAGGCTGGGGGGGTCTGCTCCAGCGACCGCTGCCTCCTACGCCACATGAGACCCTTTGGA CGCTGCCAGCCGGGAATTCCCTGCTGTAGGACC
A. scirpaceus (1)	TGCAGACAGGCTGGGGGGGTCTGCTCCAGCGACCTCTGCCTCCTACGCCACATGAGACCCTTTGGA CGCTGCCAGCCAGGAATTCCCTGCTGTAGGACC
A. scirpaceus (2)	TGCAGACAGGCTGGGGGGGTCTGCTCCAGCGACCTCTGCCTCCTACGCCACATGAGACCCTTTGGA CGCTGCCAGCCGGGAATTCCCTGCTGTAGGACC
A. scirpaceus (3)	TGCAGACAGGCTGGGGGGGTCTGCTCCAGCGACCGCTGCCTCCTGCGCCACATGAGACCCTTTGG ACGCTGCCAGCCGGGAATTCCCTGCTGTAGGACC

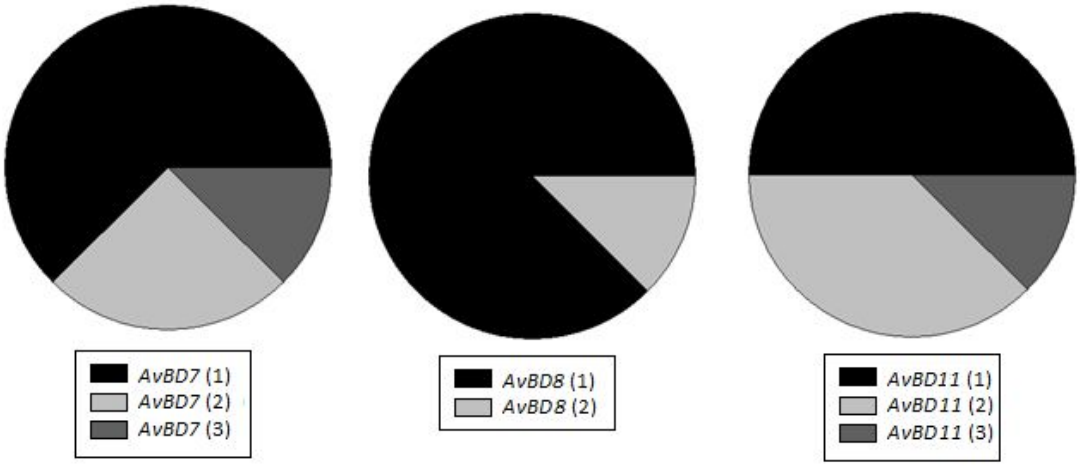
A. scirpaceus (4)	TGCAGACAGGCTGGGGGGGTCTGCTCCAGTGACCGCTGCCTCCTACGCCACATGAGACCCTTTGGA CGCTGCCAGCCGGGAATTCCCTGCTGTAGGACC
A. scirpaceus (5)	TGCAGACAGGCTGGGGGGGTCTGCTCCAGCGACCCCTGCCTCCTTCGCCACATGAGACCCTTTGGA CGCTGCCAGCCGGGAATTCCCTGCTGTAGGACC
A. scirpaceus (6)	TGCAGACAGGCTGGGGGGGTCTGCTCCAGCGACCTCTGCCTCCTTCGCCACATGAGACCCTTTGGA CGCTGCCAGCCGGGAATTCCCTGCTGTAGGACC
AvBD9	
A. sechellensis	TCCTGCTCCTTCATGCCCTGCTCTGCTCCTCTGGTTGACATCGGGACCTGCCGCGGTGGGAAGCTA
A. brevipennis	TCCTGCTCCTTCGTGCCCTGCTCTGCTCCTCTGGTTGACATCGGGACCTGCCGCGGTGGGAAGCTA
A. arundinaceus	TCCTGCTCCTTCGTGCCCTGCTCTGCTCCTCTGGTTGACATCGGGACCTGCCGCGGTGGGAAGCTA
A. taiti	TCCTGCTCCTTCGTGCCCTGCTCTGCTCCTCTGGTTGACATCGGGACCTGCCGCGGTGGGAAGCTA
A. scirpaceus	TCCTGCTCCTTCGTGCCCTGCTCTGCTCCTCTGGTTGACATCGGGACCTGCCGCGGTGGGAAGCTA
AvBD11	
A. sechellensis (1)	AGGGACACCTTGCGTTGCTTGGAATACCACGGCTACTGCTTCCATCTGAAATCCTGCCCGGAGCCAT TTGCTGCCTTTGGAACTTGCTATCGGCGCCGGAGGACCTGCTGTGTTGGT
A. sechellensis (2)	AGGGACACCTTGCGTTGCTTGGAATACCACGGCTACTGCTTCCATATGAAATCCTGCCCGGAGCCA TTTGCTGCCTTTGGAACTTGCTATCGGCGCCGGAGGACCTGCTGTGTTGGT
A. sechellensis (3)	AGGGACACCTTGCGTTGCTTGGAATACCACGGCTACTGCTTCCATCTGAAATCCTGCCCGGAGCCAT TTGCTGCCTTTGGAACTTGCTATCGGCGCCGGAGGACCTGCTGTGTTGGT
A. arundinaceus (1)	AGGGACACCTTGAGTTGCTTGGAATACCACGGCTACTGCTTCCATCTGAAATCCTGCCCGGAGCCA TTTGCTGCCTTTGGAACTTGCTATCGGCGCCGGAGGACCTGCTGTGTTGGT
A. arundinaceus (2)	AGGGACACCTTGCGTTGCTTGGAATACCACGGCTACTGCTTCCATCTGAAATCCTGCCCGGAGCCAT TTGCTGCCTTTGGAACTTGCTATCGGCGCCGGAGGACCTGCTGTGTTGGT
A. arundinaceus (3)	AGGGACACCTTGCGTTGCTTGGAATACCATGGCTACTGCTTCCATCTGAAATCCTGCCCGGAGCCAT TTGCTGCCTTTGGAACTTGCTATCGGCGCCGGAGGACCTGCTGTGTTGGT
A. taiti	AGGGACACCTTGCGTTGCTTGGAATACCACGGCTACTGCTTCCATCTGAAATCCTGCCCGGAGCCAT TTGCTGCCTTTGGAACTTGCTATCGGCGCCGGAGGACCTGCTGTGTTGGT
A. scirpaceus (1)	AGGGACACCTTGCGTTGCTTGGACTACCACGGCTACTGCTTCCATCTGAAATCCTGCCCGGAGCCAT TTGCTGCCTTTGGAACTTGCTATCGGCGCCGGAGGACCTGCTGTGTTGGT
A. scirpaceus (2)	AGGGACACCTTGCGTTGCTTGGACTACCACGGCTACTGCTTCCATCTGAAATCCTGCCCGGAGCCAT TTGCTGCCTTTGGAACTTGCTATCGGCGCCGGAGGACCTGCTGCGTTGGT
AvBD13	
A. sechellensis	CAGAAGCACCGTGGGCACTGCCGGAGGCTCTGCTTCCACATGGAGCGCTGGGAAGGGAGCTGCA GCAACGGCCGCCTG
A. brevipennis	CGTGGGCACTGCCGGAGGCTCTGCTTCCACATGGAGCGCTGGGAAGGGAGCTGCAGCAACGGCC GCCTG
A. arundinaceus (1)	CGTGGGCACTGCCGGAGGCTCTGCTTCCACATGGAGCGCTGGGAAGGGAGCTGCAGCAACGGCC GCCTG
A. arundinaceus (2)	CGTGGGCACTGCCGGAGGCTCTGCTTCCACATGGAGCGCTGGGAAGGGAGCTGCAGCAACGGCC GCCTG
A. taiti	CGTGGGCACTGCCGGAGGCTCTGCTTCCACATGGAGCGCTGGGAAGGGAGCTGCAGCAACGGCC GCCTG
A. scirpaceus (1)	CGTGGGCACTGCCGGAGGCTCTGCTTCCACATGGAGCGCTGGCAAGGGAGCTGCAGCAACGGCCG CCTG
A. scirpaceus (2)	CGTGGGCACTGCCGGAGGCTCTGCTTCCACATGGAGCGCTGGGAAGGGAGCTGCAGCAACGGCC GCCTG
A. scirpaceus (3)	CGTGGGCACTGCCGGAGGCTCTGCTTCCACATGGAGCGCTGGCAAGGGAGCTGCAGCAATGGCCG CCTG

Figure 1. Allele frequencies at each polymorphic AvBD locus screened in the five *Acrocephalus* species.

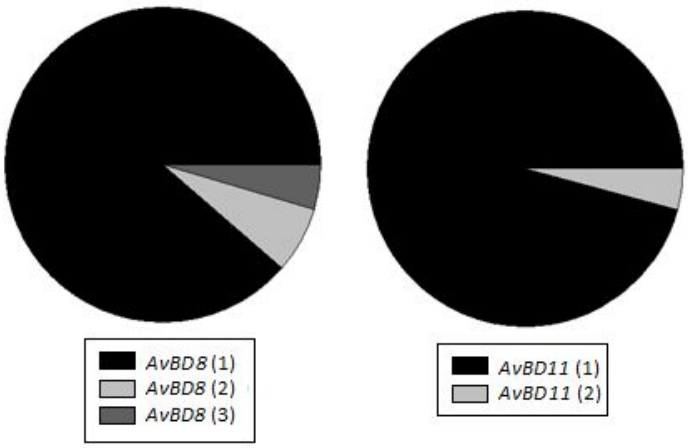
Acrocephalus scirpaceus



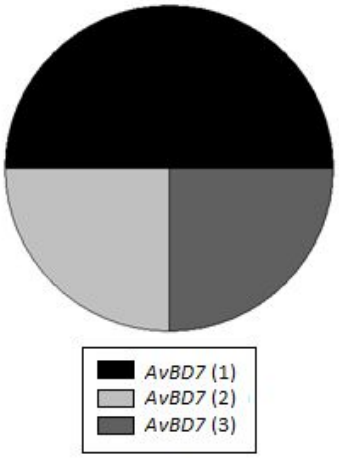
Acrocephalus arundinaceus



Acrocephalus sechellensis



Acrocephalus brevipennis



Chapter 3: Characterisation of Toll-like receptors (TLRs) in the Seychelles warbler

Abstract

In small populations loss of genetic variation due to drift can lead to inbreeding depression and a loss of adaptive potential, thus reducing short- and long-term viability. Under certain circumstances, balancing selection may be able to counteract drift and maintain variation at key loci. Characterising functional loci at which variation remains in small or in bottlenecked populations is important in potentially identifying loci critical to the persistence of these populations and providing candidates for investigations into the relative strength of different evolutionary forces in such situations. Toll-like receptor genes (TLRs) encode for molecules which play a pivotal role in innate immune defence in vertebrates. Here we characterise variation at TLR loci in the Seychelles warbler *Acrocephalus sechellensis*, an endangered passerine that recently went through a severe population bottleneck, resulting in a significant reduction in its genome-wide variation. In this species, we found that five of the seven TLR loci amplified were polymorphic, with one locus (*TLR15*) containing four functional variants. Haplotype-level tests of selection failed to detect selection at these TLRs, but site-specific tests detected signatures of negative (purifying) selection within two loci (*TLR1LB* and *TLR5*) and sites under positive (balancing) selection at *TLR3* and *TLR15*. We characterised variation at the seven TLR loci in six other *Acrocephalus* species with varying demographic backgrounds and identified more sites showing signatures of selection across the genus. Finally, we found a positive correlation between population size and TLR variation across *Acrocephalus* populations, indicating that the species existing in small isolated populations had reduced variation at TLR loci. Our results show that TLR variation does still exist within the Seychelles warbler, thus providing candidate loci for further investigation into the strength and causes of selection at these loci. However, the depauperate TLR variation observed in this small bottlenecked population suggest that even at important immunogenetic loci like the TLRs, balancing selection may, at best, only attenuate the overriding effects of drift.

Introduction

The analysis of genetic variation within and among populations can provide important insight into the evolutionary and demographic history of a species (Garrigan & Hedrick 2003; Piertney & Webster 2010; Sutton *et al.* 2011). Levels of variation also provide an indication of a population's adaptive potential and viability (Frankham *et al.* 1999). In demographically stable populations, genetic variation will reach a mutation-selection-drift balance given sufficient evolutionary time (Kimura & Ohta 1969). However, balancing selection is said to have occurred when genetic variation at a locus is maintained at a higher level than expected based on the amount of drift affecting the population (Takahata, 1990; Takahata & Nei 1990; Takahata *et al.* 1992). Identifying when and where balancing selection occurs can provide insight into the function and importance of specific loci and help us to understand the evolutionary pressures affecting a population (Oleksiak *et al.* 2002; Mitchell-Olds & Schmitt 2006). Furthermore, understanding the potential for critical genetic variation to be maintained within small, isolated populations where drift is strong (Lacy 1987; Franklin & Frankham 1998; van Oosterhout *et al.* 2006) is important from a conservation perspective (Young *et al.* 1996; Tompkins 2007; Willi *et al.* 2007; Grueber *et al.* 2013).

Pathogen-mediated selection (PMS) has been proposed to be a major driver of balancing selection given the strong co-evolutionary relationship between pathogens and their hosts (Jeffery & Bangham 2000; Bernatchez & Landry 2003). This idea is well-supported by various studies that have identified elevated levels of variation at specific immune genes across a range of taxa (Hoelzel *et al.* 1993; Luikart *et al.* 1998; Frankham *et al.* 1999; Hansson & Richardson 2005). Three main none mutually-exclusive mechanisms of PMS (i.e. heterozygote advantage, rare allele advantage and fluctuating selection) (Doherty & Zinkernagel 1975; Hill *et al.* 1991; Slade & McCallum 1992, respectively) have been put forward explain how genetic variation may be maintained at these immune genes (for reviews, see Potts & Slev 1995; Hedrick 2002; Spurgin & Richardson 2010). Other forces such as sexual selection (Fisher 1915; Andersson 1994) and selection against any mutational load associated with highly polymorphic genes (van Oosterhout 2009) can act on top of this, and have the potential to interact and exacerbate the overall effect on genetic variation (for examples, see Brouwer *et al.* 2010; Netea *et al.* 2012; Ejsmond *et al.* 2014).

Most studies which use a candidate gene approach to investigate balancing selection have focused on Major Histocompatability Complex (MHC) genes (for reviews, see Piertney & Oliver 2006; Spurgin & Richardson 2010), both because of their central role in the acquired immune response and because of the exceptional levels of polymorphism observed (Hedrick 1994; Meyer & Thomson 2001). However, population genetic inference of selection is difficult in this multigene family because of the complications caused by various phenomena including frequent gene duplication (Jeffery & Bangham 2000; Hess & Edwards 2002), gene conversion (Ohta 1995; Spurgin *et al.* 2011), epistasis, strong linkage and high mutational load (van Oosterhout 2009). In contrast, studies on variation at innate immune system genes within wild populations are relatively scarce (for review, see Sutton *et al.* 2011), yet these genes are thought to have a relatively simple genomic architecture and evolution, which reduces the confounding effects of the other factors outlined above. Furthermore, innate immune genes play a pivotal role as the first line of defence in vertebrate immunity and there is evidence that they can be under balancing selection (Schlenke & Begun 2003; Ferrer-admetlla *et al.* 2008; Mukherjee *et al.* 2009).

Toll-like receptors (TLRs) are membrane-bound sensors of the innate immune system that recognise distinctive molecular features of invading microbes (for review, see Jin & Lee 2008). They bind pathogen-associated molecular patterns (PAMPs), thus triggering an intracellular signal cascade to activate an appropriate immune response (Takeda & Akira 2005). TLRs are divided into six families based on the types of PAMPs they bind to (Roach *et al.* 2005). These include TLRs which bind to bacterial lipoproteins, lipopolysaccharides or DNA motifs (Takeuchi *et al.* 2002; Bihl *et al.* 2003; Keestra *et al.* 2010). In vertebrates, TLRs link the innate and adaptive immune system, working with both modes of immune defence (Schnare *et al.* 2001; Roach *et al.* 2005). Recent studies show that polymorphisms at TLR loci can have a direct effect on resistance/ susceptibility to pathogen infection across a range of vertebrate groups (see: Creagh & O'Neill 2006; Vinkler *et al.* 2009; Franklin *et al.* 2011). Consequently, PMS is thought to maintain variation at these genes and positive selection at TLR genes has been shown in fish (Palti 2011), mammals (Nakajima *et al.* 2008; Areal *et al.* 2011; Tschirren *et al.* 2013) and birds (Downing *et al.* 2010; Alcaide & Edwards 2011; Grueber *et al.* 2013, 2014).

Wild birds have been the focus of a disproportionate number of evolutionary and ecological studies (for review, see Kaiser 2007). The samples and data from such studies now provide excellent systems in which to investigate the causes and consequences of innate immune gene variation under natural conditions. A recent study of variation at avian TLR genes across outbred passerines found evidence that balancing selection was responsible for maintaining variation at these loci (Alcaide & Edwards 2011). Another study on a bottlenecked population of a single species showed that TLR variation was elevated compared to overall genetic diversity (Grueber *et al.* 2013).

Here we characterise variation at seven TLR genes in the Seychelles warbler *Acrocephalus sechellensis* (SW) and, for comparison, across six other *Acrocephalus* warbler species (OW). We use these data to investigate whether TLR variation exist despite the severe bottleneck that the SW population endured when its population was reduced to ca 29 individuals in the last century (Collar & Stuart 1985). We assess whether there is any evidence that selection has influenced TLR variation in the SW, or across the *Acrocephalus* genus, and include a comparison of TLR variation in relation to population size across all of the *Acrocephalus* populations characterised.

Materials and Methods

Study species and sampling

The Seychelles warbler (SW) is a small (ca 12-15 g) insectivorous passerine bird endemic to the Seychelles islands (Safford & Hawkins 2013). Due to anthropogenic effects, by the 1960s the SW was reduced to just one population of ca 26 individuals remaining on the island of Cousin (Collar & Stuart 1985). As a result, the SW effective population size was dramatically reduced from ca 6900 in the early 1800s to < 50 in the contemporary population (Spurgin *et al.* 2014). However, with effective conservation management, the population recovered to its carrying capacity of ca 320 adults on Cousin by 1982 (Komdeur 1992) and has since remained relatively stable (Brouwer *et al.* 2009; Wright *et al.* 2014). The SW has since proved to be an excellent study species for evolutionary, ecological and conservation question (Komdeur 1992; Richardson *et al.* 2003; van de Crommenacker *et al.* 2011; Barrett *et al.* 2013). Since 1997, > 96% of the Cousin population have been caught and each bird

rung with a unique combination of colour rings and a metal British Trust for Ornithology ring (Richardson *et al.* 2002). Birds are aged at first catch according to their eye-colour and behaviour: adult birds are > 10 months old with distinctive reddish-brown eyes compared to the light brown eyes of a sub-adult aged 5-10 months. Birds < 5 months old have grey eyes. Blood samples (ca 25 µl) are taken via brachial venipuncture, placed in absolute ethanol in a 2 ml screw-top Eppendorf tube and stored at 4^{o}C.

Toll-like receptor (TLR) variation in the Seychelles warbler (SW)

Samples were from unrelated adult birds (> 1 year old) chosen at random from the contemporary 2000-2008 population. Genomic DNA was extracted using a salt-extraction method (Richardson *et al.* 2001), and sex-confirmed using a molecular protocol (Griffiths *et al.* 1998). The TLR loci were selected based on their successful amplification in other passerine species - principally, the house finch *Carpodacus mexicanus,* and New Zealand robin *Petroica australis raikura* - using locus-specific primers (Alcaide & Edwards 2011; Grueber & Jamieson 2013) (Table S1). The seven TLR genes that amplified successfully (*TLR1LA, TLR1LB, TLR3, TLR4, TLR5, TLR15* and *TLR21*) were screened in 22-33 Seychelles warblers. The number of samples required to identify all variation at each locus was calculated using rarefaction curves of the number of alleles discovered with increasing sample size until the curve reached an asymptote, using HPRare v1.0 (Kalinowski 2005).

For each locus, PCRs were carried out in 10 µl volume with genomic DNA at a concentration of ca 10 ng / µl. Taq PCR Master Mix was used (Qiagen, UK) which includes: *Taq* DNA Polymerase, QIAGEN PCR Buffer, $MgCl_2$, and ultrapure dNTPs at optimised concentrations. PCRs were carried out using the following conditions: 40 s at 94^{o}C, 40 s at the locus-specific annealing temperature (Table S1), 80 s at 72^{o}C, all repeated for 34 cycles. All PCRs started with an incubation step of 3 mins at 94^{o}C and finished with an incubation step of 10 mins at 72^{o}C. All PCR products were electrophoresed on a 2% agarose gel containing ethidium bromide and visualised to determine successful amplification of the expected size fragment. Successful samples were submitted to MWG Operon (Eurofins, Germany) for Sanger sequencing. All unique sequences were confirmed by repeated sequencing across multiple individuals or, where identified in only one individual, multiple independent PCRs from that individual.

All sequences were aligned against target sequences of the given loci/exon from other passerine species available in the National Centre for Biotechnology Information (NCBI) nucleotide database using BioEdit (Hall 1999) via ClustalW codon alignment. Each chromatogram was examined by eye to identify single nucleotide polymorphisms (SNPs). Sequences with multiple SNPs had their haplotypes inferred using Bayesian PHASE algorithms (Stephens & Donnelly 2003) in the program DnaSP (Librado & Rozas 2009).

Toll-like receptor (TLR) variation across *Acrocephalus* species (OW)

To help identify signatures of selection within TLR loci and assess variation at the genus-level (*Acrocephalus*), the same TLR loci as above were screened in 4-8 individuals from each of the great reed warbler (*A. arundinaceus*), Eurasian reed warbler (*A. scirpaceus*), Australian reed warbler (*A. australis*), sedge warbler (*A. schoenobaenus*), Cape Verde warbler (*A. brevipennis*) and Henderson's Island warbler (*A. taiti*); hereafter collectively referred to as 'other warblers' (OW). The sequencing protocols outlined above for the SW were used but with different optimised annealing temperatures for each OW species (Table S1).

These *Acrocephalus* species have a range of demographic and evolutionary histories (for overview, see Schulze-Hagen & Leisler 2011). The great reed warbler, Eurasian reed warbler, Australian reed warbler and Sedge warbler are all migrant species from large outbred populations classified as 'under least concern'. Estimated European populations are 950'000, 3.1 million and 2.3 million for the great reed warbler, Eurasian reed warbler and sedge warbler respectively (after Hagemeijer & Blair 1997; BirdLife International 2015). The Australian Reed warbler is widespread across Australia, New Guinea and South-West Asia, therefore we have used an estimate of 1.5 million based on existing literature (del Hoyo *et al.* 2006; BirdLife International 2015). These four species can be categorised as 'mainland'. The Cape Verde warbler and Henderson's Island warbler are two other island species with restricted but- at the present time- stable populations estimated at 1000-1500 (Schulze-Hagen & Leisler 2011) and ca 7000 individuals (Brooke & Hartley 1995; Birdlife International 2015) respectively. These two species and the SW are categorised as 'island' species.

We tested for differences in variation between mainland and island species, in addition to investigating the relationship between consensus population size and TLR haplotype diversity using a logistical regression analysis. This was run for (i) all TLR variation observed and (ii) only TLR variation resulting in a change at the amino acid level (hereafter termed functional variation). Haplotype diversity is a measure of the uniqueness of a given haplotype in a given subset / population of individuals where its formula includes a measure of the relative haplotype frequency (x_i) in the sample of individuals and can account for differences in sample size (N) (Nei 1987).

In order to assess sequence evolution at each TLR locus within and across the different species, we constructed maximum-likelihood trees for each locus with 1000 bootstrap replications. The trees were based on nucleotide variation (given the sequences were all < 1 kb) under the general time reversible substitution model (Nei & Kumar 2000). Sequences of non-*Acrocephalus* avian species, obtained from the NCBI database, were used to root the tree: *Carpodacus mexicanus* (house finch), *Petroica australis raikura* (Stewart Island robin), *Taeniopygia guttata* (zebrafinch), *Picoides pubescens* (downy woodpecker), *Philesturnus carunculatus* (saddleback), *Accipiter cooperi* (Cooper's hawk), *Falco naumanni* (lesser kestrel), *Anas platyrynchus* (mallard) and *Gallus gallus domesticus* (domestic chicken) (Table S2).

Signatures of selection

***Haplotype level tests*:** Amino acid sequences were translated using Mega v5.1 (Tamura *et al.* 2011). In the SW, all haplotype frequencies observed at each loci were tested for linkage disequilibrium using pairwise log likelihood ratio statistics, and were tested for deviation from Hardy-Weinberg proportions using the Markov chain method available in Genepop v.2 (Raymond & Rousset 1995). Tests were based on allelic frequency measures of exact probability and should there be a deviation, subsequent testing for heterozygote deficiency / excess were also carried out (Guo & Thompson 1992). F_{IS} values are presented using Robertson and Hill's estimates (1984), which have lower variance under the null hypothesis compared to the alternative Weir and Cockerham's estimate (1984). DnaSP was used to calculate basic measures of genetic variation for each locus for each species: number of sequences (N), overall number of segregating sites (S), number of unique haplotypes (H),

haplotype diversity (Hd), nucleotide diversity (π) and ratio of synonymous (dS) to non-synonymous (dN) substitutions.

Neutrality tests were carried out on the SW sequences using DnaSP, including Tajima's D (Tajima 1989), Fu and Li's D (Fu & Li 1993) and the D-statistic (Fu 1996). Tajima's D is based on the differences between the number of segregating sites and the average number of nucleotide differences. Fu and Li's D statistic is based on the differences between mutations appearing only once among sequences and total number of mutations, whereas the F-statistic is based on η_s and the average number of nucleotide differences between pairs (k). These tests of selection are averaged over all sites in the sequence, thus they will be confounded if selection differs across sites. Moreover they lack power and are not able to detect relatively weak signatures of selection (Pond & Frost, 2005).

Z-tests of selection were carried out in Mega v5.1 (Tamura *et al.* 2007) in order to identify selection based on dN/dS across species (Kryazhimskiy & Plotkin 2008); first using the OW species and then also including the SW. This is a codon-based test that can account for selective waves with different direction or intensity on specific sites (for example, see Burgarella *et al.* 2012). The McDonald Kreitman test (MK) was also carried out for each locus comparing the ratio of dN to dS mutations both between and within pairwise species (McDonald & Kreitman 1991).

Using the haplotype sequences for all *Acrocephalus* species, the occurrence of gene conversion was estimated for each locus in DnaSP, which incorporates an algorithm (Betrán *et al.* 1997) to detect gene conversion tracts from multiple differentiated populations (referred to as subpopulations). Recombination rates were also estimated using the recombination parameter R = 4Nr (for autosomal loci of diploid organisms) (Hudson 1987) where N is the population size and r is the recombination rate per sequence (per gene). The estimator is based on the variance of the average number of nucleotide differences between pairs of sequences, S2k (Hudson 1987, equation 1). The minimum number of recombination events is estimated based on these calculations (Hudson & Kaplan 1985).

Site specific tests: We assessed evidence of selection at codons within each TLR locus across the *Acrocephalus* genus using phylogenetically-controlled selection tests. The HyPhy package available on DataMonkey (Delport *et al.* 2010) was used to run different models

(for review, see Kosakovsky Pond & Frost 2005) to identify individual sites under selection based on dN/dS ratios at each codon across: (i) SW (ii) OW and (iii) SW & OW. Two models were run: (i) MEME, a mixed effects model of evolution with a significance level threshold of 0.1 and used to detect episodic positive selection (Murrell *et al.* 2012)and (ii) FUBAR, a fast unconstrained Bayesian approximation model using a Markov chain Monte Carlo routine which has a Bayes Factor / posterior probability set at 0.9 and detects sites under putative selection (Murrell *et al.* 2013).

Results

Table 1 characterises the variation observed at the seven loci amplified in the SW. *TLR4* and *TLR21* were monomorphic in the 30 individuals screened for these loci. *TLR1LA, TLR1LB*, *TLR3, TLR5* and *TLR15* showed polymorphisms (Fig S1). *TLR15* was the only locus where all the variation observed (at three segregating sites) was non-synonymous, resulting in four different amino acid haplotypes. Table 2 characterises the variation observed at the same seven TLR loci in the other *Acrocephalus* species populations (OW). There was significantly more variation present at loci in the mainland species- A*. australis, A. arundinaceus, A. schoenobaenus* and *A. scirpaceus*- than observed in the island endemic species including *A. brevipennis, A. sechellensis* and *A. taiti* (Fig 1). This was the case for the number of segregating sites *S* (t = -2.75, df = 6, *P* = 0.032) and number of unique haplotypes *H* (t = -2.99, df = 6, *P* = 0.023). Focusing on the species, post-hoc Tukey tests show that levels of variation averaged across all TLR loci (measured as *S* and *H*) observed in the SW differ to those observed in *A. brevipennis* (Tukey HSD: mean difference in S = -0.268, *P* = 0.076; mean difference in H = -0.308, *P* = 0.074) but not in *A. taiti* (Tukey HSD: mean difference in S = -0.052, *P* = 0.891; mean difference in H = -0.148, *P* = 0.496).

Across the *Acrocephalus* species sampled, census population size significantly predicted mean haplotype diversity for all nucleotide variation (t = 3.20, df = 6, *P* = 0.02) (Fig 2a) though this pattern became a non-significant trend when only considering amino acid variation i.e. dN substitutions only (t = 1.99, df = 6, *P* = 0.10) (Fig 2b) most likely due to a lack of power. Overall, the mainland species had more variation across the TLR gene family in comparison to the island endemic species with the mean number of alleles observed per

locus for mainland *Acrocephalus* species 3 - 4, but only 2 - 3 for the island *Acrocephalus* species.

Table 1. Characterising variation at seven Toll-like receptor (TLR) loci in the Seychelles warbler (SW). Abbreviations: number of segregating sites (S), number of haplotypes (H), haplotype diversity (Hd) with standard deviations (sd), nucleotide diversity (π) with sd, number of non-synonymous polymorphisms (dN) and number of synonymous polymorphisms (dS).

Locus	N	Fragment size	S	H	Hd (sd)	π (SD)	dN	ds
TLR1LA	44	531	1	2	0.36 (0.07)	0.0007 (0.0001)	0	1
TLR1LB	66	750	2	4	0.64 (0.04)	0.0011 (0.0001)	0	2
TLR3	56	801	3	5	0.54 (0.06)	0.0012 (0.0001)	2	1
TLR4	60	648	0	1	-	-	-	-
TLR5	46	741	2	3	0.13 (0.07)	0.0003 (0.0002)	1	1
TLR15	60	528	3	4	0.69 (0.02)	0.0017 (0.0001)	3	0
TLR21	60	462	0	1	-	-	-	-

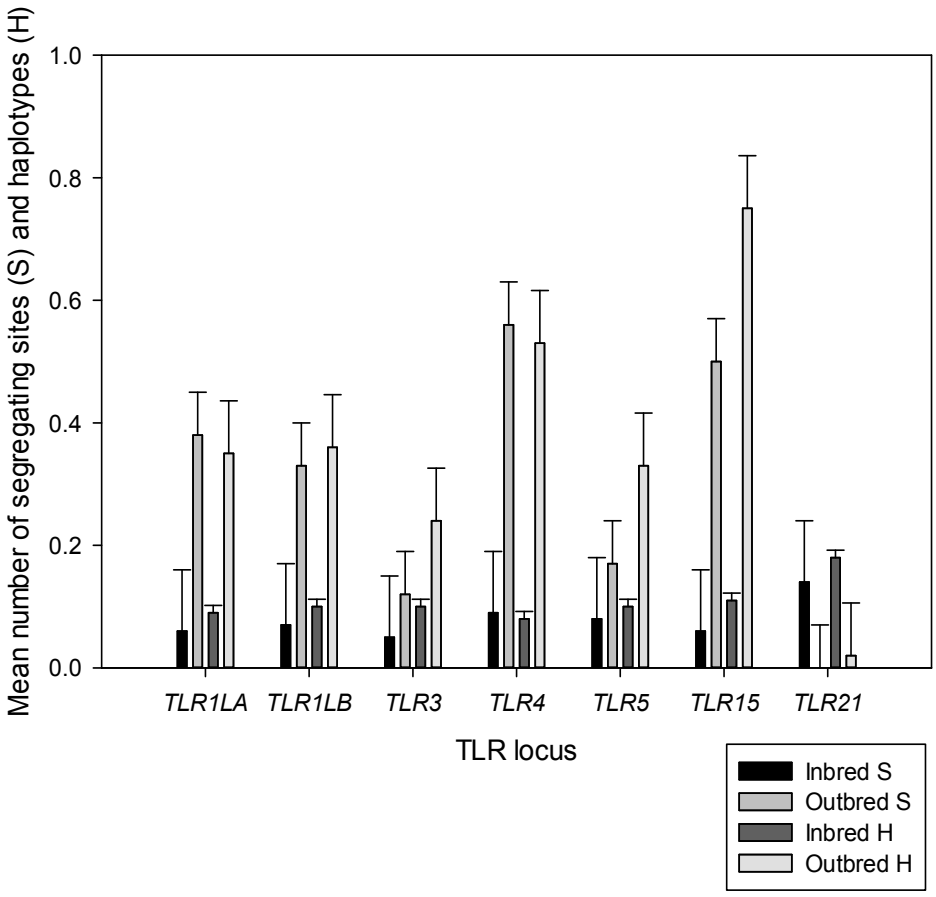


Figure 1. Levels of variation observed across Toll-like receptor (TLR) loci in island (*A. brevipennis, A. sechellensis* and *A. taiti*) compared to mainland *Acrocephalus* species (*A. arundinaceus, A. australis. A. scirpaceus* and *A. schoenobaenus*): i) number of segregating sites (S), ii) number of unique haplotypes (H), iii) haplotype diversity (Hd) and iv) nucleotide diversity (Pi).

Table 2. Characterising variation at seven Toll-like receptor (TLR) loci in a range of other *Acrocephalus* species: *A. arundinaceus, A. australis, A. brevipennis, A. scirpaceus, A. schoenobaenus* and *A. taiti*. Abbreviations include: number of segregating sites (S), number of haplotypes (H), haplotype diversity (Hd) with standard deviations (sd), nucleotide diversity (π) with sd, number of synonymous polymorphisms (dS) and number of non-synonymous polymorphisms (dN).

Locus	Species	N	Fragment size*	S	H	Hd (sd)	π (SD)	dN	dS
TLR1LA	*Acbr*	10	918	3	3	0.38 (0.181)	0.0007 (0.00036)	3	0
	Acar	10	756	4	2	0.36 (0.159)	0.0019 (0.00084)	3	1
	Acta	12	915	0	1	0.00 (0.000)	0.0000 (0.00000)	0	0
	Acau	16	843	6	7	0.82 (0.005)	0.0017 (0.00024)	5	1
TLR1LB	*Acbr*	10	948	3	3	0.51 (0.164)	0.0008 (0.00034)	2	1
	Acar	8	558	6	3	0.71 (0.123)	0.0051 (0.00088)	3	3
	Acta	12	951	1	2	0.49 (0.106)	0.0005 (0.00011)	1	0
	Acsc	10	780	2	3	0.69 (0.104)	0.0011 (0.00023)	1	1
	Acsch	8	792	1	2	0.49 (0.169)	0.0005 (0.00021)	1	0
	Acau	16	954	5	7	0.74 (0.105)	0.0012 (0.00026)	5	0
TLR3	*Acbr*	10	942	0	1	0.00 (0.000)	0.0000 (0.00000)	0	0
	Acar	10	642	2	3	0.64 (0.101)	0.0011 (0.00026)	0	2
	Acta	12	777	1	2	0.30 (0.147)	0.0004 (0.00019)	1	0
	Acsc	8	720	1	2	0.25 (0.180)	0.0004 (0.00025)	0	1
	Acsch	4	836	0	1	0.00 (0.000)	0.0000 (0.00000)	0	0
	Acau	12	888	1	2	0.30 (0.147)	0.0003 (0.00017)	1	0
TLR4	*Acbr*	8	659	6	3	0.46 (0.200)	0.0034 (0.00143)	0	6
	Acar	4	660	1	2	0.50 (0.265)	0.0008 (0.00040)	0	1
	Acta	12	672	1	2	0.17 (0.134)	0.0003 (0.00020)	1	0
	Acsc	10	655	11	9	0.98 (0.054)	0.0059 (0.00076)	6	5
	Acsch	6	617	6	5	0.93 (0.122)	0.0040 (0.00123)	2	4
	Acau	16	660	2	3	0.51 (0.126)	0.0009 (0.00024)	2	0
TLR5	*Acbr*	6	423	0	1	0.00 (0.000)	0.0000 (0.00000)	0	0
	Acar	8	459	1	2	0.43 (0.169)	0.0009 (0.00037)	0	1
	Acta	10	501	3	2	0.36 (0.159)	0.0021 (0.00100)	3	0
	Acsc	4	504	1	2	0.67 (0.204)	0.0013 (0.00041)	0	1
TLR15	*Acbr*	2	892	1	2	1.00 (0.500)	0.0011 (0.00056)	1	0
	Acar	4	807	1	2	0.67 (0.204)	0.0008 (0.00025)	1	0
	Acta	2	519	0	1	0.00 (0.000)	0.0000 (0.00000)	0	0
	Acsc	2	901	1	2	1.00 (0.500)	0.0011 (0.00055)	1	0
	Acsch	2	808	2	2	1.00 (0.500)	0.0025 (0.00124)	1	1
TLR21	*Acar*	8	462	1	2	0.25 (0.180)	0.0006 (0.00046)	1	0
	Acau	14	462	2	2	0.26 (0.136)	0.0011 (0.00059)	1	1

Neighbour-joining trees showed distinct segregation to the level of genus in all the TLR loci. All Acrocephalus species screened were clearly out-grouped from all non-passerine species, including raptors and galliformes, but also from other passerine families like the finches and thrushes (Fig S2). Within the Acrocephalus, variation within each TLR locus did not separate out by individual species but all the branching at this level was supported by

low bootstrapping values. None of the loci showed evidence of gene conversion as no conversion events were identified for any of the pairwise combinations of alleles tested in any of the species for any of the loci. However, at least one recombination event appears to have occurred in each of four TLR genes in the evolution of these Acrocephalus warblers (TLR1LB, TLR3, TLR4 and TLR15) (minimum number of recombination events identified between specific sites = 2, 1, 2, 2 respectively).

Figure 2a.

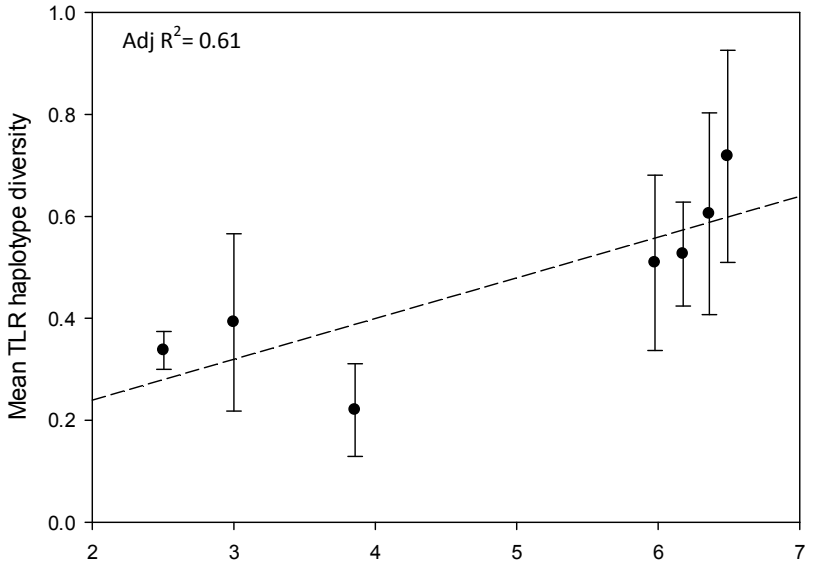


Figure 2b.

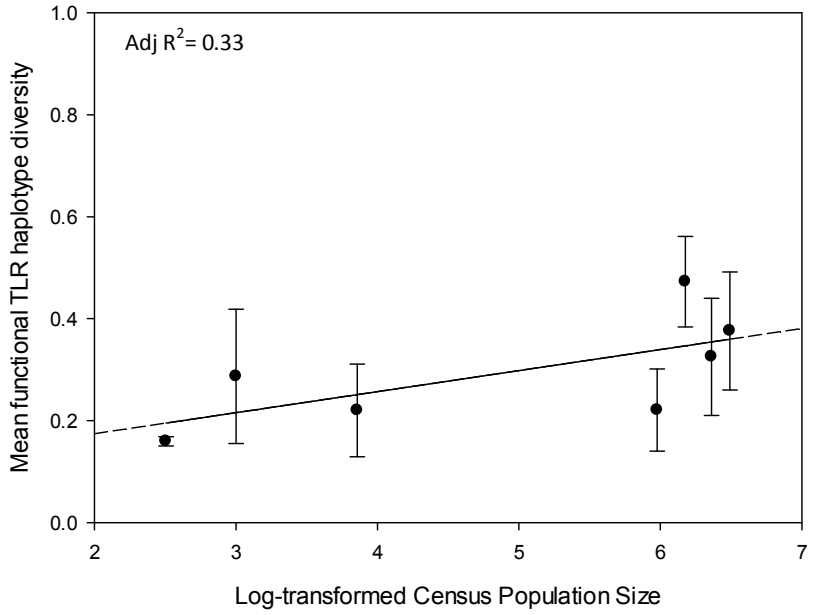


Figure 2. Mean Toll-like receptor (TLR) haplotype diversity (Hd) in relation to census population size across seven *Acrocephalus* species. Regression lines are denoted by dashed lines.

***Haplotype-level selection tests*:** In the SW, three loci deviated from Hardy-Weinberg proportions: *TLR1LB* and *TLR3* had heterozygote excess (*TLR1LB:* F_{IS} = 0.372, *P* = 0.002; *TLR3:* F_{IS} = 0.186, *P* = 0.031) and *TLR15* deviated based on Fisher's exact test of allele frequency probabilities (F_{IS} = -0.061, *P* = 0.017). By plotting observed and expected haplotype frequencies under Hardy-Weinberg equilibrium for *TLR15* (and other loci), it is clear that *TLR15* is the most variable locus and some alleles show signs of heterozygote deficiency whereas others show signs of heterozygote excess (Fig S1). None of the pairwise combinations of loci tested positive for linkage disequilibrium. At the haplotype level, none of the tests could reject neutral evolution when performed on the limited numbers of alleles found at each of the five polymorphic TLR loci within the SW (all tests *P* > 0.1) (Table S2). The z-tests based on dN / dS across whole sequences also failed to detect selection in both the SW and in OW (Table S3). The McDonald-Kreitman test, looking at pairwise dN / dS comparisons between each species, failed to detect any significant signatures of selection with the exception of *TLR4* for the Seychelles warbler and the Australian reed warbler, which had five fixed synonymous substitutions between the two species (*P* = 0.048) (Table S4).

Site-specific selection tests: The codon-based test of selection was performed at three levels: within the Seychelles warblers only (SW), across the other warblers excluding the Seychelles warbler (OW), and finally across the entire dataset (SW + OW). Using the SW TLR sequences, both *TLR1LB* and *TLR5* each had a single site identified as being under putative purifying selection according to the FUBAR model, which also identified a single site at both the *TLR3* and *TLR15* loci under putative positive (balancing) selection (Table 3).

When the OW TLR sequences were examined (excluding the SW) the same site at *TLR1LB* was identified as being under purifying selection along with an additional site, and a total of three other sites were identified to be under putative positive selection in the FUBAR model (Table 3). One of these sites was confirmed using the episodic positive selection MEME model. As for *TLR5,* the one site under purifying selection identified in the SW was not identified in the OW, but re-appeared when considering all *Acrocephalus* species, so must be SW-specific . The one site under positive selection at *TLR3* in the SW was also detected when considering OW, in addition to a couple of sites under purifying selection. Likewise, the site under positive selection at *TLR15* in the SW was repeatedly

identified for OW and OW-including-SW by both MEME and FUBAR models (Table 3). Furthermore, an additional site under positive selection was identified in OW and a total of five sites under purifying selection, making *TLR15* the locus under the most selection according to dN/dS – based signatures.

Table 3. Site-specific dN/dS analysis of TLR loci to identify sites under putative selection using the fast unconstrained Bayesian approximation model (FUBAR) with a) in the Seychelles warbler, b) within and across other *Acrocephalus* species (OW) including *A. arundinaceus, A. australis, A. brevipennis, A. scirpaceus, A. schoenobaenus* and *A. taiti*, and c) combining both the SW with OW. Sites also identified by the mixed effects model of evolution (MEME) under episodic positive selection only, are denoted with *.

Locus	Group	# Positive codons	Mean dN-dS	Mean post prob dN > dS	# Negative codons	Mean dN-dS	Mean post prob dN < dS
TLR1LA	SW	NA	NA	NA	NA	NA	NA
	OW	1	7.07	0.96	4	-5.64	0.92
	ALL	1	6.28	0.95	4	-5.68	0.93
TLR1LB	SW	0	0	0	1	-7.25	0.97
	OW	3*	7.20	0.96	2	-5.51	0.92
	ALL	3*	5.82	0.95	3	-6.02	0.94
TLR3	SW	1	6.33	0.93	0	0	0
	OW	1	3.89	0.91	2	-5.50	0.92
	ALL	1	5.29	0.92	2	-5.06	0.93
TLR4	SW	NA	NA	NA	NA	NA	NA
	OW	0	0	0	1	-3.16	0.96
	ALL	0	0	0	4	-3.51	0.97
TLR5	SW	0	0	0	1	-4.77	0.91
	OW	0	0	0	0	0	0
	ALL	0	0	0	1	-5.25	0.91
TLR15	SW	1	6.38	0.93	0	0	0
	OW	1*	7.43	0.97	2	-5.36	0.91
	ALL	2*	7.33	0.95	5	-4.70	0.90
TLR21	SW	NA	NA	NA	NA	NA	NA
	OW	0	0	0	2	-5.29	0.93
	ALL	0	0	0	2	-5.22	0.93

The same analyses could not be done for *TLR1LA* since < 3 unique haplotype sequences were detected in the SW. However, in the OW including the SW, one site was identified to be under positive selection and four sites under purifying selection at this locus across the genus. At the loci *TLR4* and *TLR21* (monomorphic in the SW) no sites were identified as being under positive selection across the genus, but had a handful of sites were identified to be under purifying selection at each locus (Table 3). All sites detected by MEME under episodic positive selection were also detected by the putative selection FUBAR model.

Discussion

We characterised variation at seven TLR innate immune genes in the bottlenecked population of the Seychelles warbler (SW), and compared this to the variation observed in six other congeneric species. In the SW, five out of seven TLR genes were polymorphic with 2-5 alleles at each locus. This level of variation per locus is higher than that reported in a previous study where seven out of ten neutral (microsatellite) markers were polymorphic with an average of less than three alleles per locus (Hansson & Richardson 2005). Although, in comparison to the other *Acrocephalus* species screened in the present study, the SW variation was relatively low and hence our haplotype-level neutrality tests failed to detect any signatures of selection. When using haplotype-level tests across all *Acrocephalus* species there was still no strong evidence of selection detected. However, site-specific tests were able to detect individual sites at several TLR loci under both putatively negative (purifying) and positive (balancing) selection in the SW and further sites were identified when looking across the *Acrocephalus* genus. Finally, phylogenetic analyses showed that the different TLR genes evolve independently without the inter-locus complications observed in other immune genes such as the MHC.

Despite the lack of evidence of strong signatures of selection we did find some interesting patterns within specific loci. For example, *TLR15* appeared to have a bias towards potentially functional (amino-acid) variants remaining, with all four sequence variants detected encoding different amino acid sequences. It was the only locus to significantly deviate from Hardy-Weinberg proportions and this may be a result of heterozygote advantage (Figure S1), a mechanism of balancing selection that has been found to act on immune gene variation in various species (Hedrick 2002; Worley *et al.* 2010; Niskanen *et al.* 2013). We suggest this possibility based on the relative excess of heterozygotes at *TLR15* compared to the six other TLR loci examined, though heterozygote excess when tested specifically did not, on its own, explain the deviation from HWE observed. Haplotype-based neutrality tests have been much criticised for their limitations and lack of power for detecting selection (Vasemagi & Primmer 2005; Leffler *et al.* 2012; Li *et al.* 2012), which, given the limited sequence data available from the genetically depauperate population of the SW, may explain our results. Furthermore, haplotype-level tests based on the allele frequency spectrum make strong inferences about the populations'

demography, such as constant population size (Zhai *et al.* 2009). The Seychelles warbler population, which has been expanding rapidly since it was reduced to ca 26 individuals in the 1960's (Wright *et al.* 2014) does not comply with these assumptions.

Selection was identified at individual sites within the exons of the TLR loci examined. In the SW, both *TLR3* and *TLR15* had individual sites identified as being under positive selection, and these sites were confirmed to be under selection across the *Acrocephalus* genus. An additional site at the *TLR15* locus was also identified in the other warbler species but not in the SW, which is evidence of species-specific selection. *TLR15* was also the locus under the greatest amount of selection overall with sites for both positive and purifying selection. This pattern of different sites within the exon showing signatures of different types of selection is probably because some of the sites are directly involved in PAMP binding while others may be important in determining the overall shape and configuration on the molecule and thus conserved (Bell *et al.* 2003; Werling *et al.* 2009; Kawai & Akira 2010). Given that the codons in the same exon interference selection has the potential to mask opposing selection forces (Good *et al.* 2013).

While such codon-based tests across species provide considerable power, there are caveats. For example, the signatures they detect will be of past selection caused by a selective pressure that may no longer be acting (Yang & Bielawski 2000). The tests cannot resolve whether the variation observed is *currently* under selection in the contemporary population. Many sites shown to be under negative (purifying) selection across the other *Acrocephalus* species were also found to be under negative selection in the SW. Population genetic theory predicts that while the intensity of (positive) selection on the innate immune genes may fluctuate in space and time, depending on the selective pressures exerted by pathogens, purifying selection is a constantly operating evolutionary force that preserves the functionality of these genes (Kimura & Ohta 1969; Ohta 2002; Mukherjee *et al.* 2009). This may therefore explain the different results we obtained for positively and negatively selected sites in the SW, in that signals of negative selection were clearer because they overwhelm the potentially weaker signals of balancing selection.

When comparing the SW to OW, it is clear that TLR polymorphism in the SW (and other island populations) is reduced compared to that of the large populations of mainland

Acrocephalus species (Figure 1). This implicates genetic drift associated with the small size / bottlenecked history of these isolated island populations as the main force shaping this genetic variation. While census population size did predict overall levels of nucleotide variation at TLR loci within different populations, this association was weaker (not significant) when testing amino acid variation. This difference is because the island species have much lower levels of overall nucleotide variation than mainland species but not so much lower levels of (apparently functional) amino acid variants (Figure 1). This may be because a greater proportion of functional variation, compared to neutral variation, is retained in the bottlenecks populations, perhaps as a result of balancing selection mitigating the effect of drift on these functional variants

One may ask why pathogen-mediated selection has not been maintained more variation at these important immune loci in the SW. Importantly, despite considerable screening efforts, no gastro-intestinal parasites or virus infections and only one blood parasite - a single strain of avian-malaria (GRW1) - have been detected in the SW population (Hutchings 2009). This contrasts markedly with the diversity of pathogens found in most mainland avian populations, but is normal for remote (bottlenecked) island populations (Steadman *et al.* 1990; Coltman *et al.* 1999; Vögeli *et al.* 2011). Since pathogen-mediated balancing selection is thought to be the force maintaining variation at immune genes (Turner *et al.* 2012; Westerdahl *et al.* 2012; Grueber *et al.* 2014), the paucity of pathogens could help explain why drift appears to be the predominant evolutionary force shaping TLR variation in our SW population (Vögeli *et al.* 2011).

On the other hand, a restricted pathogen in the SW fauna may have actively contributed to the loss of immunogenetic variation. For example, the lack of variation observed at *TLR4* in this study is perhaps particularly surprising as this locus has been shown to be involved in the recognition of Protozoan's such as haemosporidian (malaria-like) parasites (Franklin *et al.* 2011; Basu *et al.* 2012), and the only pathogen detected in the SW was a *Haemoproteus* (Hutchings 2009). It is possible that the single *TLR4* allele remaining in the SW population might have offered the best protection (or tolerance) against GRW1. In the absence of multiple strains exerting selection pressures favouring different alleles, selection may have driven this allele to fixation at *TLR4*. So while the most parsimonious explanation for a lack of variation may be genetic drift, we highlight the possibility that PMS

could reach a new equilibrium in small isolated populations in the form of the complete fixation of a single allele. This effect of selection could have important implications for conservation genetics of post-bottlenecked populations with limited pathogens because immunogenetic variation could be lost faster than expected based on drift alone. In support of this idea, several other studies have found that immunogenetic variation eroded faster than neutral variation in island / fragmented populations (for example, see Bollmer *et al.* 2011; Sutton *et al.* 2011; Eimes *et al.* 2011).

While TLR polymorphism may be low in the SW, it may be significant that some variation has been maintained in five of the seven loci given its demographic history. All TLRs recognise specific pathogen-associated molecular patterns (PAMPs) and their different levels of haplotype variation could reflect the biodiversity of the pathogens they protect against in a particular population or species. *TLR1LA* and *TLR1LB* recognise lipoproteins in the cell walls of bacteria, fungi and protozoans (Brownlie & Allan 2011). In the SW, *TLR1LB* was the only TLR gene that showed evidence for purifying selection. *TLR1LA* only had two synonymous alleles, while *TLR5*,which has been shown to recognise the flagella of bacteria in other species (Andersen-Nissen *et al.* 2007) had three apparently functional alleles. *TLR3* and *TLR15* had the highest levels of variation and these were the only loci at which any sites were identified as being positively selected in the SW. *TLR3* is involved in sensing viral RNA (Uematsu & Akira 2008), whereas *TLR15,* a gene unique to birds, appears to be important in the recognition of intracellular parasites, including haemosporidian parasites (Boyd *et al.* 2007). Further investigations into how individual variation at *TLR15* influences resistance or resilience to haemosporidian infection in the SW may be worthwhile.

There has been much debate on the relative roles of genetic drift and selection in shaping functional variation in natural populations and importantly, whether balancing selection can maintain important functional variation even in the face of strong drift (e.g. Alcaide 2010; Sutton *et al.* 2011; Strand *et al.* 2012). There is, however, considerable evidence that drift outweighs selection even at immunologically-important loci where balancing selection would be expected to be most effective (Miller & Lambert 2004; Willi *et al.* 2007; Kuo *et al.* 2009; Grueber *et al.* 2013, 2015). Our data on TLR variation in the SW and other *Acrocephalus* species concurs with this general view of the overriding effect of drift. However potentially functional variation does still exist within the SW at some of the

TLR loci. Thus it is possible that selection may have played a role in maintaining this variation, though more in-depth studies are now required to investigate this possibility. Furthermore, approaches will need to consider how to delineate the relative effects of drift and selection at these candidate loci during the bottleneck. Studies undertaken during bottleneck events, and which identify the cause of selection, will therefore be required before we can fully understand these dynamics.

In summary, the bottleneck suffered by the SW population appears to have reduced the levels of variation observed in TLR genes in this species. However, some potentially functional variation still remains (most noticeably at the *TLR15* locus) possibly as a result of balancing selection. The limited amount of variation detected does, however, undermine our ability to test the significance of such variation. Sequence-based tests of selection have low statistical power and restrictive assumptions. Only studies assessing the impact of genetic variation on individual fitness within the contemporary population and/ or simulating how drift and selection shape variation across bottlenecks will be able to robustly identify the effects of selection. That some TLR loci have maintained variation in spite of the recent bottleneck makes them an ideal candidate for analysing the association between immunological variation and fitness at the individual level in this wild population.

Acknowledgments

We thank Nature Seychelles for facilitating the work on Cousin Island and the Seychelles Bureau of Standards and the Department of Environment gave permission for sampling and fieldwork. We thank a number of collaborators for providing *Acrocephalus* DNA samples: Drs Deborah Dawson, Juan Carlos Illera, Andrew Dixon, Bengt Hansson, Michael Brooke and Ian Hartley.

Data Accession Statement

All sequences used in the study have been published and are available in GenBank (accession numbers KM657646 - KM657768 & KP814140).

References

Alcaide M (2010) On the relative roles of selection and genetic drift in shaping MHC variation. *Molecular Ecology*, **19**, 3842–4.

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

Andersen-Nissen E, Smith KD, Bonneau R, Strong RK, Aderem A (2007) A conserved surface on Toll-like receptor 5 recognizes bacterial flagellin. *Journal of Experimental Medicine*, **204**, 393–403.

Andersson MB (1994) *Sexual selection*. Princeton University Press.

Areal H, Abrantes J, Esteves PJ (2011) Signatures of positive selection in Toll-like receptor (TLR) genes in mammals. *BMC evolutionary biology*, **11**, 368.

Bandelt HJ, Röhl A (1999) Median-joining networks for inferring intraspecific phylogenies. *Molecular Biology and Evolution*, **16**, 37–48.

Barrett ELB, Burke T a, Hammers M, Komdeur J, Richardson DS (2013) Telomere length and dynamics predict mortality in a wild longitudinal study. *Molecular Ecology*, **22**, 249–59.

Basu M, Das T, Ghosh A *et al.* (2012) Gene-gene interaction and functional impact of polymorphisms on innate immune genes in controlling Plasmodium falciparum blood infection level. *PLoS One*, **7**, e46441.

Bell JK, Mullen GED, Leifer C a. *et al.* (2003) Leucine-rich repeats and pathogen recognition in Toll-like receptors. *Trends in Immunology*, **24**, 528–533.

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

Betrán E, Rozas J, Navarro A, Barbadilla A (1997) The estimation of the number and the length distribution of gene conversion tracts from population DNA sequence data. *Genetics*, **146**, 89–99.

Bihl F, Salez L, Beaubier M *et al.* (2003) Overexpression of Toll-like receptor 4 amplifies the host response to lipopolysaccharide and provides a survival advantage in transgenic mice. *Journal of Immunology*, **170**, 6141–50.

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

Boyd A, Philbin VJ, Smith AL (2007) Conserved and distinct aspects of the avian Toll-like receptor (TLR) system: implications for transmission and control of bird-borne zoonoses. *Biochemical Society Transactions*, **35**, 1504–1507.

Brooke MDL, Hartley IR (1995) Nesting Henderson Reed-Warblers (Acrocephalus vaughani taiti) Studied by DNA Fingerprinting : Unrelated Coalitions in a Stable Habitat? *The Auk*, **112**, 77–86.

Brouwer L, Barr I, van de Pol M *et al.* (2010) MHC-dependent survival in a wild population: evidence for hidden genetic benefits gained through extra-pair fertilizations. *Molecular Ecology*, **19**, 3444–3455.

Brouwer L, Tinbergen JM, Both C *et al.* (2009) Experimental evidence for density-dependent reproduction in a cooperatively breeding passerine. *Ecology*, **90**, 729–741.

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

Burgarella C, NavascuÉs M, Zabal-Aguirre M *et al.* (2012) Recent population decline and selection shape diversity of taxol-related genes. *Molecular Ecology*, **21**, 3006–3021.

Collar NJ, Stuart SN (1985) *Threatened birds of Africa and related islands.* tional Council for Bird Preservation, Cambridge.

Coltman DW, Pilkington JG, Smith JA, Josephine M (1999) Parasite-Mediated Selection against Inbred Soay Sheep in a Free-Living, Island Population. *Evolution*, **53**, 1259–1267.

Creagh EM, O'Neill L a J (2006) TLRs, NLRs and RLRs: a trinity of pathogen sensors that co-operate in innate immunity. *Trends in Immunology*, **27**, 352–7.

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

Delport W, Poon AFY, Frost SDW, Kosakovsky Pond SL (2010) Datamonkey 2010: a suite of phylogenetic analysis tools for evolutionary biology. *Bioinformatics (Oxford, England)*, **26**, 2455–7.

Dobson AP (1988) Restoring Island Ecosystems: The potential of Parasites to Control Introduced Mammals. Conservation Biology. *Conservation Biology*, **2**, 31–39.

Doherty PC, Zinkernagel RM (1975) Enhanced immunological surveillance in mice heterozygous at the H-2 gene complex. *Nature*, **256**, 50–52.

Downing T, Lloyd AT, O'Farrelly C, Bradley DG (2010) The differential evolutionary dynamics of avian cytokine and TLR gene classes. *Journal of Immunology*, **184**, 6993–7000.

Eimes J a, Bollmer JL, Whittingham L a *et al.* (2011) Rapid loss of MHC class II variation in a bottlenecked population is explained by drift and loss of copy number variation. *Journal of Evolutionary Biology*, **24**, 1847–56.

Ejsmond MJ, Radwan J, Wilson AB, Wilson AB (2014) Sexual selection and the evolutionary dynamics of the major histocompatibility complex. *Proceedings Biological sciences / The Royal Society*, **281**, 20141662.

Ferrer-admetlla A, Bosch E, Sikora M *et al.* (2008) Balancing Selection Is the Main Force Shaping the Evolution of Innate Immunity. *Journal of Immunology*, **181**, 1315–1322.

Fisher R a (1915) The evolution of sexual preference. *The Eugenics review*, **7**, 184–192.

Frankham R, Lees K, Montgomery ME *et al.* (1999) Do population size bottlenecks reduce evolutionary potential ? *Animal Conservation*, **2**, 255–260.

Franklin IR, Frankham R (1998) How large must populations be to retain evolutionary potential? *Animal Conservation*, **1**, 69–70.

Franklin BS, Ishizaka ST, Lamphier M *et al.* (2011) Therapeutical targeting of nucleic acid-sensing Toll-like receptors prevents experimental cerebral malaria. *Proceedings of the National Academy of Sciences USA*, **108**, 3689–3694.

Fu Y-X (1996) New statistical tests of neutrality for DNA samples from a population. *Genetics*, **143**, 557–70.

Fu Y-X, Li W-H (1993) Statistical tests of neutrality of mutations. *Genetics*, **133**, 693–709.

Garrigan D, Hedrick PW (2003) Perspective: detecting adaptive molecular polymorphism: lessons from the MHC. *Evolution*, **57**, 1707–22.

Good BH, Walczak AM, Neher R a, Desai MM (2013) Interference limits resolution of selection pressures from linked neutral diversity. *PLoS Genetics*, **10**, e1004222.

Griffiths R, Double MC, Orr K, Dawson RJG (1998) A DNA test to sex most birds. *Molecular Ecology*, **7**, 1071–1075.

Grueber CE, Jamieson IG (2013) Primers for amplification of innate immunity toll-like receptor loci in threatened birds of the Apterygiformes, Gruiformes, Psittaciformes and Passeriformes. *Conservation Genetics Resources*, **5**, 1043–1047.

Grueber CE, Knafler GJ, King TM *et al.* (2015) Toll-like receptor diversity in 10 threatened bird species : relationship with microsatellite heterozygosity. *Conservation Genetics*, **16**, 595–611.

Grueber CE, Wallis GP, Jamieson IG (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**, 4470–4482.

Grueber CE, Wallis GP, Jamieson IG (2014) Episodic positive selection in the evolution of avian toll-like receptor innate immunity genes. *PloS one*, **9**, e89632.

Guo SW, Thompson EA (1992) Performing the exact test of Hardy-Weinberg proportion for multiple alleles. *Biometrics*, **48**, 361–372.

Hagemeijer WJM, Blair MJ (1997) *The EBCC Atlas of European Breeding Birds.* T & AD Poyser, London.

Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium*, **41**, 95–98.

Hansson B, Richardson DS (2005) Genetic variation in two endangered Acrocephalus species compared to a widespread congener: estimates based on functional and random loci. *Animal Conservation*, **8**, 83–90.

Hedrick PW (1994) Evolutionary genetics of the Major Histocompatibility Complex. *The American Naturalist*, **143**, 945–964.

Hedrick PW (2002) Pathogen resistance and genetic variation at MHC loci. *Evolution*, **56**, 1902–1908.

Hess CM, Edwards S V (2002) The Evolution of the Major Histocompatibility Complex in Birds. *BioScience*, **52**, 423–431.

Hill AVS, Allsopp CEM, Kwiatkowski D *et al.* (1991) Common West African HLA antigens are associated with protection from severe malaria. *Nature*, **352**, 595–560.

Hoelzel a R, Halley J, O'Brien SJ *et al.* (1993) Elephant seal genetic variation and the use of simulation models to investigate historical population bottlenecks. *Journal of Heredity*, **84**, 443–9.

Del Hoyo J, Elliott A, Christie D (2006) *Handbook of the Birds of the World, vol. 11: Old World Flycatchers to Old World Warblers.* Lynx Edicions, Barcelona, Spain.

Hudson RR (1987) Estimating the recombination parameter of a finite population model without selection. *Genetical research*, **50**, 245–250.

Hudson RR, Kaplan NL (1985) Statistical properties of the number of recombination events in the history of a sample of DNA sequences. *Genetics*, **111**, 147–164.

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

Jeffery KJM, Bangham CRM (2000) Do infectious diseases drive MHC diversity? *Microbes and Infection*, **2**, 1335–1341.

Jin MS, Kim SE, Heo JY *et al.* (2007) Crystal structure of the TLR1-TLR2 heterodimer induced by binding of a tri-acylated lipopeptide. *Cell*, **130**, 1071–1082.

Jin MS, Lee JO (2008) Structures of the toll-like receptor family and its ligand complexes. *Immunity*, **29**, 182–191.

Kaiser P (2007) The avian immune genome--a glass half-full or half-empty? *Cytogenetic Genome Research*, **117**, 221–230.

Kalinowski ST (2005) Hp-Rare 1.0: a Computer Program for Performing Rarefaction on Measures of Allelic Richness. *Molecular Ecology Notes*, **5**, 187–189.

Kawai T, Akira S (2010) The role of pattern-recognition receptors in innate immunity: update on Toll-like receptors. *Nature Immunology*, **11**, 373–384.

Keestra M, de Zoete MR, Bouwman LI, van Putten JPM (2010) Chicken TLR21 is an innate CpG DNA receptor distinct from mammalian TLR9. *Journal of Immunology*, **185**, 460–7.

Kim HM, Park BS, Kim JI *et al.* (2007) Crystal structure of the TLR4-MD-2 complex with bound endotoxin antagonist Eritoran. *Cell*, **130**, 906–917.

Kimura M, Ohta T (1969) The Number of Heterozygous Nucleotide Sites Maintained in a Finite Population Due to Steady Flux of Mutations. *Genetics*, **61**, 763–771.

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

Kosakovsky Pond SL, Frost SDW (2005) Not so different after all: a comparison of methods for detecting amino acid sites under selection. *Molecular Biology and Evolution*, **22**, 1208–22.

Kryazhimskiy S, Plotkin JB (2008) The population genetics of dN/dS. *PLoS genetics*, **4**, e1000304.

Kuo C-H, Moran N a, Ochman H (2009) The consequences of genetic drift for bacterial genome complexity. *Genome Research*, **19**, 1450–4.

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

Leffler EM, Bullaughey K, Matute DR *et al.* (2012) Revisiting an old riddle: what determines genetic diversity levels within species? *PLoS Biology*, **10**, e1001388.

Li J, Li H, Jakobsson M *et al.* (2012) Joint analysis of demography and selection in population genetics: where do we stand and where could we go? *Molecular Ecology*, **21**, 28–44.

Librado P, Rozas J (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics*, **25**, 1451–2.

Luikart G, Sherwin WB, Steele BM, Allendorf FW (1998) Usefulness of molecular markers for detecting population bottlenecks via monitoring genetic change. *Molecular Ecology*, **7**, 963–74.

McDonald JH, Kreitman M (1991) Adaptive protein evolution at the Adh locus in Drosophila. *Nature*, **351**, 652–654.

Meyer D, Thomson G (2001) How selection shapes variation of the human major histocompatibility complex: a review. *Annals of Human Genetics*, **65**, 1–26.

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

Mitchell-Olds T, Schmitt J (2006) Genetic mechanisms and evolutionary significance of natural variation in Arabidopsis. *Nature*, **441**, 947–952.

Mukherjee S, Sarkar-roy N, Wagener DK, Majumder PP (2009) Signatures of natural selection are not uniform across genes of innate immune system , but purifying. *PNAS*, **106**, 7073–7078.

Murrell B, Moola S, Mabona A *et al.* (2013) FUBAR: a fast, unconstrained bayesian approximation for inferring selection. *Molecular Biology and Evolutioniology and evolution*, **30**, 1196–205.

Murrell B, Wertheim JO, Moola S *et al.* (2012) Detecting individual sites subject to episodic diversifying selection. *PLoS genetics*, **8**, e1002764.

Nakajima T, Ohtani H, Satta Y *et al.* (2008) Natural selection in the TLR-related genes in the course of primate evolution. *Immunogenetics*, **60**, 727–35.

Nei M (1987) *Molecular evolutionary genetics.* Columbia University Press, New York.

Netea MG, Wijmenga C, O'Neill L a J (2012) Genetic variation in Toll-like receptors and disease susceptibility. *Nature Immunology*, **13**, 535–42.

Niskanen a. K, Kennedy LJ, Ruokonen M *et al.* (2013) Balancing selection and heterozygote advantage in MHC loci of the bottlenecked Finnish wolf population. *Molecular Ecology*, n/a–n/a.

Ohta T (1995) Gene Conversion vs Point Mutation in Generating Variability at the Antigen Recognition Site of Major Histocompatibility Complex Loci. *Journal of Molecular Evolution*, **41**, 115–119.

Ohta T (2002) Near-neutrality in evolution of genes and gene regulation. *Proceedings of the National Academy of Sciences USA*, **99**, 16134–16137.

Oleksiak MF, Churchill G a, Crawford DL (2002) Variation in gene expression within and among natural populations. *Nature Genetics*, **32**, 261–266.

Van Oosterhout C (2009) A new theory of MHC evolution: beyond selection on the immune genes. *Proceedings. Biological sciences / The Royal Society*, **276**, 657–65.

Van Oosterhout C, Joyce DA, Cummings SA *et al.* (2006) Balancing Selection, Random Genetic Drift, and Genetic Variation at the Major Histocompatibility Complex in Two Wild Populations of Guppies (Poecilia reticulata). *Evolution*, **60**, 2562–2574.

Palti Y (2011) Toll-like receptors in bony fish: from genomics to function. *Developmental and Comparative Immunology*, **35**, 1263–72.

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

Piertney SB, Webster LM (2010) Characterising functionally important and ecologically meaningful genetic diversity using a candidate gene approach. *Genetica*, **138**, 419–432.

Potts WK, Slev PR (1995) Pathogen-based models favoring MHC genetic diversity. *Immunological Reviews*, **143**, 181–97.

Raymond M, Rousset F (1995) GENEPOP (Version 1.2): Population Genetics Software for Exact Tests and Ecumenicism. *Heredity*, **86**, 248–249.

Richardson DS, Burke T, Komdeur J (2002) Direct benefits and the evolution of female-biased cooperative breeding in Seychelles warblers. *Evolution*, **56**, 2313–21.

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

Richardson DS, Jury FL, Blaakmeer K, Komdeur J, Burke T (2001) Parentage assignment and extra-group paternity in a cooperative breeder: the Seychelles warbler (Acrocephalus sechellensis). *Molecular Ecology*, **10**, 2263–73.

Roach JC, Glusman G, Rowen L *et al.* (2005) The evolution of vertebrate Toll-like receptors. *Proceedings of the National Academy of Sciences USA*, **102**, 9577–9582.

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

Safford R, Hawkins F (2013) *The Birds of Africa, Volume 8: The Malagasy Region*.

Schlenke TA, Begun DJ (2003) Natural Selection Drives Drosophila Immune System Evolution. *Genetics*, **164**, 1471–1480.

Schnare M, Barton GM, Holt a C *et al.* (2001) Toll-like receptors control activation of adaptive immune responses. *Nature Immunology*, **2**, 947–50.

Schulze-Hagen K, Leisler B (2011) *Reed Warblers: Diversity in a Uniform Bird Family.* KNNV Uitgeverij; 1 edition.

Slade RW, McCallum HI (1992) Overdominant vs. Frequency-Dependent Selection at MHC Loci. *Genetics*, **132**, 861–862.

Spurgin LG, van Oosterhout C, Illera JC *et al.* (2011) Gene conversion rapidly generates major histocompatibility complex diversity in recently founded bird populations. *Molecular Ecology*, **20**, 5213–5225.

Spurgin LG, Richardson DS (2010) How pathogens drive genetic diversity: MHC, mechanisms and misunderstandings. *Proceedings of the Royal Society B*, **277**, 979–988.

Spurgin LG, Wright DJ, van der Velde M *et al.* (2014) Museum DNA reveals the demographic history of the endangered Seychelles warbler. *Evolutionary Applications*, 1–10.

Steadman DW, Greiner EC, Wood CS (1990) Absence of Blood Parasites in Indigenous and Introduced Birds from the Cook Islands , Pacific South. *Conservation Biology*, **4**, 398–404.

Stephens M, Donnelly P (2003) A comparison of bayesian methods for haplotype reconstruction from population genotype data. *American Journal of Human Genetics*, **73**, 1162–9.

Strand TM, Segelbacher G, Quintela M *et al.* (2012) Can balancing selection on MHC loci counteract genetic drift in small fragmented populations of black grouse? *Ecology and Evolution*, **2**, 341–353.

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

Tajima F (1989) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics*, **123**, 585–95.

Takahata N (1990) A simple genealogical structure of strongly balanced allelic lines and trans-species evolution of polymorphism. *Proceedings of the National Academy of Sciences of the United States of America*, **87**, 2419–2423.

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

Takahata N, Satta Y, Klein J (1992) Polymorphism and balancing selection at major histocompatibility complex oci. *Genetics*, **130TSK92**, 925–938.

Takeda K, Akira S (2005) Toll-like receptors in innate immunity. *International Immunology*, **17**, 1–14.

Takeuchi O, Sato S, Horiuchi T *et al.* (2002) Cutting Edge: Role of Toll-Like Receptor 1 in Mediating Immune Response to Microbial Lipoproteins. *The Journal of Immunology*, **169**, 10–14.

Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Molecular Biology and Evolution*, **24**, 1596–9.

Tompkins DM (2007) Population bottlenecks and avian immunity: implications for conservation. *Animal Conservation*, **10**, 11–13.

Tschirren B, Andersson M, Scherman K *et al.* (2013) Polymorphisms at the innate immune receptor TLR2 are associated with Borrelia infection in a wild rodent population Polymorphisms at the innate immune receptor TLR2 are associated with Borrelia infection in a wild rodent population.

Turner AK, Begon M, Jackson JA, Paterson S (2012) Evidence for selection at cytokine loci in a natural population of field voles (Microtus agrestis). *Molecular Ecology*, **21**, 1632–1646.

Uematsu S, Akira S (2008) Toll-Like receptors (TLRs) and their ligands. *Handbook of Experimental pharmacology*, 1–20.

Vasemagi A, Primmer CR (2005) Challenges for identifying functionally important genetic variation: the promise of combining complementary research strategies. *Molecular Ecology*, **14**, 3623–3642.

Vinkler M, Bryjova A, Albrecht T, Bryja J (2009) Identification of the first Toll-like receptor gene in passerine birds: TLR4 orthologue in zebra finch (Taeniopygia guttata). *Tissue Antigens*, **74**, 32–41.

Vögeli M, Lemus J a, Serrano D, Blanco G, Tella JL (2011) An island paradigm on the mainland: host population fragmentation impairs the community of avian pathogens. *Proceedings of the Royal Society B*, **278**, 2668–76.

Weir BS, Cockerham CC (1984) Estimating F-Statistics for the Analysis of Population Structure. *Evolution*, **38**, 1358–1370.

Werling D, Jann OC, Offord V, Glass EJ, Coffey TJ (2009) Variation matters: TLR structure and species-specific pathogen recognition. *Trends in Immunology*, **30**, 124–130.

Westerdahl H, Asghar M, Hasselquist D, Bensch S (2012) Quantitative disease resistance: to better understand parasite-mediated selection on major histocompatibility complex. *Proceedings of the Royal Society B*, **279**, 577–584.

Willi Y, Van Buskirk J, Schmid B, Fischer M (2007) Genetic isolation of fragmented populations is exacerbated by drift and selection. *Journal of Evolutionary Biology*, **20**, 534–42.

Worley K, Collet J, Spurgin LG *et al.* (2010) MHC heterozygosity and survival in red junglefowl. *Molecular Ecology*, **19**, 3064–3075.

Wright DJ, Spurgin LG, Collar NJ *et al.* (2014) The impact of translocations on neutral and functional genetic diversity within and among populations of the Seychelles warbler. *Molecular Ecology*, **23**, 2165–2177.

Yang Z, Bielawski J (2000) Statistical methods for detecting molecular adaptation. *Trends in Ecology & Evolution*, **15**, 496–503.

Young A, Boyle T, Brown T (1996) The population genetic consequences of habitat fragmentation for plants. *Trends in Ecology & Evolution*, **11**, 413–418.

Supplementary material

Table S1. Primers and PCR annealing temperatures used to amplify TLR loci in seven *Acrocephalus* species.

Table S2. Haplotype-level tests for selection based on the allele frequency spectrum for each TLR locus for the Seychelles warbler. Significant *P*-values are in bold.

Table S3. Z-tests of selection based upon dN/dS for each TLR locus for both the Seychelles warbler (SW) and all other *Acrocephalus* species (OW): *A. arundinaceus, A. australis, A. brevipennis, A. scirpaceus, A. schoenobaenus* and *A. taiti*. Significant *P*-values are in bold.

Table S4. McDonald-Kreitman's test for selection within and between species for each TLR locus and all pairwise combinations of all *Acrocephalus* species: *A. arundinaceus, A. australis, A. brevipennis, A. scirpaceus, A. schoenobaenus, A. sechellensis* and *A. taiti*. Significant *P*-values are in bold. 'NA' denotes when the McDonald-Kreitman contingency table not be computed as not all components of the table have sufficient data.

Figure S1. Observed and expected haplotype frequency charts for each polymorphic TLR locus amplified in the Seychelles warbler.

Figure S2. Maximum-likelihood trees for each Toll-like receptor (TLR) locus to show the relationship between alleles at each locus across different avian lineages. Bootstrapping is applied to each relationship with 1000 repetitions and the tree is drawn to scale, with branch lengths measured in number of substitutions per site. Trees include all sequences obtained for the Seychelles warbler (SW) and six other *Acrocephalus* species (OW) and reference sequences of other passerines and non-passerine species to root the trees.

Table S1.

Locus	Primer Name	Primer Sequence 5'-3'	Species	Anneal T °C
TLR1LA	avTLR1LAF	GATGGAATGAGCACTTCAGA	*Acrocephalus brevipennis*	62
	avTLR1LAR	CTTCGTCTGCGTCCACTG	*Acrocephalus arundinaceus*	62
			Acrocephalus taiti	62
			Acrocephalus australis	62
			Acrocephalus sechellensis	60
TLR1LB	avTLR1LBF	TCCAGGYTWCAAAATCTGACAC	*Acrocephalus brevipennis*	60
	avTLR1LBR	CGGCACRTCCARGTAGATG	*Acrocephalus arundinaceus*	62
			Acrocephalus taiti	60
			Acrocephalus scirpaceus	60
			Acrocephalus schoenobaenus	60
			Acrocephalus australis	60
			Acrocephalus sechellensis	60
TLR3	avTLR3F	CAAWGTTGAACTTGGTGAAAAT	*Acrocephalus brevipennis*	57
	avTLR3R	TCACAGGTRCAATCAAANGG	*Acrocephalus arundinaceus*	58
			Acrocephalus taiti	57
			Acrocephalus scirpaceus	57
			Acrocephalus schoenobaenus	57
			Acrocephalus australis	57
			Acrocephalus sechellensis	55
TLR4	PauTLR4F	GCTTTCCTTGAACAACATAAAGTCC	*Acrocephalus brevipennis*	55
	PauTLR4R	GGGACAGAAAGACAGGGTAGG	*Acrocephalus arundinaceus*	58
			Acrocephalus taiti (HW)	55
			Acrocephalus scirpaceus	55
			Acrocephalus schoenobaenus	55
			Acrocephalus australis	55
			Acrocephalus sechellensis	58
TLR5	avTLR5F	GTAATCTTACCAGCTTCCAAGG	*Acrocephalus taiti*	61
	avTLR5R	GCTGGAGTTCATCTTCATC	*Acrocephalus arundinaceus*	62
			Acrocephalus scirpaceus	61
			Acrocephalus schoenobaenus	61
			Acrocephalus sechellensis	55
TLR15	FinchTLR15F	GATCTCCCATCCCACCTGA	*Acrocephalus brevipennis*	58
	avTLR15R	AAGGAGATCTTATTCCCTG	*Acrocephalus arundinaceus*	58
			Acrocephalus scirpaceus	60
			Acrocephalus australis	57
			Acrocephalus sechellensis	57
TLR21	FinchTLR21F	TTGACAACAACCTGCTCACTG	*Acrocephalus arundinaceus*	58-60
	FinchTLR21R	TACGCAGCTCGTTCTTGG	*Acrocephalus australis*	60
			Acrocephalus brevipennis	58
			Acrocephalus taiti	58
			Acrocephalus sechellensis	58

Table S2.

Locus	Number of individuals	Tajima's D	Fu & Li's D	Fu & Li's F
TLR1LA	22	0.78 (>0.1)	0.55 (>0.1)	0.72 (>0.1)
TLR1LB	33	1.54 (>0.1)	0.72 (>0.1)	1.13 (>0.1)
TLR3	28	0.85 (>0.1)	0.88 (>0.1)	1.02 (>0.1)
TLR4	30	-	-	-
TLR5	23	-0.98 (>0.1)	0.76 (>0.1)	0.29 (>0.1)
TLR15	30	0.84 (>0.1)	0.87 (>0.1)	1.01 (>0.1)
TLR21	30	-	-	-

Table S3.

Locus	Species Group	Z-test of positive selection (dN > dS)	Z-test of negative selection (dN < dS)
		dN-dS	dS-dN
TLR1LA	SW	-1.02 (1.00)	1.06 (0.15)
	OW	1.07 (0.14)	-1.11 (1.00)
TLR1LB	SW	-1.38 (1.00)	1.39 (0.08)
	OW	1.12 (0.13)	-1.15 (1.00)
TLR3	SW	-0.62 (1.00)	0.62 (0.27)
	OW	-0.80 (1.00)	0.82 (0.21)
TLR4	SW	-	-
	OW	-2.38 (1.00)	2.44 (0.08)
TLR5	SW	-0.57 (1.00)	0.57 (0.29)
	OW	0.19 (0.43)	-0.18 (1.00)
TLR15	SW	1.17 (0.12)	-1.14 (1.00)
	OW	-0.71 (1.00)	0.70 (0.24)
TLR21	SW	-	-
	OW	-0.62 (1.00)	0.62 (0.27)

***P*-values given in brackets

Table S4.

Locus	Species paired with SW	dS		dN		P-value
		Fixed differences between species	S	Fixed differences between species	S	
TLR1LA	A. arundinaceus	1	2	3	2	1.00
	A. australis	1	1	3	4	1.00
	A. brevipennis	2	1	4	2	1.00
	A. taiti	1	1	3	0	0.40
TLR1LB	A. arundinaceus	0	5	0	3	NA
	A. australis	0	2	1	5	1.00
	A. brevipennis	0	2	2	2	0.47
	A. schoenobaenus	2	2	4	1	0.53
	A. scirpaceus	1	3	1	1	1.00
	A. taiti	0	2	1	1	1.00
TLR3	A. arundinaceus	1	3	1	2	1.00
	A. australis	1	1	1	3	1.00
	A. brevipennis	1	1	2	2	1.00
	A. schoenobaenus	2	0	2	2	0.47
	A. taiti	1	1	2	3	1.00
TLR4	A. arundinaceus	5	1	0	0	NA
	A. australis	**5**	**0**	**0**	**2**	**0.05**
	A. brevipennis	4	6	0	0	NA
	A. schoenobaenus	9	4	2	2	0.58
	A. scirpaceus	**6**	**5**	**0**	**6**	**0.04**
	A. taiti	5	0	0	1	0.17
TLR5	A. arundinaceus	0	1	0	0	NA
	A. brevipennis	0	0	0	0	NA
	A. scirpaceus	0	1	3	0	0.25
	A. taiti	0	0	0	3	NA
TLR15	A. arundinaceus	0	0	1	3	NA
	A. australis	6	1	9	3	1.00
	A. brevipennis	1	0	0	3	0.25
	A. scirpaceus	0	0	1	3	NA
	A. taiti	0	0	1	3	NA
TLR21	A. arundinaceus	1	0	0	1	1.00
	A. australis	0	1	0	1	NA

Figure S1.

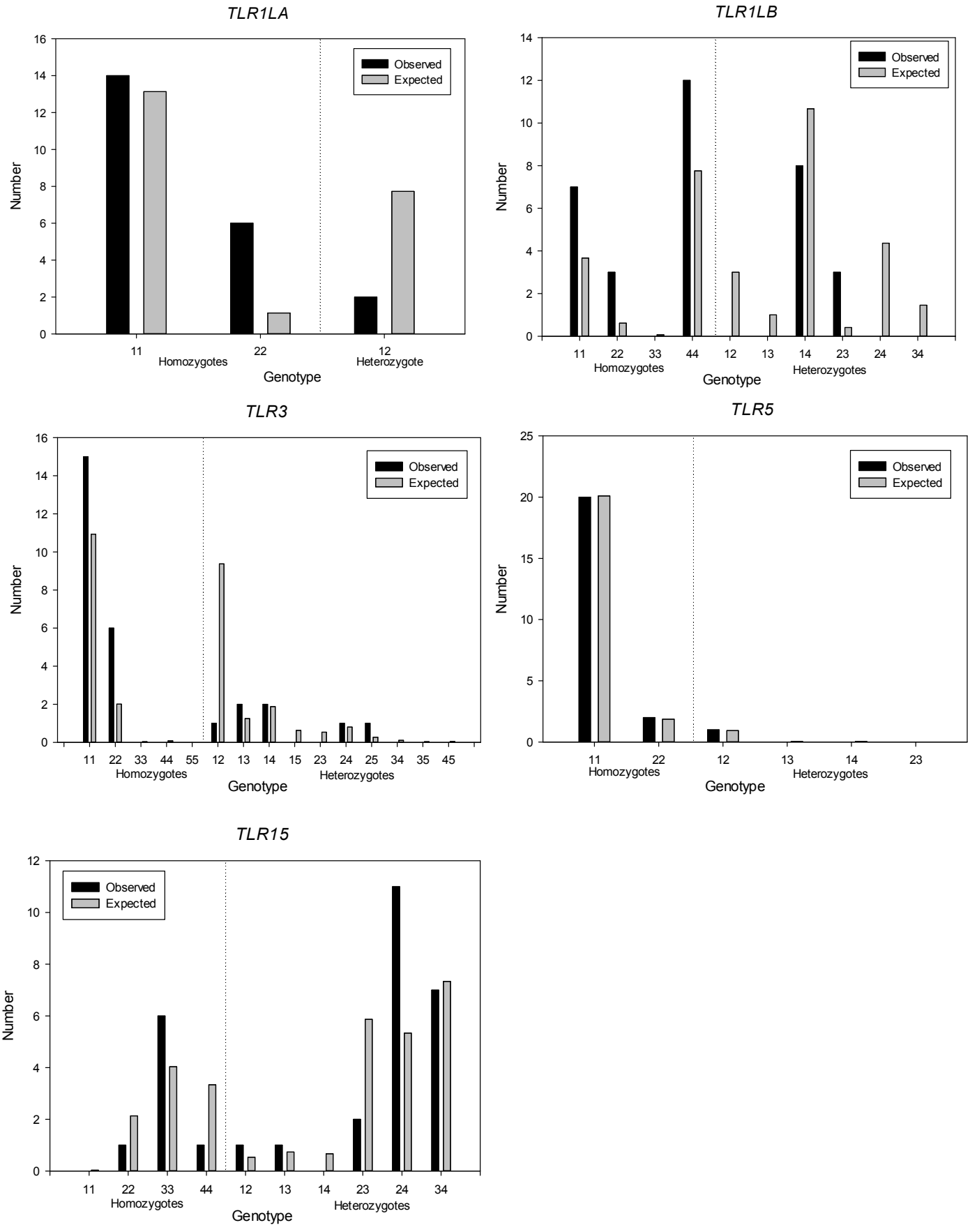
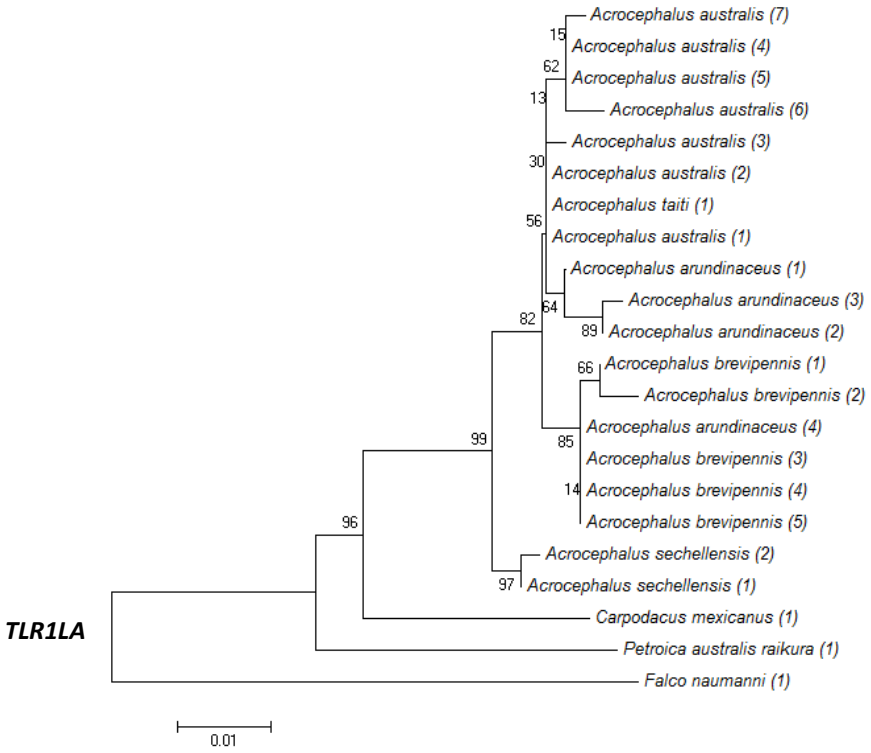
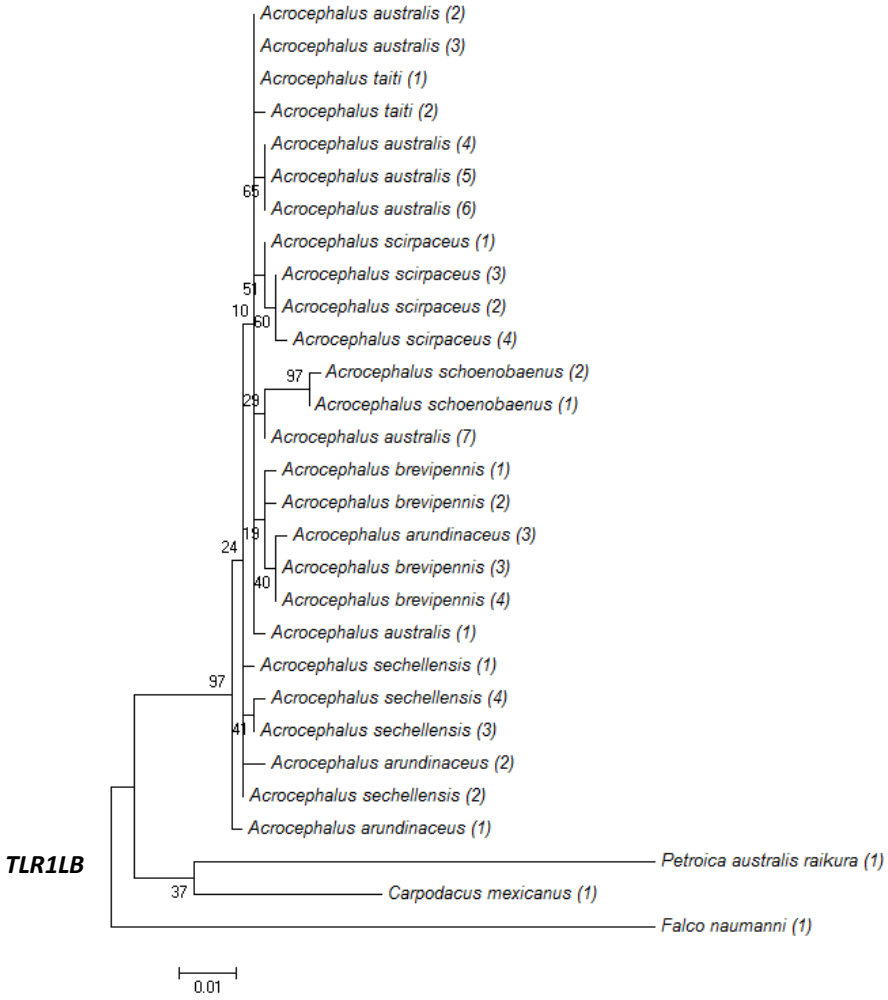
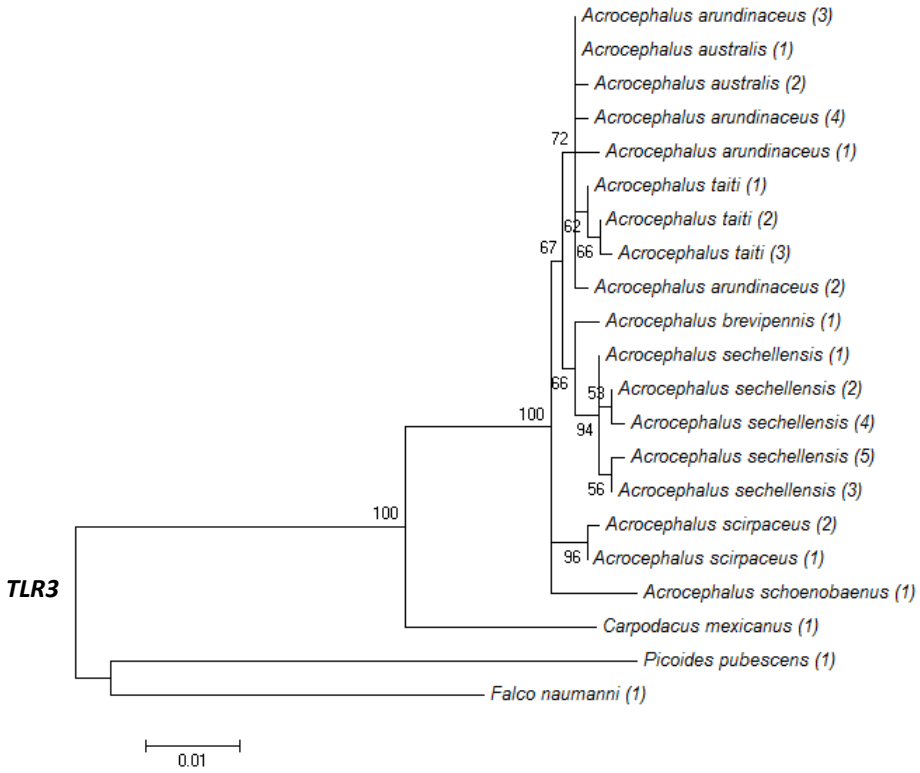
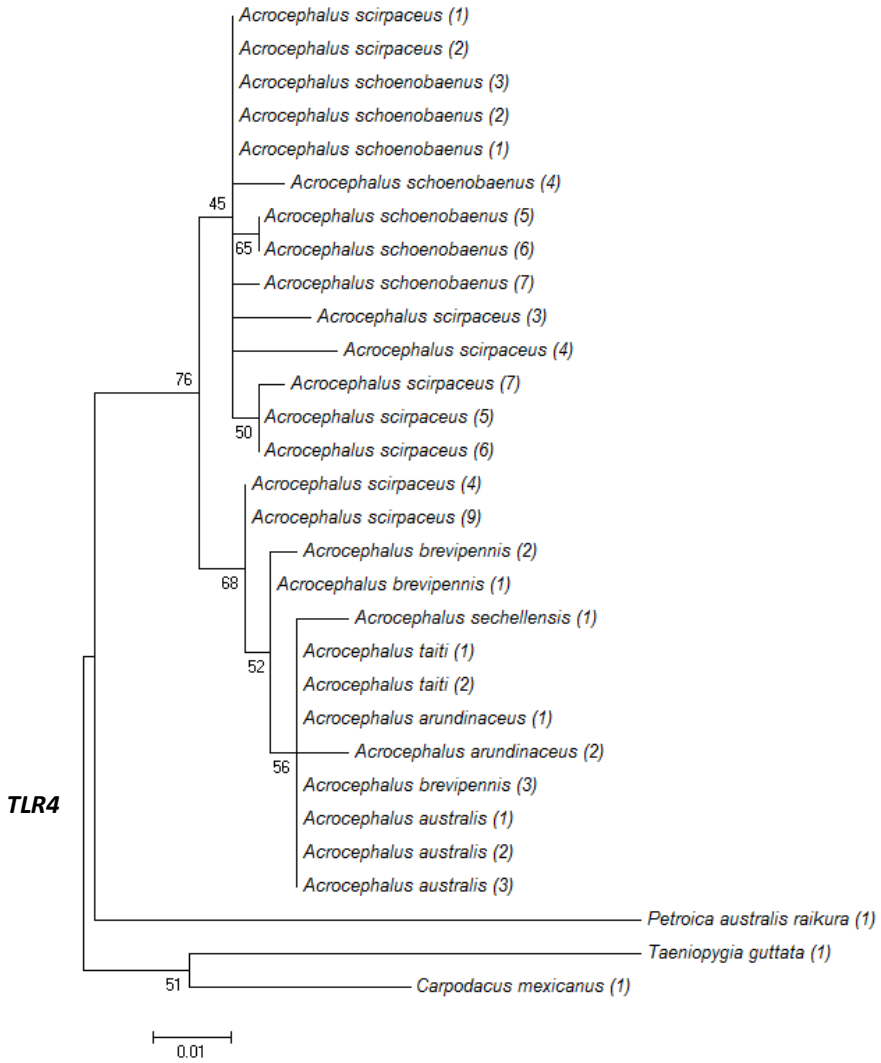

TLR1LA
Observed
Expected
Number
Genotype
11
22
12
Homozygotes
Heterozygote
TLR1LB
Observed
Expected
Number
Genotype
11
22
33
44
12
13
14
23
24
34
Homozygotes
Heterozygotes
TLR3
Observed
Expected
Number
Genotype
11
22
33
44
55
12
13
14
15
23
24
25
34
35
45
Homozygotes
Heterozygotes
TLR5
Observed
Expected
Number
Genotype
11
22
12
13
14
23
Homozygotes
Heterozygotes
TLR15
Observed
Expected
Number
Genotype
11
22
33
44
12
13
14
23
24
34
Homozygotes
Heterozygotes

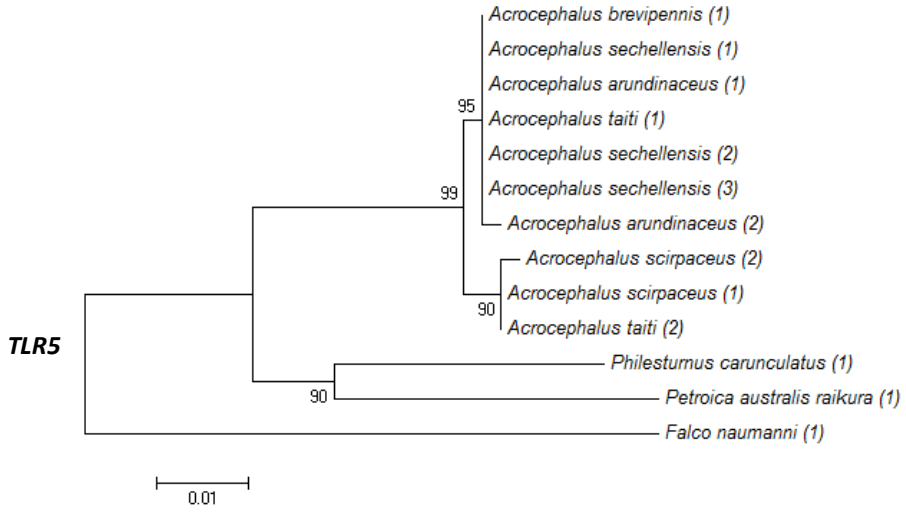
Figure S2.

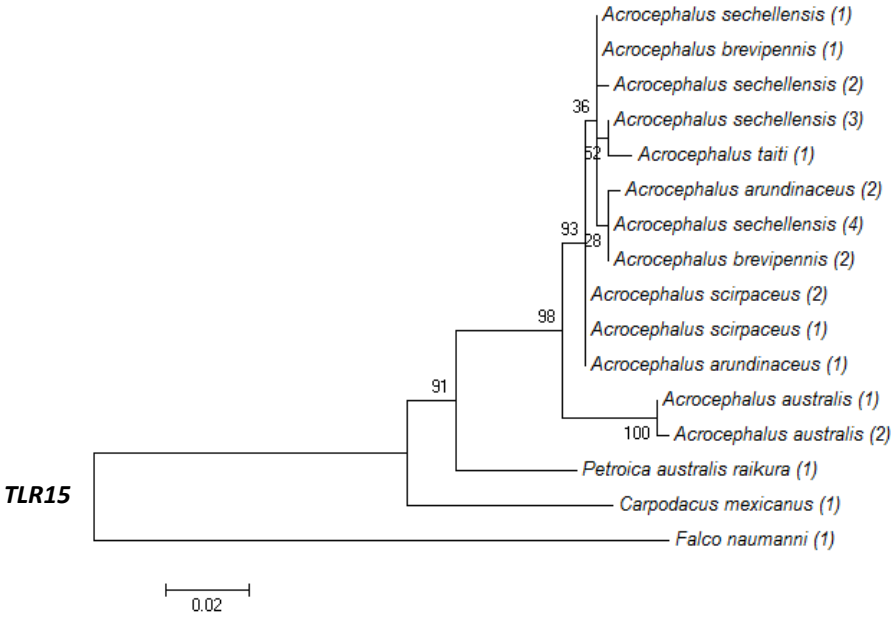

TLR1LA
Acrocephalus australis (7)
Acrocephalus australis (4)
Acrocephalus australis (5)
Acrocephalus australis (6)
Acrocephalus australis (3)
Acrocephalus australis (2)
Acrocephalus taiti (1)
Acrocephalus australis (1)
Acrocephalus arundinaceus (1)
Acrocephalus arundinaceus (3)
Acrocephalus arundinaceus (2)
Acrocephalus brevipennis (1)
Acrocephalus brevipennis (2)
Acrocephalus arundinaceus (4)
Acrocephalus brevipennis (3)
Acrocephalus brevipennis (4)
Acrocephalus brevipennis (5)
Acrocephalus sechellensis (2)
Acrocephalus sechellensis (1)
Carpodacus mexicanus (1)
Petroica australis raikura (1)
Falco naumanni (1)
15
62
13
30
56
64
89
82
66
85
99
14
96
97
0.01

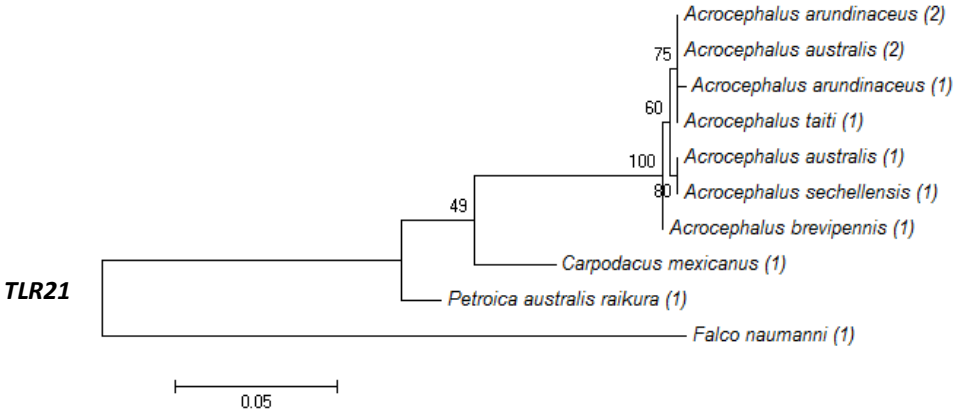

TLR1LB
Acrocephalus australis (2)
Acrocephalus australis (3)
Acrocephalus taiti (1)
Acrocephalus taiti (2)
Acrocephalus australis (4)
Acrocephalus australis (5)
Acrocephalus australis (6)
Acrocephalus scirpaceus (1)
Acrocephalus scirpaceus (3)
Acrocephalus scirpaceus (2)
Acrocephalus scirpaceus (4)
Acrocephalus schoenobaenus (2)
Acrocephalus schoenobaenus (1)
Acrocephalus australis (7)
Acrocephalus brevipennis (1)
Acrocephalus brevipennis (2)
Acrocephalus arundinaceus (3)
Acrocephalus brevipennis (3)
Acrocephalus brevipennis (4)
Acrocephalus australis (1)
Acrocephalus sechellensis (1)
Acrocephalus sechellensis (4)
Acrocephalus sechellensis (3)
Acrocephalus arundinaceus (2)
Acrocephalus sechellensis (2)
Acrocephalus arundinaceus (1)
Petroica australis raikura (1)
Carpodacus mexicanus (1)
Falco naumanni (1)
65
51
10
60
97
29
19
24
40
97
41
37
0.01


TLR3
Acrocephalus arundinaceus (3)
Acrocephalus australis (1)
Acrocephalus australis (2)
Acrocephalus arundinaceus (4)
Acrocephalus arundinaceus (1)
Acrocephalus taiti (1)
Acrocephalus taiti (2)
Acrocephalus taiti (3)
Acrocephalus arundinaceus (2)
Acrocephalus brevipennis (1)
Acrocephalus sechellensis (1)
Acrocephalus sechellensis (2)
Acrocephalus sechellensis (4)
Acrocephalus sechellensis (5)
Acrocephalus sechellensis (3)
Acrocephalus scirpaceus (2)
Acrocephalus scirpaceus (1)
Acrocephalus schoenobaenus (1)
Carpodacus mexicanus (1)
Picoides pubescens (1)
Falco naumanni (1)
72
62
67
66
66
53
100
94
56
100
96
0.01


TLR4
Acrocephalus scirpaceus (1)
Acrocephalus scirpaceus (2)
Acrocephalus schoenobaenus (3)
Acrocephalus schoenobaenus (2)
Acrocephalus schoenobaenus (1)
Acrocephalus schoenobaenus (4)
Acrocephalus schoenobaenus (5)
Acrocephalus schoenobaenus (6)
Acrocephalus schoenobaenus (7)
Acrocephalus scirpaceus (3)
Acrocephalus scirpaceus (4)
Acrocephalus scirpaceus (7)
Acrocephalus scirpaceus (5)
Acrocephalus scirpaceus (6)
Acrocephalus scirpaceus (4)
Acrocephalus scirpaceus (9)
Acrocephalus brevipennis (2)
Acrocephalus brevipennis (1)
Acrocephalus sechellensis (1)
Acrocephalus taiti (1)
Acrocephalus taiti (2)
Acrocephalus arundinaceus (1)
Acrocephalus arundinaceus (2)
Acrocephalus brevipennis (3)
Acrocephalus australis (1)
Acrocephalus australis (2)
Acrocephalus australis (3)
Petroica australis raikura (1)
Taeniopygia guttata (1)
Carpodacus mexicanus (1)
45
65
76
50
68
52
56
51
0.01


TLR5
Acrocephalus brevipennis (1)
Acrocephalus sechellensis (1)
Acrocephalus arundinaceus (1)
Acrocephalus taiti (1)
Acrocephalus sechellensis (2)
Acrocephalus sechellensis (3)
Acrocephalus arundinaceus (2)
Acrocephalus scirpaceus (2)
Acrocephalus scirpaceus (1)
Acrocephalus taiti (2)
Philesturnus carunculatus (1)
Petroica australis raikura (1)
Falco naumanni (1)
95
99
90
90
0.01


TLR15
Acrocephalus sechellensis (1)
Acrocephalus brevipennis (1)
Acrocephalus sechellensis (2)
Acrocephalus sechellensis (3)
Acrocephalus taiti (1)
Acrocephalus arundinaceus (2)
Acrocephalus sechellensis (4)
Acrocephalus brevipennis (2)
Acrocephalus scirpaceus (2)
Acrocephalus scirpaceus (1)
Acrocephalus arundinaceus (1)
Acrocephalus australis (1)
Acrocephalus australis (2)
Petroica australis raikura (1)
Carpodacus mexicanus (1)
Falco naumanni (1)
36
52
93
28
98
91
100
0.02


TLR21
Acrocephalus arundinaceus (2)
Acrocephalus australis (2)
Acrocephalus arundinaceus (1)
Acrocephalus taiti (1)
Acrocephalus australis (1)
Acrocephalus sechellensis (1)
Acrocephalus brevipennis (1)
Carpodacus mexicanus (1)
Petroica australis raikura (1)
Falco naumanni (1)
75
60
100
80
49
0.05

Additional information

Figure 1. Rarefaction curve of number of unique alleles observed with increasing sample size of individuals sampled, calculated in the program HpRare v1.0 (Kalinowski 2005).

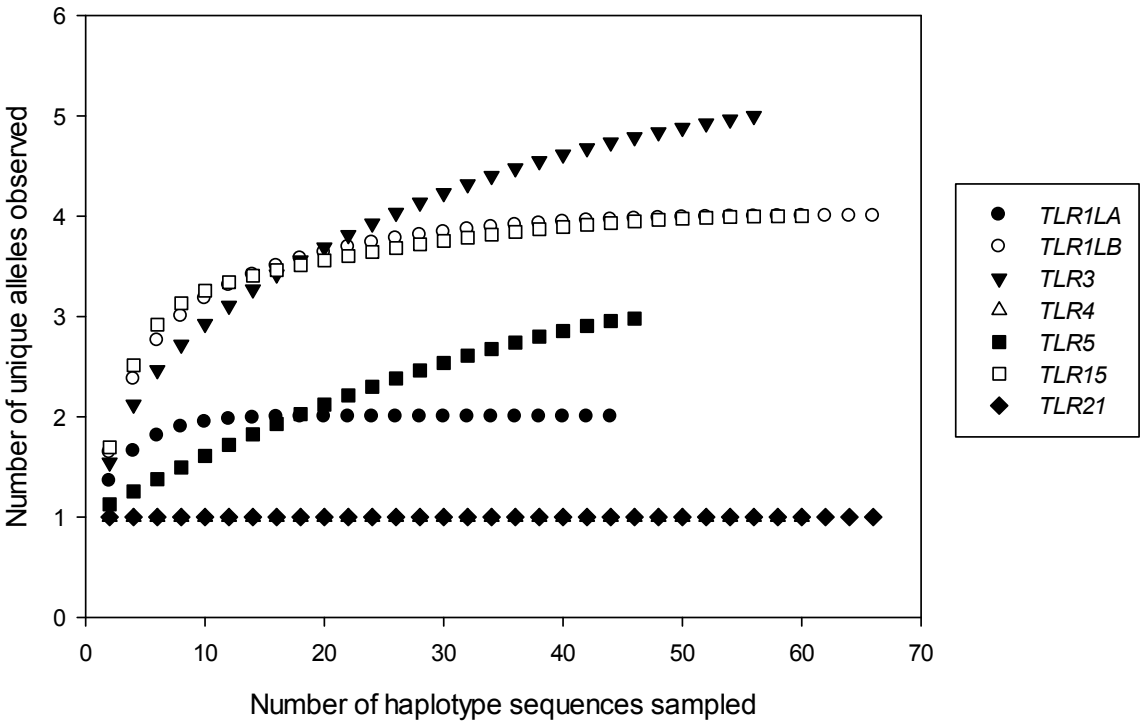
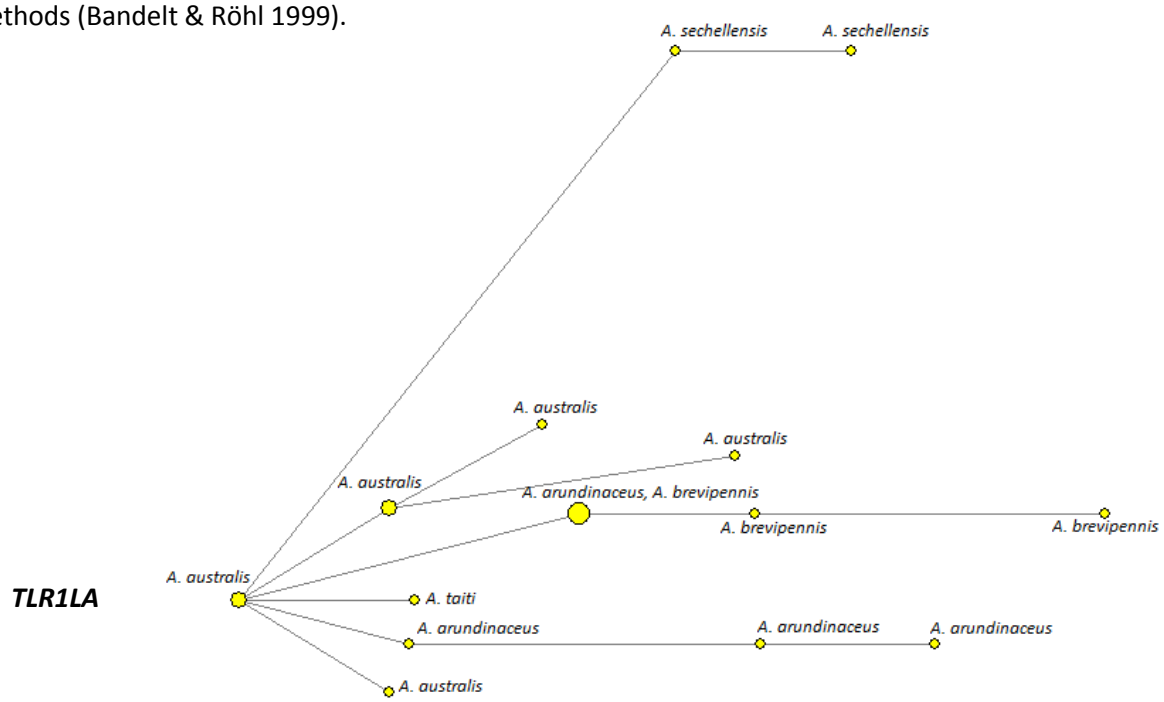
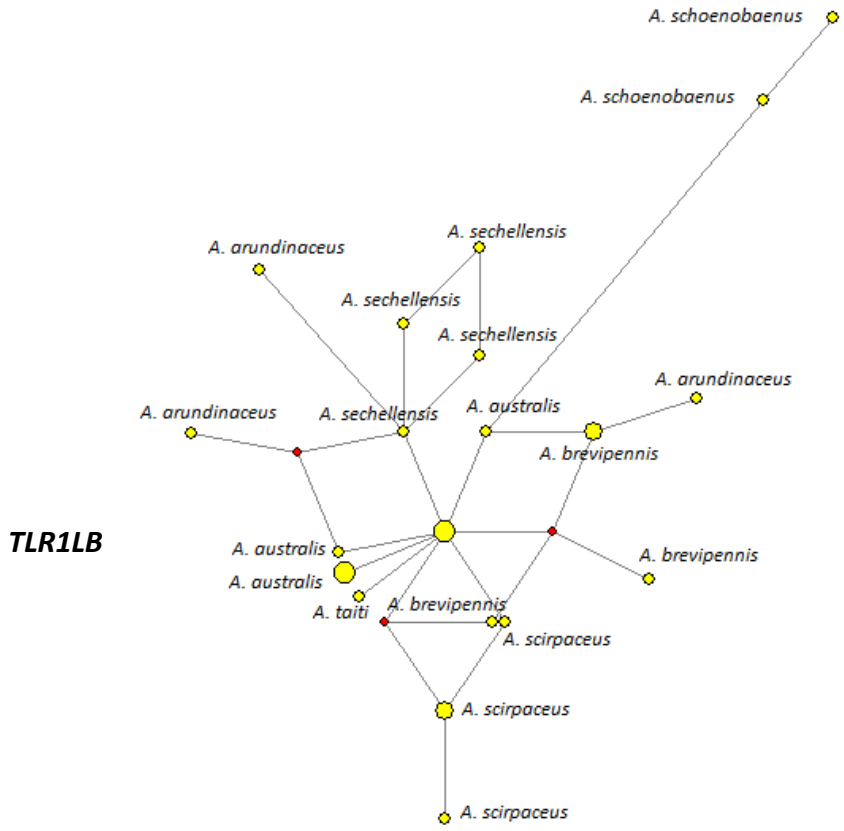
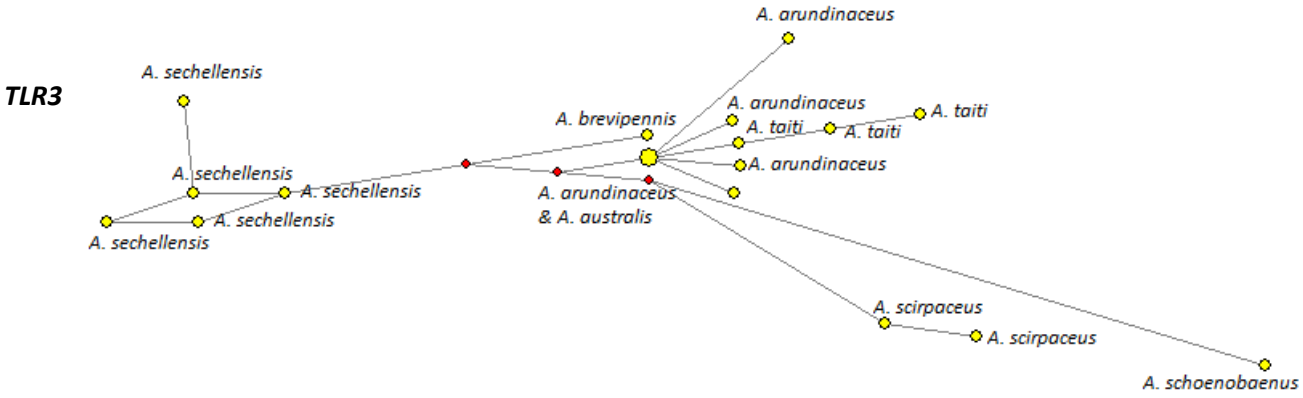
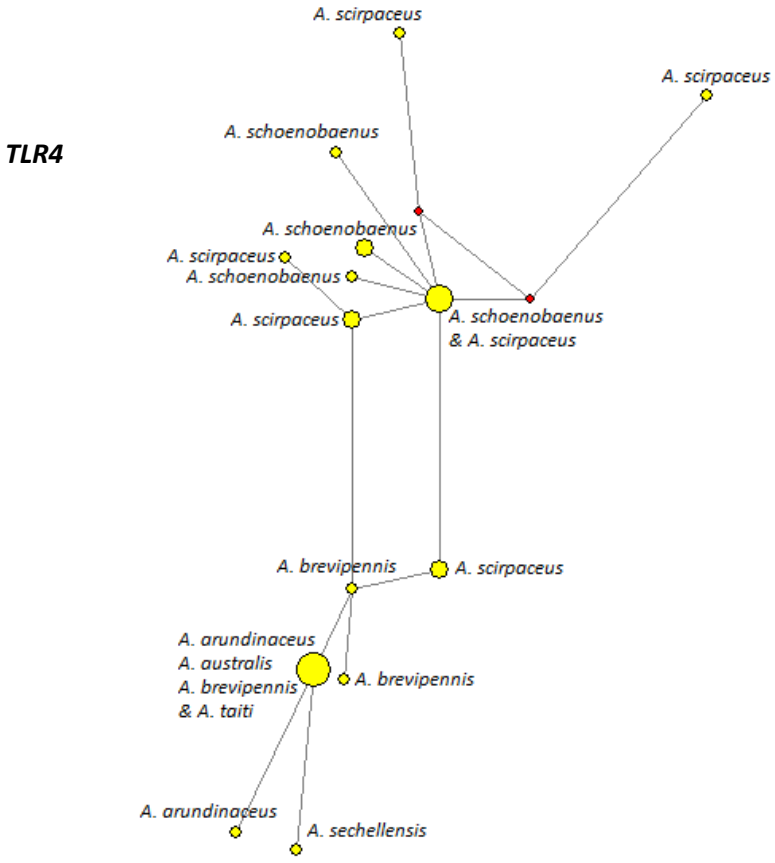


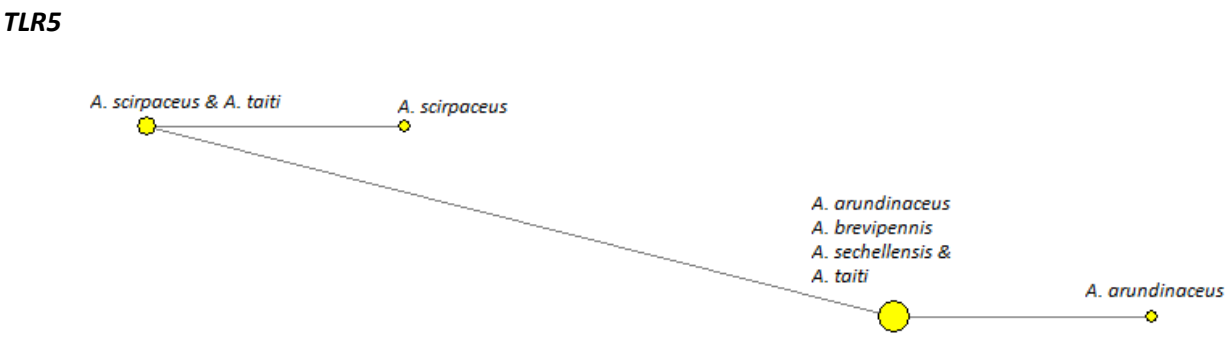
Figure 2. Haplotype networks for each TLR locus characterised in the *Acrocephalus* genus. Nodes are proportional to frequency of haplotypes observed and joining branches represent genetic distances over evolutionary time. Red dots donate key mutational steps between haplotypes All networks were constructed in the program Fluxus (fluxus-engineering.com) based on neighbour-joining methods (Bandelt & Röhl 1999).



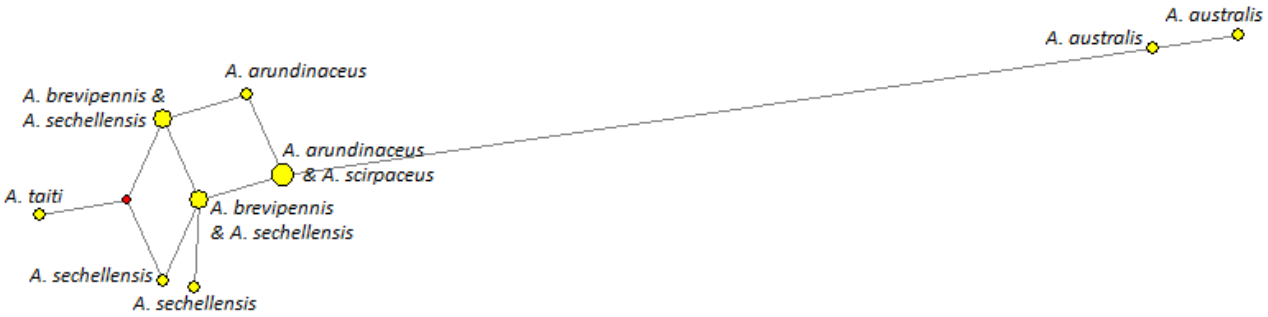



TLR3
A. arundinaceus
A. sechellensis
A. brevipennis
A. arundinaceus
A. taiti
A. taiti
A. taiti
A. arundinaceus
A. sechellensis
A. sechellensis
A. sechellensis
A. sechellensis
A. arundinaceus
& A. australis
A. scirpaceus
A. scirpaceus
A. schoenobaenus

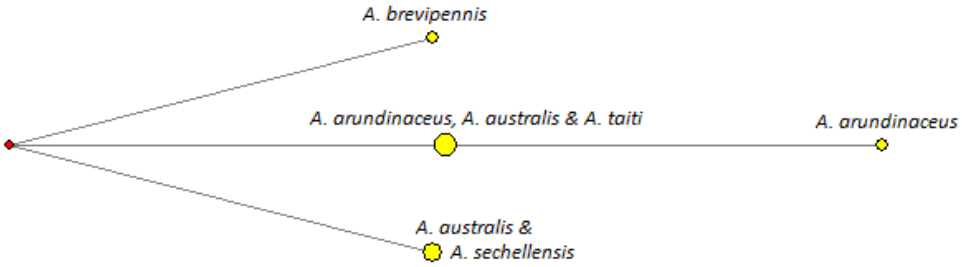

TLR4
A. scirpaceus
A. scirpaceus
A. schoenobaenus
A. schoenobaenus
A. scirpaceus
A. schoenobaenus
A. scirpaceus
A. schoenobaenus
& A. scirpaceus
A. brevipennis
A. scirpaceus
A. arundinaceus
A. australis
A. brevipennis
& A. taiti
A. brevipennis
A. arundinaceus
A. sechellensis


TLR5
A. scirpaceus & A. taiti
A. scirpaceus
A. arundinaceus
A. brevipennis
A. sechellensis &
A. taiti
A. arundinaceus

TLR15


A. australis
A. australis
A. arundinaceus
A. brevipennis &
A. sechellensis
A. arundinaceus
& A. scirpaceus
A. taiti
A. brevipennis
& A. sechellensis
A. sechellensis
A. sechellensis

TLR21


A. brevipennis
A. arundinaceus, A. australis & A. taiti
A. arundinaceus
A. australis &
A. sechellensis

Chapter 4: Simulating selection at Toll-like receptors (TLRs) in the Seychelles warbler

Abstract

Pathogen-mediated selection (PMS) can maintain immunogenetic variation within host populations, but how PMS acts on such variation within bottlenecked isolated populations with a depauperate parasite fauna remains a subject of considerable debate. Toll-like receptor genes (TLRs) play a fundamental role in vertebrate immune defence and are predicted to be under PMS. We previously characterised variation at TLR loci in the Seychelles warbler (*Acrocephalus sechellensis*), an endemic passerine that has undergone a severe recent bottleneck. We found that five out of seven TLR loci were polymorphic, which is in sharp contrast to the low levels of genome-wide variation observed in this species. As is often the case however, standard population genetic statistical methods failed to detect a contemporary signature of selection at any of the TLR loci. Therefore, we applied forward-in-time computer simulations to delineate demographic effects from the effects of selection. Our simulations rejected neutral evolution in all five polymorphic TLR genes in this species. Weak balancing selection appears to have acted in the recent past on the five TLR genes with estimated selection coefficients ranging from $0.005 < S < 0.03$. The model could not discern whether balancing selection has been acting during the actual bottleneck, with drift being the overriding evolutionary force. Forecast models predict that immunogenetic variation in the Seychelles warbler will continue to erode, but only if PMS has ceased to operate. Such 'drift debt' occurs when a genepool reaches its new equilibrium level of polymorphism, and this loss is likely to be an important threat to many recently-bottlenecked populations.

Introduction

Understanding how random drift and selection affect genetic diversity in populations is important from a conservation perspective, given that genetic variation is key to a population's short and long term health though it links to inbreeding depression and adaptive potential, respectively (Frankham *et al.* 1999). Selection can potentially maintain variation in a gene pool in the face of drift, yet the relative role of these different evolutionary forces in bottlenecked populations in shaping variation in the wild remains unclear (Acevedo-Whitehouse & Cunningham 2006). Genetic drift is often thought to outweigh the effects of selection in small (bottlenecked) populations, resulting in a loss of genetic variation and leading to population differentiation and isolation (Miller & Lambert 2004; Grueber *et al.* 2013). However, even though the post-bottleneck population may have lost significant variation compared to the ancestral gene pool, this does not preclude the possibility that balancing selection has occurred. Furthermore, given that recently bottlenecked populations are unlikely to be in mutation-drift-selection equilibrium, the currently observed levels of polymorphism might overestimate future genetic diversity.

Immune genes are ideal candidates with which to investigate the link between genetic variation and fitness because of their direct effects on survival (e.g. Sorci & Moller 1997; Merino *et al.* 2000; Sol *et al.* 2003; Moller & Saino 2004) and reproductive success (e.g. Pedersen & Greives 2008; Kalbe *et al.* 2009; la Puente *et al.* 2010; Radwan *et al.* 2012). Furthermore, variation at immune genes can have an important impact on the demographic structure of populations (e.g. Hudson 1986; Redpath *et al.* 2006; Deter *et al.* 2007; Pedersen & Greives 2008) and they are thought to evolve faster than the rest of the genome as a result of host-pathogen co-evolution (Trowsdale & Parham 2004). Much work has been done investigating how pathogen-mediated selection (PMS) can maintain genetic variation at genes of the Major Histocompatibility Complex (MHC) (for reviews, see Piertney & Oliver 2006; Spurgin *et al.* 2011). However, the MHC is a large multigene family with a complex evolutionary history and with many evolutionary forces acting simultaneously on the various gene members (e.g. van Oosterhout 2009) and events such as gene conversion (e.g. Spurgin *et al.* 2011), which can complicate population genetic analysis of wild populations. Other important immune genes exist that remain relatively understudied, and these genes are increasingly recognised as excellent candidates for investigating functional variation and

selection (e.g. Bollmer *et al.* 2011; Turner *et al.* 2012; Grueber *et al.* 2013; for review, see Acevedo-Whitehouse & Cunningham 2006).

Toll-like receptors (TLRs) are membrane-bound sensors of the vertebrate immune system that function in recognising pathogen-associated molecular patterns (PAMPs) and triggering an appropriate immune response (Akira *et al.* 2001; Werling & Jungi 2003). Vertebrate TLRs fall into six different families depending on the specific PAMPs they recognise (Takeda & Akira 2005; Kawai & Akira 2010). Different TLRs bind to different elements, ranging from bacterial lipoproteins (Takeuchi *et al.,* 2002; Jin *et al.,* 2007), lipopolysaccharides (Bihl *et al.,* 2003; Kim *et al.,* 2007), DNA motifs (Keestra *et al.,* 2010; Brownlie & Allan, 2011) and viral RNA (Yoneyama & Fujita 2010). Studies have shown evidence of positive selection acting within TLR loci across a range of vertebrate taxa (e.g. Ferrer-admetlla *et al.* 2008; Nakajima *et al.* 2008; Areal *et al.* 2011; Palti 2011; Grueber *et al.* 2014). It appears that this selection largely targets the TLR extracellular domain responsible for binding PAMPs (for reviews, seeTakeda & Akira, 2005; Kawai & Akira, 2010). Assuming that TLRs are involved in a co-evolutionary arms race with pathogens, it is likely that balancing selection operates at these genes. This idea is also supported by the direct links that have been made between *in vitro* nucleotide variation at these genes with differential disease outcome (Hellgren *et al.* 2010; Basu *et al.* 2012; Netea *et al.* 2012).

Avian models are widely used for looking into patterns of functional variation (e.g. Hellgren & Ekblom 2010; Bonneaud *et al.* 2011; Kyle *et al.* 2014; Staley & Bonneaud 2015). A study on the entire TLR multigene family in seven phylogenetically-diverse avian species has inferred polymorphic TLRs to be under strong balancing selection (Alcaide & Edwards 2011). However, it can be difficult to examine the causes and consequences of functional variation in wild avian populations. The Seychelles warbler, *Acrocephalus sechellensis*, is an island endemic passerine species that was, because of anthropogenic effects, reduced to the verge of extinction with less than 30 individuals remaining on a single island during the last century (Collar & Stuart 1985). In a previous study, we characterised variation at TLR genes in the bottlenecked population of the Seychelles warbler (Chapter 3). We found that despite the considerable losses in genome-wide variation due to the bottleneck (Spurgin *et al.* 2014), considerable polymorphism remained at five different TLR loci (*TLR1LA, TLR1LB,*

TLR3, TLR5 and *TLR15*) while two loci were monomorphic (*TLR4* and *TLR21*). Remarkably, four functional variants (alleles) were found at a single locus (*TLR15*) (Table 1).

Due to the overwhelming effect of stochastic processes, detecting any possible signature of selection in bottlenecked populations using standard population genetic statistical methods is difficult since they make unrealistic demographic assumptions, such as constant population size and no population structure. Furthermore, they fail to distinguish historic selection from current selection. Possibly due to this limitation, a previous study failed to detect the evidence of balancing selection acting on TLR variation in the Seychelles warbler (Chapter 3). Additionally, rejection of neutral evolution only indicates that a population is not in mutation-drift equilibrium. Such deviation from equilibrium is consistent with both the effects of selection as well as a post-bottleneck population expansion, thus making the interpretation of such tests complicated (Ramírez-Soriano *et al.* 2008). This problem is particularly acute in relation to conservation genetics, given that by definition, endangered populations are not in equilibrium.

In an attempt to resolve this problem, we designed an individual-based model that uses forward-in-time simulations to account for the stochasticity during the population bottleneck. By simulating the exact bottleneck scenario, as previously inferred through neutral markers and historic data (Spurgin *et al.* 2014), we estimate the strength of balancing selection acting on these genes before, during, and after the bottleneck. In addition, we estimate the predicted future loss of genetic variation at the TLRs, i.e. the 'drift debt', which is likely to occur until the Seychelles warbler has reached its new mutation-drift-selection equilibrium state. After using this new and considerably more robust statistical analysis with more realistic demographic assumptions to detect selection, we then discuss the applicability of such an approach and the insights it provides in relation to the forces acting to shape present and future patterns of genetic variation within bottlenecked populations in general.

Materials and Methods

Molecular methods

Blood samples (ca 25 µl) are taken at each catch via brachial venipuncture, placed in absolute ethanol in a 2 ml screw-top Eppendorf tube and kept in the fridge at 4^{o}C. The blood samples used in the present study were from randomly-selected adult birds (> 1 year old) chosen at random from the contemporary 2000-2008 population. Genomic DNA was extracted using a salt extraction method (Richardson *et al.* 2001) and sex was confirmed using a molecular sexing protocol (Griffiths *et al.* 1998). The following TLR loci were amplified: *TLR1LA, TLR1LB, TLR3, TLR4, TLR5, TLR15* and *TLR21,* in 22-30 individuals, as detailed in Chapter 3.

Forward-in-Time Computer Simulations

A model was built to simulate the loss of genetic variation at an autosomal locus under balancing selection (symmetric overdominance) in a diploid population that experienced a bottleneck of known size and duration. A synthetic nucleotide sequence representing several different TLR loci with a known number of base-pairs was simulated. This locus was first allowed to accumulate polymorphisms in an ancestral population of a given *Ne* until the genetic diversity (expressed as the effective number of haplotypes) in this population had reached a mutation-drift-selection equilibrium (Fig 1). We assumed an effective population size (*Ne*) of 6900 based on work done by Spurgin *et al.* (2014), where *Ne* was estimated by analysis of microsatellite data from samples taken from both the pre- (museum samples) and post-bottleneck (contemporary) population of the Seychelles warbler. We also explored the minimum and maximum estimates (*Ne* = 2600 and 9700). We assumed a constant mutation rate equal to $\mu = 10^{-9}$ (Kumar & Subramanian 2002). Furthermore, we modelled balancing selection across a narrow range of selection coefficients (S), based on preliminary findings (Chapter 3) of $0 < S < 0.1$. We simulated a selection coefficient that was constant over time within each model. We compared the results of these simulations to the loss in genetic variation in simulations without selection acting during or after the bottleneck.

In our simulations, once the ancestral population had reached equilibrium it underwent a bottleneck consisting of 22 generations of *Ne* = 50, followed by three generations of population expansion with *Ne* = 100, 150 and 200, and finishing with nine generations at *Ne* = 250 (Spurgin *et al.* 2014). The bottleneck scenario was started at different (random) points in time after the burn-in had completed and after the ancestral

population had reached equilibrium. Finally, a subsample equal to the number of genotyped birds was randomly drawn from the simulated contemporary population, and haplotype diversity (H_{sim}) in this sub-sample was assessed. These precautions were taken to account for the random levels of stochasticity in genetic evolution over time (Fig S1 & S2).

Simulations were first run with a selection coefficient S = 0 to test whether the null model of neutral evolution can be rejected. If neutral evolution was rejected, different selection coefficients were then explored to examine the parameter space and determine the strength of selection that was most consistent with the observed heterozygosity in the post-bottleneck Seychelles warbler population. Therefore, we compared the simulated value (H_{sim}) to the observed heterozygosity (H_{obs}) in the contemporary Seychelles warbler population using a total of 1,000 independent sampling points to calculate the distribution of H_{sim} for each selection coefficient S of each locus.

Given that isolated and bottlenecked populations may lose a component of their parasite fauna (Bergstrom *et al.* 1999), we ran a second set of simulations to examine the effect of reducing PMS on TLR variation. We focused on *TLR15,* the most polymorphic locus in the Seychelles warbler, and simulated neutral evolution during and after the bottleneck. Observing a continued decline in gene diversity, we then performed forecast modelling to predict the future loss of genetic variation at TLR loci in the Seychelles warbler assuming no PMS. This part of the study was performed to quantify the 'drift debt' by analysing the loss of genetic polymorphism still required to reach the novel mutation-drift-selection equilibrium value. In these simulations, we assume a 'no-change' scenario regarding the future demography and population size of the Seychelles warbler. The model predicted the amount of genetic variation at *TLR15* in 2050 and 2100. We assumed an average four-year generation time and an effective population size *Ne* = 250 (Wright 2014). We simulated the future loss of genetic variation caused by drift with and without balancing selection (S = 0.00 and S = 0.03), and a mutation rate $\mu = 10^{-9}$ (Kumar & Subramanian 2002).

Results

Simulations show that an equilibrium is reached after 100 000 generations with $\theta = 2.76 \times 10^5$ (which is equivalent to *Ne* = 6900 and $\mu = 10^9$), and that H_{sim} reaches a plateau (Fig 1).

Table 1. Polymorphism statistics for TLR loci in the Seychelles warbler. Adapted from Chapter 3.

Locus	Number of sequences	Fragment size (bp)	H	S	Heterozygosity	dN / dS
TLR1LA	44	531	2	1	0.35	0 / 1
TLR1LB	66	750	4	2	0.63	0 / 2
TLR3	58	801	5	3	0.53	2 / 1
TLR4	60	648	1	0	0.00	0 / 0
TLR5	48	741	2	1	0.12	1 / 0
TLR15	60	528	5	3	0.68	3 / 0
TLR21	60	453	1	0	0.00	0 / 0

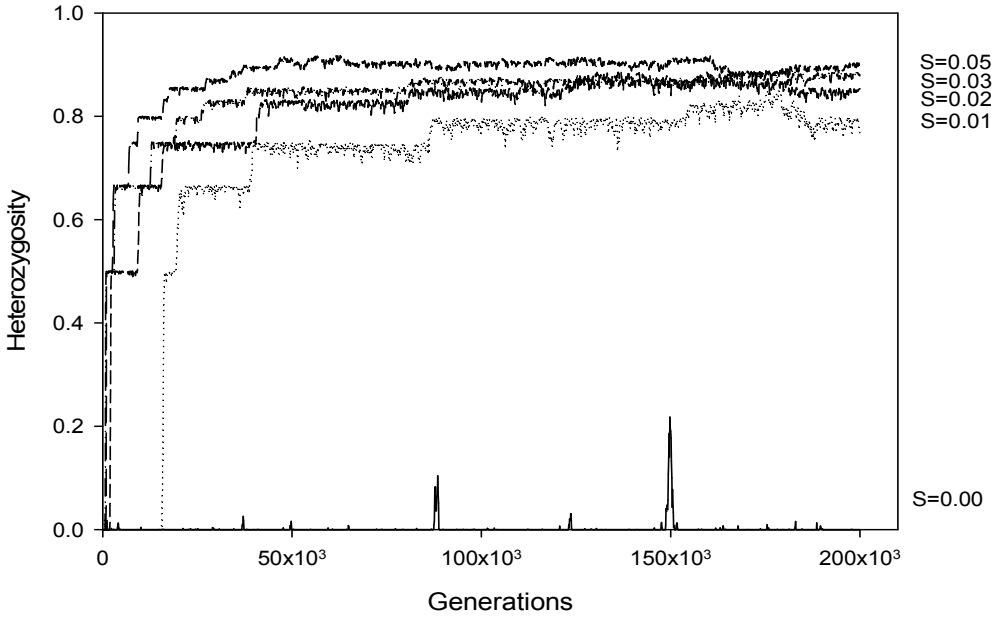


Figure 1. Simulated heterozygosity of a gene subject to overdominance selection and which consists of 528 base pairs in a population with an effective population size Ne = 6900 and a mutation rate μ =10-9 across a range of selection coefficients (S = 0.00, 0.01, 0.02, 0.03 and 0.05). The equilibrium heterozygosity is reached after ca 100 000 generations.

The figure shows that the equilibrium values of H_{sim} increase with an increased coefficient of balancing selection (S). When simulating neutral evolution, the mutation-drift equilibrium value of haplotype diversity is as close to, if not, zero ($H_{sim} \approx 0$). This contrasts with the fact that the post-bottleneck population of the Seychelles warbler showed significant levels of polymorphism at some TLR loci, with five haplotypes observed at one locus.

As predicted, the simulation population bottleneck rapidly leads to the erosion of variation over the 34 generations as a consequence of increased genetic drift. Despite this in our observed data, five out of seven TLR loci remained polymorphic in the Seychelles warbler population. According to our models, moderate- weak balancing selection ($0.005 \leq S \leq 0.03$) is likely to act on *TLRL1A, TLR1LB, TLR3, TLR5* and *TLR15,* which would explain the observed level of polymorphism at these five loci. In contrast, the two monomorphic loci (*TLR4* and *TLR21*) appear to evolve neutrally (S=0) according to the simulations (Table 2, Fig 2).

Table 2. Selection coefficient (S) estimates based on plotted simulated TLR heterozygosity (H) following specific demographic scenario after 100 000 generations under a constant selective pressure. P-values indicate any significant difference between observed and simulated *H*.

Locus	S based on *H*	S based on number of haplotypes	Difference between *H* and number of haplotypes	Significance (*P)*
TLR1LA	0.01	0.01	0.00	N/A
TLR1LB	0.03	0.05	0.02	> 0.1
TLR3	0.02	0.08	0.06	> 0.1
TLR4	0.00	0.00	0.00	N/A
TLR5	0.01	0.01	0.00	N/A
TLR15	0.03	0.09	0.06	> 0.1
TLR21	0.00	0.00	0.00	N/A

Our estimates of the strength of selection based on H_{sim} and H_{obs} consistently reject neutral evolution for five out of the seven TLR loci (Fig 2). The *Ne* of the ancestral population also does not appear to have a significant effect on the overall conclusions (Fig S3). However, if we assume a 10 times higher mutation rate ($\mu = 10^{-8}$), only the two most diverse TLR loci (*TLR1LB* and *TLR15*) have a level of genetic polymorphism that is inconsistent with neutral evolution (Fig S4).

Next we assumed that there would be no PMS during and after the bottleneck so that the immunogenetic variation would be subject only to drift and not subject to balancing selection. We simulated *TLR15*, the most polymorphic TLR locus in our study, and found that the initial decline of heterozygosity is similar between the scenario with and without balancing selection (S = 0.03 and S = 0.00 respectively) (Fig 3). However, without balancing selection, the gene diversity continues to decline into the future.

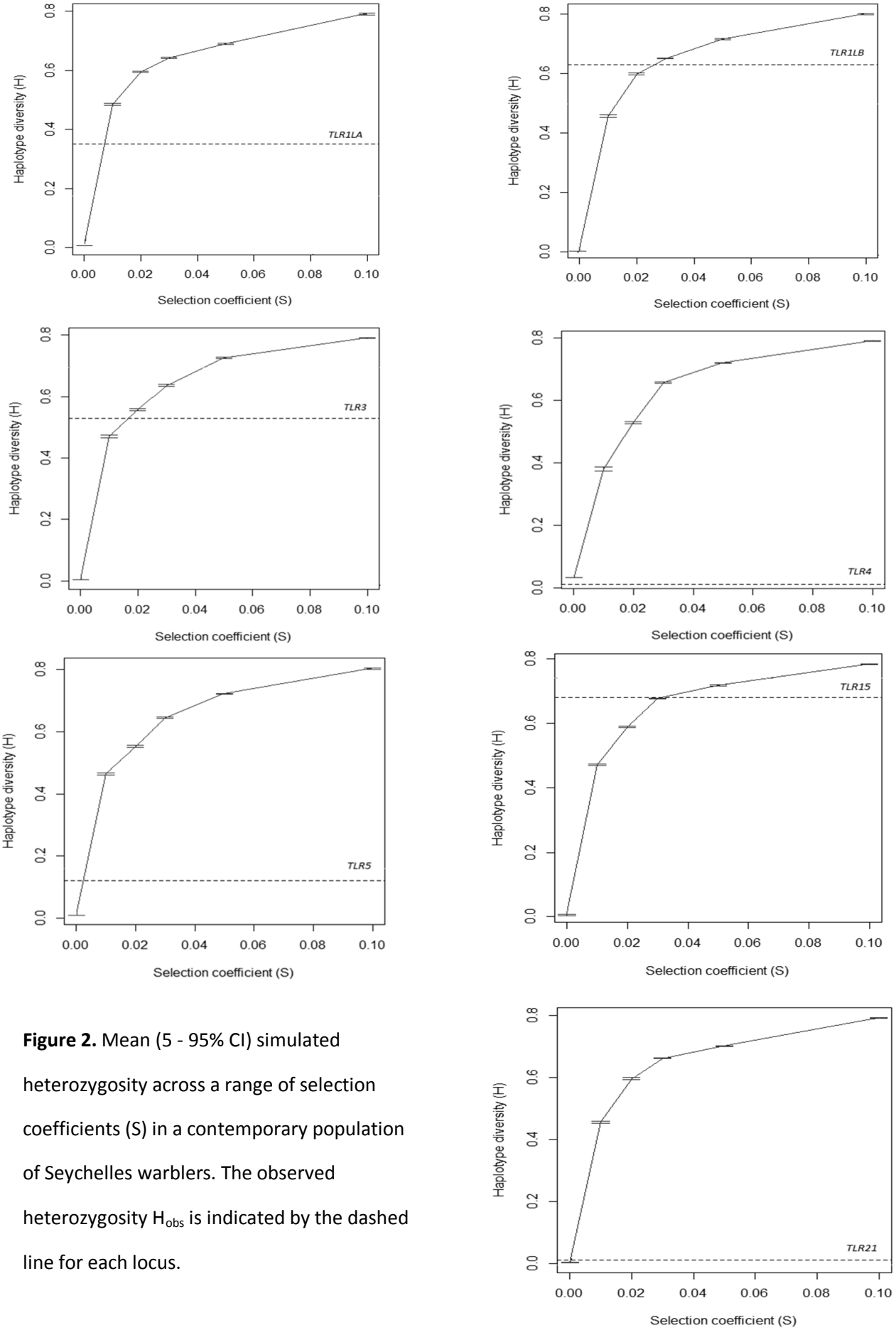


Figure 2. Mean (5 - 95% CI) simulated heterozygosity across a range of selection coefficients (S) in a contemporary population of Seychelles warblers. The observed heterozygosity H_{obs} is indicated by the dashed line for each locus.

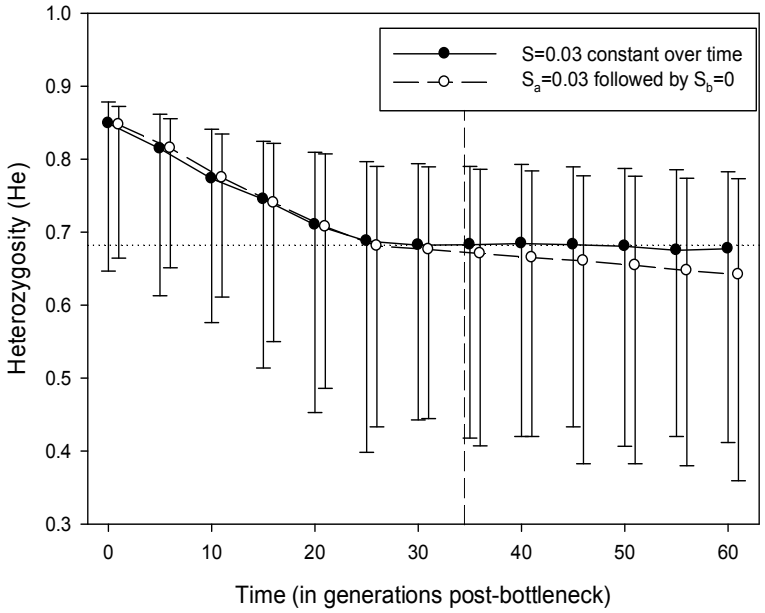


Figure 3. Heterozygosity in simulations with a constant balancing selection S = 0.03 (solid) and balancing selection in the ancestral population (Sa = 0.03) followed by neutral evolution during / after the bottleneck (Sb = 0) (open symbols). Dotted horizontal line indicate actual heterozygosity (*He* = 0.682) observed in the population sample and the contemporary sample was collected at generation 34 (dashed vertical line). In the absence of contemporary balancing selection (open symbols), the population continues to lose genetic diversity due to the 'drift debt'.

To test this further and quantify the 'drift debt', we conducted forecast modelling to predict the amount of genetic variation at *TLR15* that will remain in the Seychelles warbler population over the next decades if balancing selection has ceased to operate. Simulations show that in this scenario, genetic variation continues to decline by a further 2.2% and 4.5% by 2050 and 2100 respectively, even with the future Seychelles warbler population size remaining constant at present day levels (Fig S5). In contrast, if PMS continues to act at a constant intensity (S = 0.03), the TLR variation is not expected to decline any further.

Discussion

A previous study detected genetic polymorphisms in five out of seven TLR loci in the post-bottlenecked population of the Seychelles (Chapter 3) despite the generally low levels of genome-wide variation observed in this species (Richardson & Westerdahl 2003; Hansson &

Richardson 2005). However, even though one of the loci (*TLR15*) carried four allelic variants, population genetic statistics were unable to detect evidence of balancing selection at any of the loci. In order to analyse the evolution of TLRs in this bottlenecked population in more detail, we developed forward-in-time computer simulations to test whether balancing selection has been (and / or still is) acting on these genes. The model was parameterised based on data from a previous study of this species (Spurgin *et al.* 2014), which used microsatellite data from contemporary and historic samples to determine the extent and duration of the bottleneck that this species endured. Our simulations suggest that balancing selection has been acting on five out of seven TLR genes. The strength of selection inferred in the Seychelles warbler differs between TLR loci but is generally relatively weak ($S \leq 0.03$) compared to what has been found for TLRs in other avian species (Alcaide & Edwards 2011; Grueber *et al.* 2012). *TLR15* is the most polymorphic TLR locus in the Seychelles warbler, despite a considerably low S-value ($S = 0.03$). A similar conclusion was drawn in previous studies on several other phylogenetically-distant avian species (Alcaide *et al.* 2007; Brownlie & Allan 2011; Boyd *et al.* 2012). The house finch, *Carpodacus mexicanus*, shows high levels of polymorphism with at least 16 alleles at the *TLR15* locus, which probably reflects the species' large effective population size as well as the effect of balancing selection (Alcaide & Edwards 2011). More similar to the Seychelles warbler is the New Zealand Stewart Island robin, *Petroica australis raikura*, another island endemic that has undergone a recent bottleneck. In this species *TLR15* was found to possess two functional variants and was inferred to be under balancing selection (Grueber *et al.* 2012).

Simulations showed that the initial rate of decline in heterozygosity during the bottleneck is almost identical in scenarios with and without balancing selection. This conclusion is consistent with previous studies that have suggested that drift overrides the effect of balancing selection during bottlenecks (e.g. Willi *et al.* 2007; Bollmer *et al.* 2011; Strand *et al.* 2012). A forecast model showed that genetic variation might continue to decline due to a 'drift debt' in which the post-bottlenecked population reaches a new and considerably lower equilibrium level of polymorphism. However, such continued erosion will only take place if pathogen-mediated selection (PMS) has ceased to operate after the bottleneck (Fig 3). If the selection coefficient of PMS in the post-bottlenecked Seychelles warbler population remains similar to that in the ancestral population, TLR variation in the

contemporary gene pool will have reached its new equilibrium value and is not expected to decline any further.

The intensity of PMS operating in the Seychelles warbler population is difficult to establish. Even though our computer simulations show that the TLR polymorphism observed in the Seychelles warbler is consistent with balancing selection, it was impossible to discern whether selection was operating during the bottleneck. The rate of decline in genetic diversity is similar in simulated scenarios with and without balancing selection (S=0.03 and S=0, respectively), which shows that genetic drift tends to override the effect of balancing selection during population bottlenecks (see also Alcaide 2010; Grueber *et al.* 2013). Nevertheless, it is important to understand whether or not balancing selection continues to operate because this will determine the amount of genetic variation that remains in the population in the future. Our simulations show that TLR variation in the Seychelles warbler is expected to continue to decline if PMS has ceased to operate, whereas the genepool will have already reached a new equilibrium level of variation if PMS remains at the level inferred.

The contemporary population of the Seychelles warbler is very pathogen depauperate. No gastro-intestinal parasites have ever been detected in this population despite considerable efforts to screen a large number of birds sampled at different times (Hutchings 2009). Furthermore, there is no evidence of detrimental bacterial, viral or fungal infections in this species. Indeed only one type of blood pathogen has ever been detected; the GRW1 strain of avian malaria (Hutchings 2009). Given this low pathogen diversity in the Seychelles warbler, contemporary balancing selection might play little or no role in maintaining TLR polymorphism in this species. However, it is important to predict how the population would respond to novel introduced pathogens. In order to thoroughly explore whether PMS is still operating on TLR variation, we will need to examine the link between individual TLR characteristics and fitness in the contemporary Seychelles warbler population.

It is interesting to consider why the previous assessment of TLR sequence variation in the Seychelles warbler failed to detect positive selection. There has been much criticism on the relatively poor power underlying the use of genetic markers in molecular ecology

(Waples & Gaggiotti 2006; Vasemagi & Primmer 2005; Sutton *et al.* 2011). Indeed, sequence-based tests of selection come with several caveats such as low statistical power and restrictive assumptions (for review, see Ford 2002). Sharp changes in demography and population size, as well as the limited number of samples available for analysis, are issues that are particularly problematic in studies of endangered species. For this reason, forward-in-time simulations might be a better alternative to understand the evolutionary forces that have shaped genetic variation within endangered populations (see also Carvajal-Rodríguez 2010). In all likelihood, many studies will have concluded that selection has not been operating in their study species due to the insufficient statistical power of the most commonly used population genetic statistics.

A computer simulation approach also offers a further important advantage over population genetic statistics in that it enables researchers to estimate the future loss of genetic variation that may occur in endangered species. Such information allows conservation managers to make informed decisions by anticipating deleterious changes in genepools and strategically plan interventions such as genetic supplementation (Lynch & Hely 2001; van Oosterhout *et al.* 2007). Forecast modelling of the Seychelles warbler indicated that the genetic variation at *TLR15* might continue to decline depending on the presence or absence of PMS. Moreover, more recently bottlenecked populations than the Seychelles warbler are expected to show a continued decline in genetic variation even if the population has recovered and is demographically stable and even when PMS is operating. The reason for this is that the amount of genetic variation present in a recently bottlenecked population will still significantly exceed the level expected in a genepool that is in a mutation-drift-selection equilibrium. Analogous to the 'extinction debt' (Kuussaari *et al.* 2009), genetic variation is expected to be lost under a 'no change' scenario. We have referred to this as the 'drift debt', and we believe this is likely to affect many recently bottlenecked populations. We advocate the use of computer simulations in conservation biology to quantify the anticipated future decline in genetic variation in endangered species.

References

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

Akira S, Takeda K, Kaisho T (2001) Toll-like receptors: critical proteins linking innate and acquired immunity. *Nature immunology*, **2**, 675–680.

Alcaide M (2010) On the relative roles of selection and genetic drift in shaping MHC variation. *Molecular Ecology*, **19**, 3842–4.

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

Alcaide M, Edwards S V, Negro JJ (2007) Characterization, polymorphism, and evolution of MHC class II B genes in birds of prey. *Journal of Molecular Biology*, **65**, 541–554.

Areal H, Abrantes J, Esteves PJ (2011) Signatures of positive selection in Toll-like receptor (TLR) genes in mammals. *BMC evolutionary biology*, **11**, 368.

Basu M, Das T, Ghosh A *et al.* (2012) Gene-gene interaction and functional impact of polymorphisms on innate immune genes in controlling Plasmodium falciparum blood infection level. *PLoS One*, **7**, e46441.

Bergstrom CT, McElhany P, Real L a (1999) Transmission bottlenecks as determinants of virulence in rapidly evolving pathogens. *Proceedings of the National Academy of Sciences of the United States of America*, **96**, 5095–5100.

Bihl F, Salez L, Beaubier M *et al.* (2003) Overexpression of Toll-like receptor 4 amplifies the host response to lipopolysaccharide and provides a survival advantage in transgenic mice. *Journal of Immunology*, **170**, 6141–50.

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

Bonneaud C, Balenger SL, Russell AF *et al.* (2011) Rapid evolution of disease resistance is accompanied by functional changes in gene expression in a wild bird. *PNAS*, **108**, 7866–7871.

Boyd AC, Peroval MY, Hammond JA *et al.* (2012) TLR15 is unique to avian and reptilian lineages and recognizes a yeast-derived agonist. *Journal of Immunology*, **189**, 4930–4938.

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

Carvajal-rodríguez A (2010) Simulation of Genes and Genomes Forward in Time. *Current Genomics*, **11**, 58–61.

Collar NJ, Stuart SN (1985) *Threatened birds of Africa and related islands.* tional Council for Bird Preservation, Cambridge.

Deter J, Charbonnel N, Cosson J-F, Morand S (2007) Regulation of vole populations by the nematode Trichuris arvicolae: insights from modelling. *European Journal of Wildlife Research*, **54**, 60–70.

Ferrer-admetlla A, Bosch E, Sikora M *et al.* (2008) Balancing Selection Is the Main Force Shaping the Evolution of Innate Immunity. *Journal of Immunology*, **181**, 1315–1322.

Ford MJ (2002) Applications of selective neutrality tests to molecular ecology. *Molecular Ecology*, **11**, 1245–62.

Frankham R, Lees K, Montgomery ME *et al.* (1999) Do population size bottlenecks reduce evolutionary potential ? *Animal Conservation*, **2**, 255–260.

Griffiths R, Double MC, Orr K, Dawson RJG (1998) A DNA test to sex most birds. *Molecular Ecology*, **7**, 1071–1075.

Grueber CE, Wallis GP, Jamieson IG (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**, 4470–4482.

Grueber CE, Wallis GP, Jamieson IG (2014) Episodic positive selection in the evolution of avian toll-like receptor innate immunity genes. *PloS one*, **9**, e89632.

Grueber CE, Wallis GP, King TM, Jamieson IG (2012) Variation at innate immunity Toll-like receptor genes in a bottlenecked population of a New Zealand robin. *PLoS One*, **7**, e45011.

Hansson B, Richardson DS (2005) Genetic variation in two endangered Acrocephalus species compared to a widespread congener: estimates based on functional and random loci. *Animal Conservation*, **8**, 83–90.

Hellgren O, Ekblom R (2010) Evolution of a cluster of innate immune genes (beta-defensins) along the ancestral lines of chicken and zebra finch. *Immunome Research*, **6**, 3.

Hellgren O, Sheldon BC, Buckling A (2010) In vitro tests of natural allelic variation of innate immune genes (avian β-defensins) reveal functional differences in microbial inhibition. *Journal of Evolutionary Biology*, **23**, 2726–2730.

Hudson PJ (1986) The Effect of a Parasitic Nematode on the Breeding Production of Red Grouse. *Journal of Animal Ecology*, **55**, 85–92.

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

Jin MS, Kim SE, Heo JY *et al.* (2007) Crystal structure of the TLR1-TLR2 heterodimer induced by binding of a tri-acylated lipopeptide. *Cell*, **130**, 1071–1082.

Kalbe M, Eizaguirre C, Dankert I *et al.* (2009) Lifetime reproductive success is maximized with optimal major histocompatibility complex diversity. *Proceedings of the Royal Society B*, **276**, 925–934.

Kawai T, Akira S (2010) The role of pattern-recognition receptors in innate immunity: update on Toll-like receptors. *Nature Immunology*, **11**, 373–384.

Keestra M, de Zoete MR, Bouwman LI, van Putten JPM (2010) Chicken TLR21 is an innate CpG DNA receptor distinct from mammalian TLR9. *Journal of Immunology*, **185**, 460–7.

Kim HM, Park BS, Kim JI *et al.* (2007) Crystal structure of the TLR4-MD-2 complex with bound endotoxin antagonist Eritoran. *Cell*, **130**, 906–917.

Kumar S, Subramanian S (2002) Mutation rates in mammalian genomes. *Proceedings of the National Academy of Sciences of the United States of America*, **99**, 803–8.

Kuussaari M, Bommarco R, Heikkinen RK *et al.* (2009) Extinction debt: a challenge for biodiversity conservation. *Trends in Ecology and Evolution*, **24**, 564–571.

Kyle CJ, Rico Y, Castillo S *et al.* (2014) Spatial patterns of neutral and functional genetic variation reveal patterns of local adaptation in raccoon (Procyon lotor) populations exposed to raccoon rabies. *Molecular ecology*.

Lynch M, Hely MO (2001) Captive breeding and the genetic tness of natural populations. *Conservation Genetics*, **2**, 363–378.

Merino S, Moreno J, Sanz JJ, Arriero E (2000) Are avian blood parasites pathogenic in the wild? A medication experiment in blue tits (Parus caeruleus). *Proceedings of the Royal Society B*, **267**, 2507–10.

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

Moller AP, Saino N (2004) Immune response and survival. *Oikos*, **104**, 299–304.

Nakajima T, Ohtani H, Satta Y *et al.* (2008) Natural selection in the TLR-related genes in the course of primate evolution. *Immunogenetics*, **60**, 727–35.

Netea MG, Wijmenga C, O'Neill L a J (2012) Genetic variation in Toll-like receptors and disease susceptibility. *Nature Immunology*, **13**, 535–42.

Van Oosterhout C (2009) A new theory of MHC evolution: beyond selection on the immune genes. *Proceedings. Biological sciences / The Royal Society*, **276**, 657–65.

Van Oosterhout C, Smith AM, Hänfling B *et al.* (2007) The guppy as a conservation model: Implications of parasitism and inbreeding for reintroduction success. *Conservation Biology*, **21**, 1573–1583.

Palti Y (2011) Toll-like receptors in bony fish: from genomics to function. *Developmental and Comparative Immunology*, **35**, 1263–72.

Pedersen AB, Greives TJ (2008) The interaction of parasites and resources cause crashes in a wild mouse population. *Journal of Animal Ecology*, **77**, 370–377.

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

La Puente JM, Merino S, Tomás G *et al.* (2010) The blood parasite Haemoproteus reduces survival in a wild bird: a medication experiment. *Biology Letters*, **6**, 663–5.

Radwan J, Zagalska-Neubauer M, Cichon M *et al.* (2012) MHC diversity, malaria and lifetime reproductive success in collared flycatchers. *Molecular Ecology*, **21**, 2469–2479.

Ramírez-Soriano A, Ramos-Onsins SE, Rozas J, Calafell F, Navarro A (2008) Statistical power analysis of neutrality tests under demographic expansions, contractions and bottlenecks with recombination. *Genetics*, **179**, 555–67.

Redpath SM, Mougeot F, Leckie FM, Elston D a, Hudson PJ (2006) Testing the role of parasites in driving the cyclic population dynamics of a gamebird. *Ecology Letters*, **9**, 410–418.

Richardson DS, Jury FL, Blaakmeer K, Komdeur J, Burke T (2001) Parentage assignment and extra-group paternity in a cooperative breeder: the Seychelles warbler (Acrocephalus sechellensis). *Molecular Ecology*, **10**, 2263–73.

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

Sol D, Jovani R, Torres J (2003) Parasite mediated mortality and host immune response explain age-related differences in blood parasitism in birds. *Oecologia*, **135**, 542–7.

Sorci G, Moller a. P (1997) Comparative evidence for a positive correlation between haematozoan prevalence and mortality in waterfowl. *Journal of Evolutionary Biology*, **10**, 731–741.

Spurgin LG, van Oosterhout C, Illera JC *et al.* (2011) Gene conversion rapidly generates major histocompatibility complex diversity in recently founded bird populations. *Molecular Ecology*, **20**, 5213–5225.

Spurgin LG, Wright DJ, van der Velde M *et al.* (2014) Museum DNA reveals the demographic history of the endangered Seychelles warbler. *Evolutionary Applications*, 1–10.

Staley M, Bonneaud C (2015) Immune responses of wild birds to emerging infectious diseases. *Parasite Immunology*, **37**, 242–254.

Strand TM, Segelbacher G, Quintela M *et al.* (2012) Can balancing selection on MHC loci counteract genetic drift in small fragmented populations of black grouse? *Ecology and Evolution*, **2**, 341–353.

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

Takeda K, Akira S (2005) Toll-like receptors in innate immunity. *International Immunology*, **17**, 1–14.

Takeuchi O, Sato S, Horiuchi T *et al.* (2002) Cutting Edge: Role of Toll-Like Receptor 1 in Mediating Immune Response to Microbial Lipoproteins. *The Journal of Immunology*, **169**, 10–14.

Trowsdale J, Parham P (2004) Defense strategies and immune-related genes. *European Journal of Immunology*, **34**, 7–17.

Turner AK, Begon M, Jackson JA, Paterson S (2012) Evidence for selection at cytokine loci in a natural population of field voles (Microtus agrestis). *Molecular Ecology*, **21**, 1632–1646.

Vasemagi A, Primmer CR (2005) Challenges for identifying functionally important genetic variation: the promise of combining complementary research strategies. *Molecular Ecology*, **14**, 3623–3642.

Waples RS, Gaggiotti O (2006) What is a population? An empirical evaluation of some genetic methods for identifying the number of gene pools and their degree of connectivity. *Molecular Ecology*, **15**, 1419–1439.

Werling D, Jungi TW (2003) Toll-like receptors linking innate and adaptive immune response. *Veterinary Immunology and Immunopathology*, **91**, 1–12.

Willi Y, Van Buskirk J, Schmid B, Fischer M (2007) Genetic isolation of fragmented populations is exacerbated by drift and selection. *Journal of Evolutionary Biology*, **20**, 534–42.

Wright DJ (2014) Evolutionary and conservation genetics of the Seychelles warbler (Acrocephalus sechellensis). *Thesis*.

Yoneyama M, Fujita T (2010) Recognition of viral nucleic acids in innate immunity. *Reviews in Medical Virology*, **20**, 4–22.

Supplementary material

Figure S1. Pilot simulations to show that due to large stochasticity in number of haplotypes (*H*) in the ancestral population, the post-bottleneck population can start from very different levels of diversity by chance and thus, all post-bottleneck replicates are affected. Runs **i)** and **ii)** are at the identical settings of *Ne*= 100, $\mu = 10^{-5}$ for t= 5000 generations, followed by a bottleneck of N=100 for t= 100 generations. Runs **iii)** and **iv)** are at the identical settings of *Ne*=1000, $\mu = 10^{-6}$ for t= 5000 generations, followed by a bottleneck of N=1000 for t= 100 generations. These repeat runs with identical conditions reflect a large degree of variance due to 'chance' over evolutionary time.

Figure S2. Pilot simulations to show that by taking an average of multiple sample from the gene pool in the ancestral population, the post-bottleneck population now starts at more similar levels of diversity and thus, outputs are now similar for different *Ne* / μ combinations, which are numerically the same for runs **i)** to **iii)**. Run **i)** is at *Ne*= 10 000 and $\mu = 10^{-7}$, run **ii)** is at *Ne*= 1000 and $\mu = 10^{-6}$, and run **iii)** is at *Ne*= 100 and $\mu = 10^{-5}$. These repeat runs with identical conditions now have considerably less variance in *H,* given the new sampling methods written into the simulation instructions.

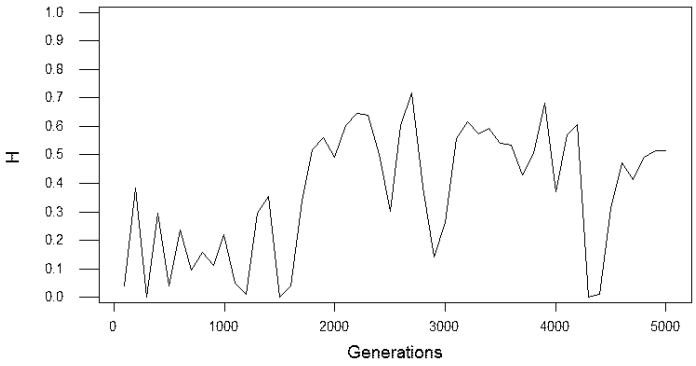
Figure S3. Upper and lower bounds of effective population size (*Ne*) to show the sensitivity of this parameter in detecting selection (S) within TLR loci in a simulated bottlenecked population of Seychelles warblers, based on TLR haplotype diversity (H_{sim} and H_{obs}).

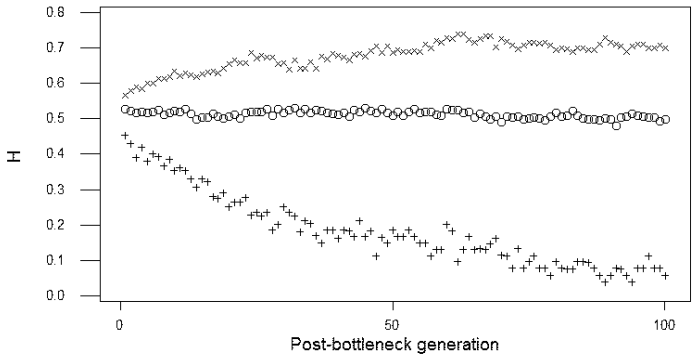
Figure S4. Upper and lower bounds of mutation rate (μ) to show the sensitivity of this parameter in detecting selection (S) within TLR loci in a simulated bottlenecked population of Seychelles warblers, based on TLR haplotype diversity (H_{sim} and H_{obs}).

Figure S5. Estimating selection coefficients (S) in the contemporary population of Seychelles warbler based on TLR haplotype diversity observed when selection is applied to the population before a bottleneck, but kept at S=0 both during and after the bottleneck. Parameters include: *Ne* (260, 690, 970), μ (10^{-7}, 10^{-8}, 10^{-9}) and 'bottle' to indicate the simulations ran where S only applies before the bottleneck, and is set at zero for during and after the bottleneck.

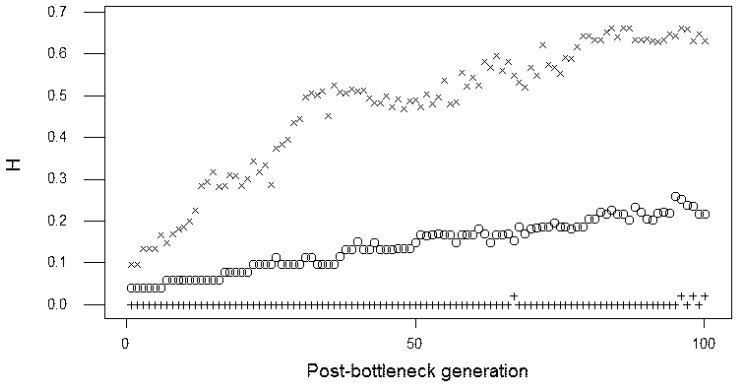
Figure S1.

i)

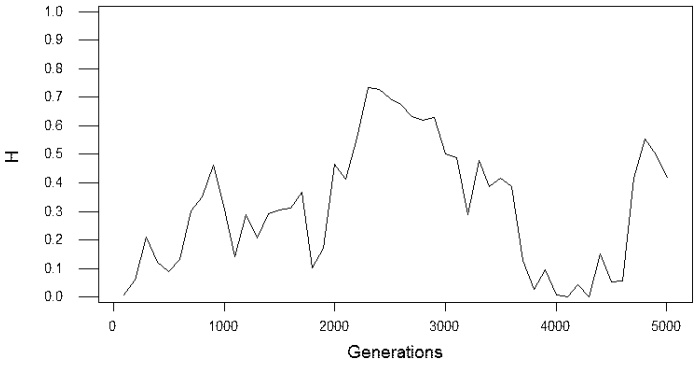


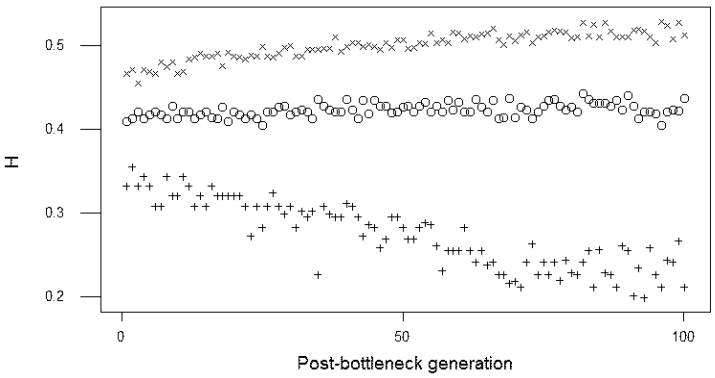


ii)

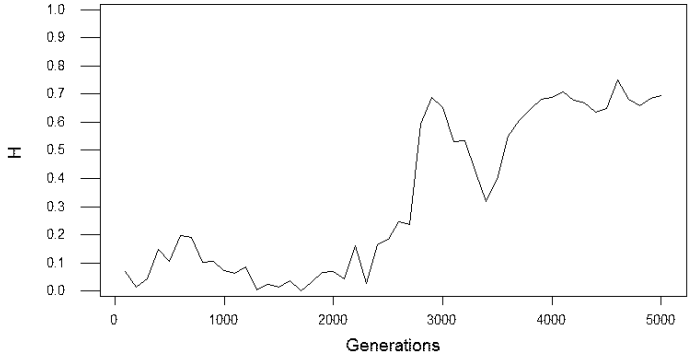


iii)





iv)



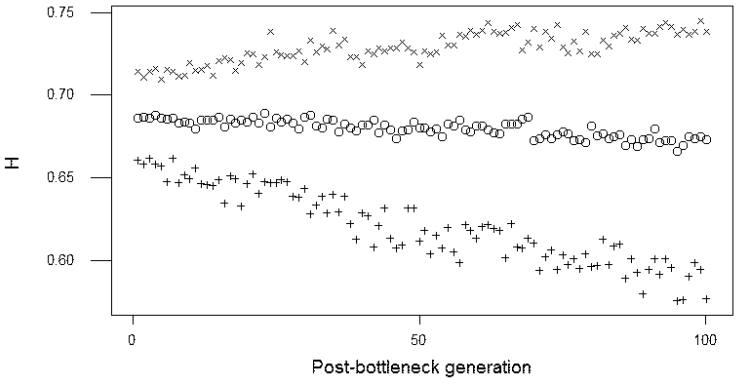
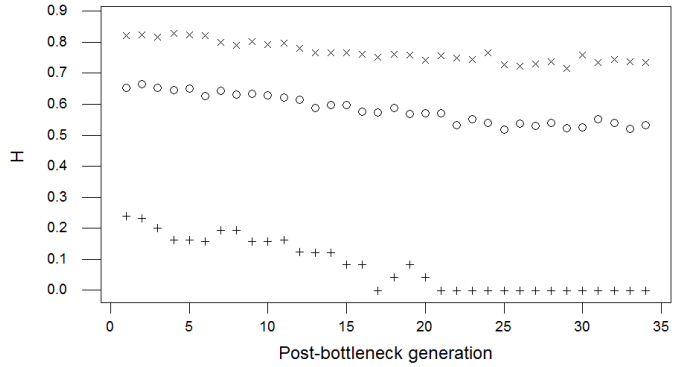
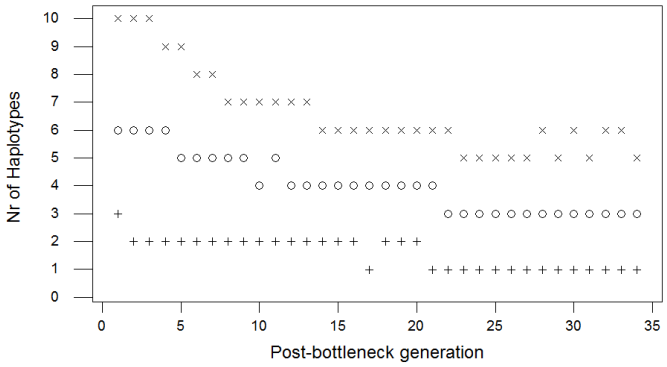


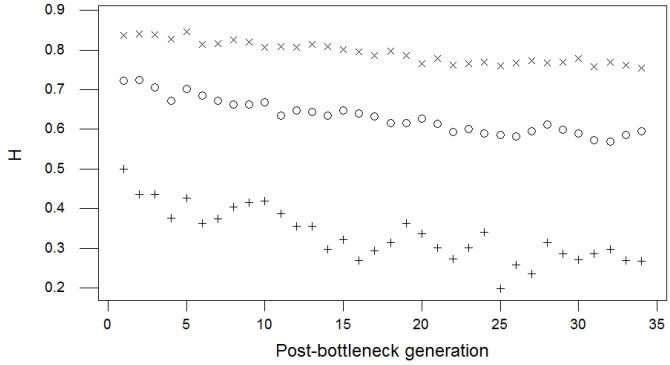
Figure S2.

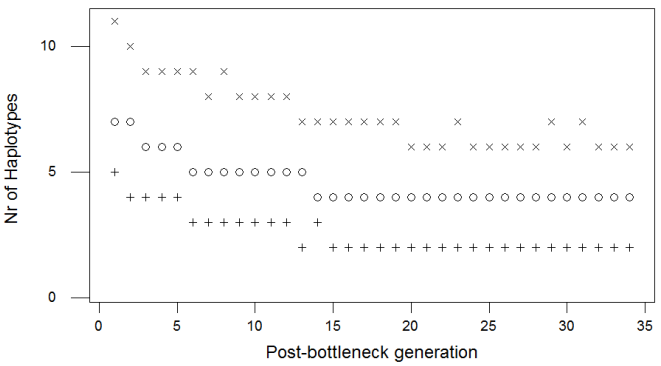
i)



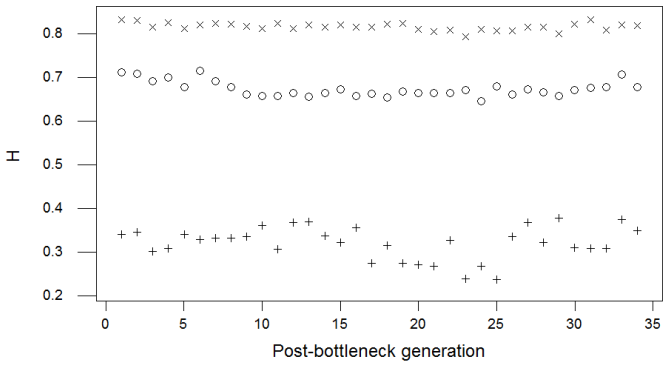


ii)





iii)



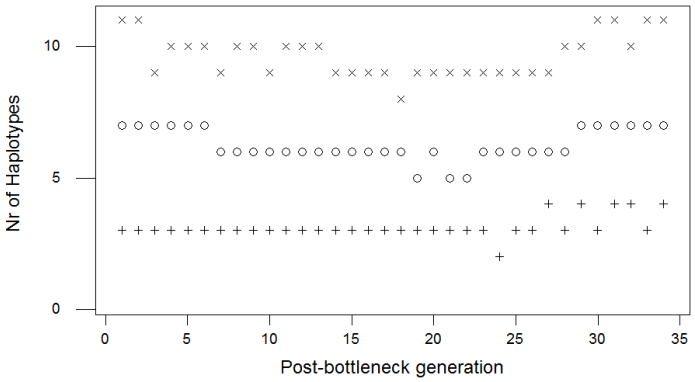
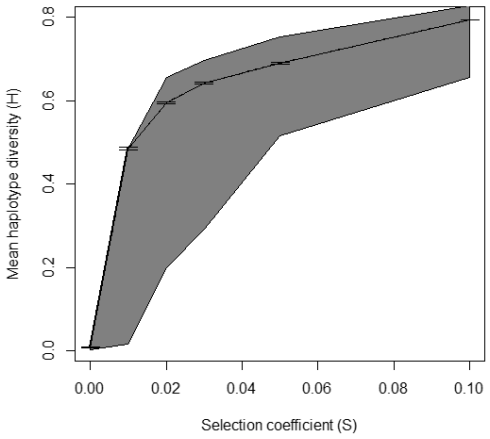
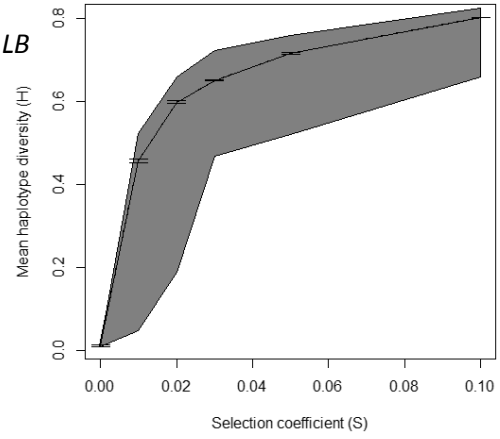


Figure S3.

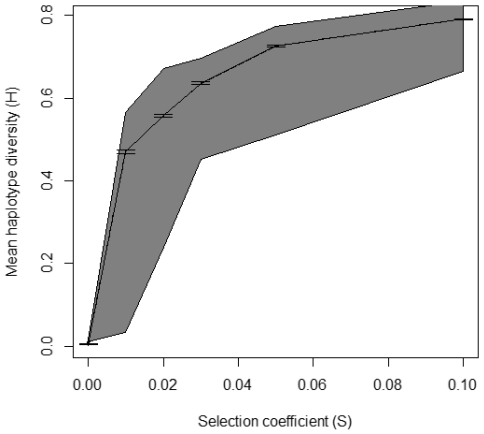
TLR1LA



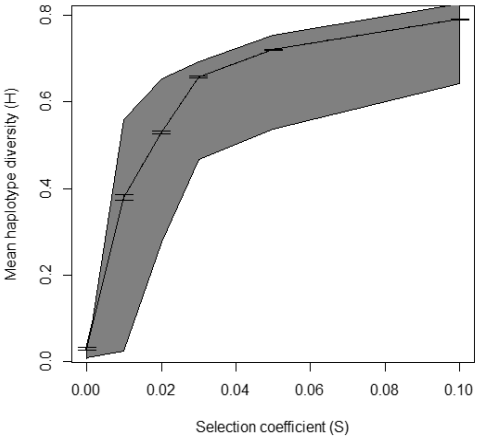
TLR1LB



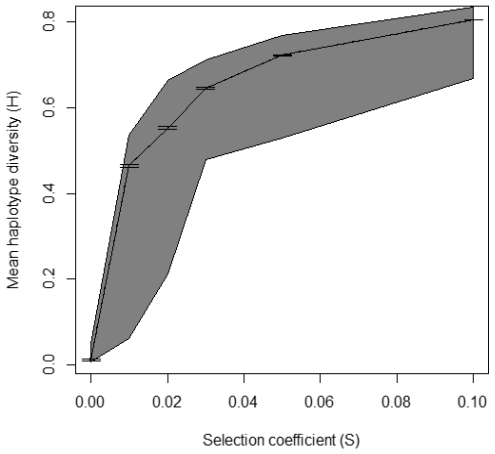
TLR3



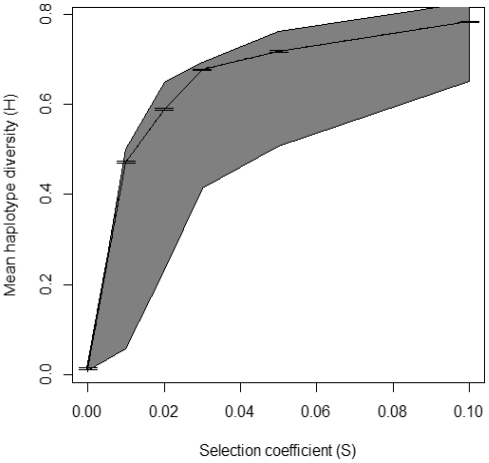
TLR4



TLR5



TLR15



TLR21

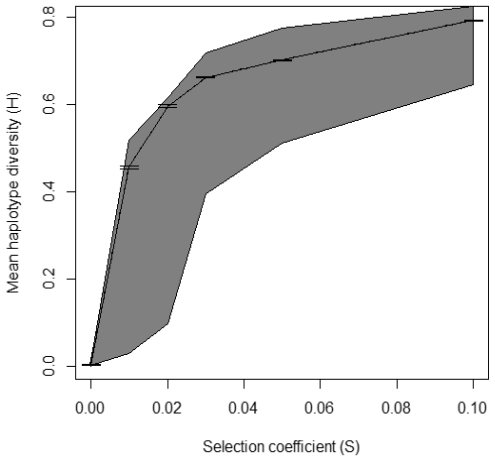
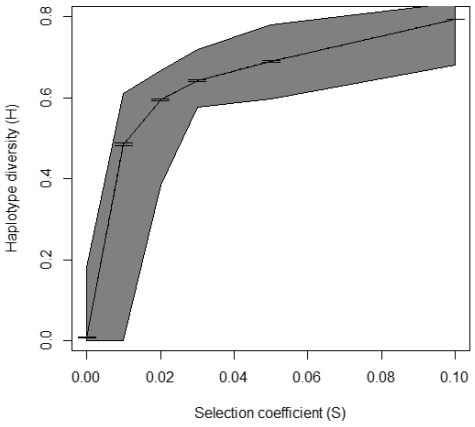
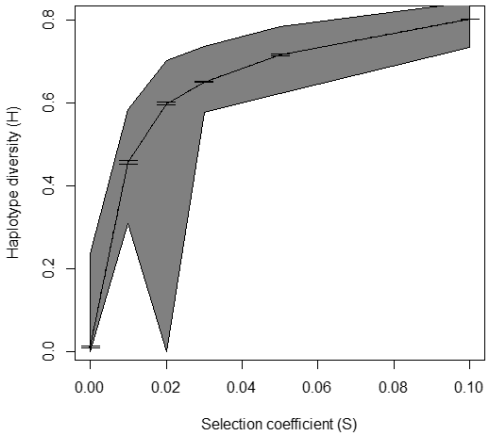


Figure S4.

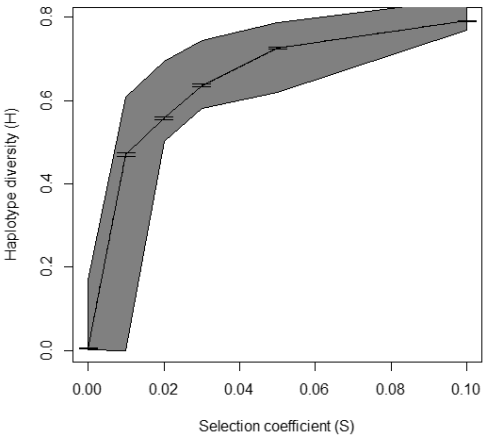
TLR1LA



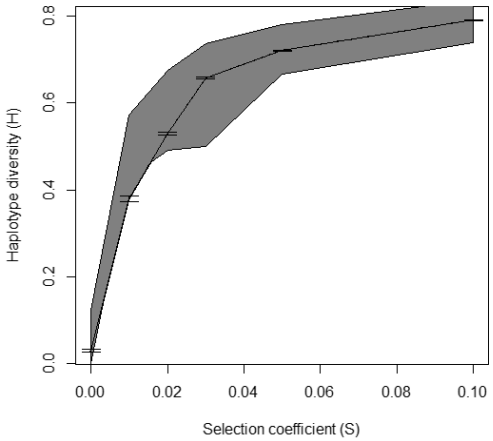
TLR1LB



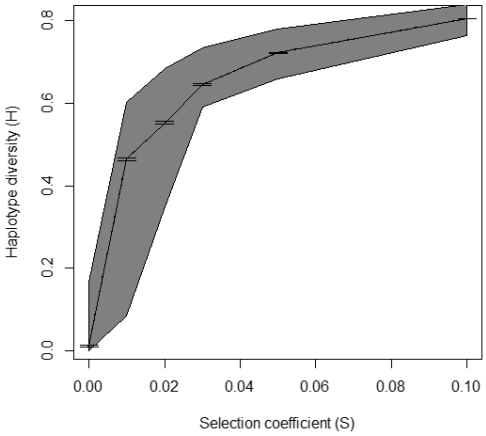
TLR3



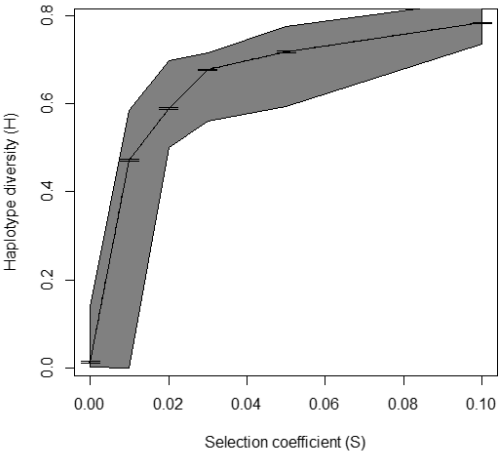
TLR4



TLR5



TLR15



TLR21

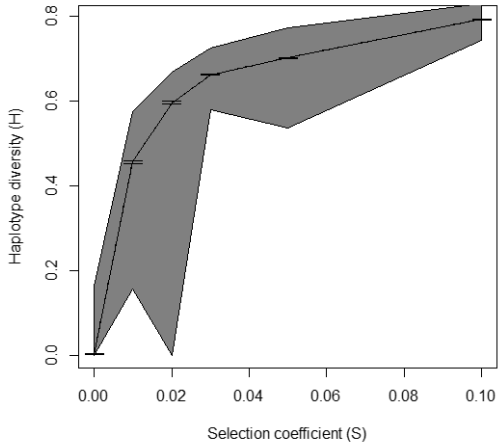
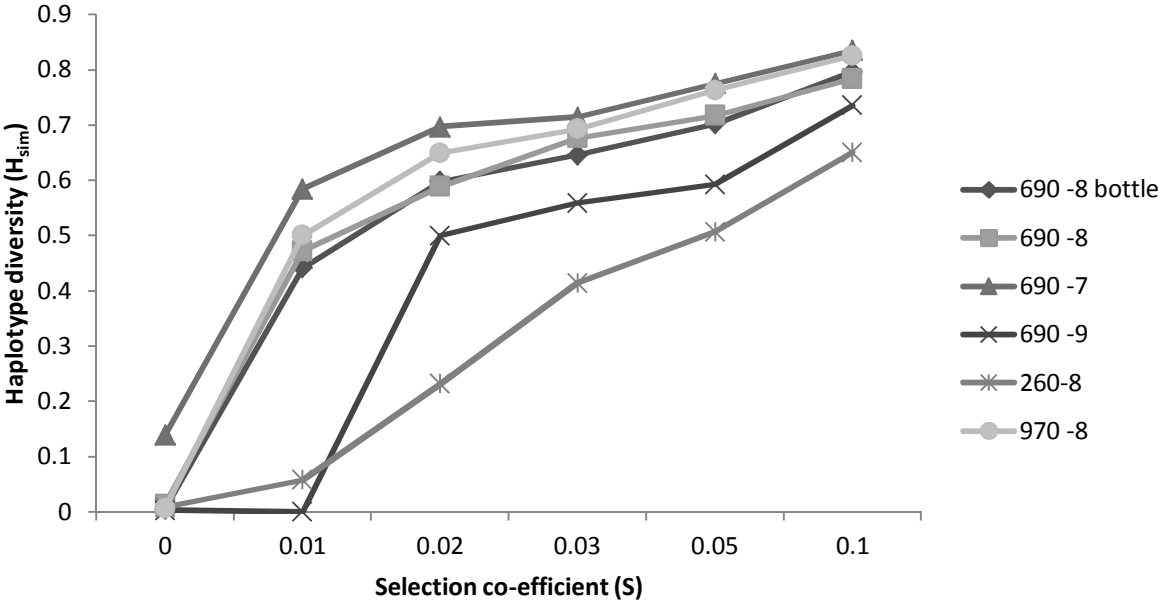


Figure S5.


0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
Haplotype diversity (H_{sim})
0
0.01
0.02
0.03
0.05
0.1
Selection co-efficient (S)
690 -8 bottle
690 -8
690 -7
690 -9
260-8
970 -8

Chapter 5: The effect of Immunogenetic variation at *TLR15*, on individual malaria infection and survival in the Seychelles warbler

Abstract

Pathogens can exercise strong selection on hosts, which makes them key drivers of evolutionary and demographic processes in natural populations. As a result of pathogen-mediated selection, specific immune genes may show relatively elevated levels of variation in small bottlenecked populations that have lost genome-wide variation due to drift. The Seychelles warbler *Acrocephalus sechellensis* is an endemic passerine that underwent a recent bottleneck and resulting loss of genome-wide variation. However, five toll-like receptor genes remain polymorphic in the population, with *TLR15* showing the highest level of functional variation and some signatures of positive selection. This study examines the association between individual *TLR15* variation and disease status and survival in a cohort of birds followed throughout their entire lives. We find that individuals with a specific allele at *TLR15* appear to have increased resilience to malarial infection, are more likely to survive into adulthood, and appear to have gained some immunity to subsequent reinfection. Furthermore our analysis also shows that resilience /resistance may be influenced by other previously screened immune loci, as MHC diversity predicted adult survival, and a specific MHC allele, *Ase-ua4,* influenced the ability to resist and / or tolerate malaria once infected.. Overall, this study suggests that complex interactions between Seychelles warbler hosts and malaria pathogens may result in the maintenance of *TLR15* variation in this population, even in the face of considerable drift.

Introduction

Pathogens (parasites and infectious diseases) exert strong selection pressures and thus are extremely important in shaping the evolution of their hosts, affecting everything from levels of genetic variation in individuals and populations, to behavioural and life history characteristics, and even organisational level process (Hudson 1986; Redpath *et al.* 2006; Deter *et al.* 2007; Pedersen & Greives 2008) such as demography, range and species persistence (for reviews, see Stahl & Bishop 2000; McDonald & Linde 2002; Barreiro & Quintana-Murci 2010). Given the impact that pathogens can have on their hosts' evolution it is important to understand the role of host genetic variation in combating pathogens. Many studies have shown that reduced genetic variation has negative effects on individual host fitness (e.g. Hedrick 2002; Biedrzycka & Radwan 2008; Lampert *et al.* 2009). Reduced genetic variation can also be detrimental at the population level, as demonstrated by the impact malarial parasites had on endemic Hawaiian passerine species (van Riper III *et al.* 1986; Atkinson *et al.* 1995; Fonesca *et al.* 2000)**.** This can lead to an increased extinction risk known as the extinction vortex (for reviews, see Newman & Pilson 1997; Spielman *et al.* 2004).

The level of polymorphism across the genome varies considerably due to a variety of evolutionary forces acting none-exclusively and underlying demography (Lande 1976, 1988). Selection can alter the level of genetic variation at any given locus from that observed on average across the genome, as a result of purifying (negative) or positive selection reducing variation compared to neutrally evolving regions (Kimura 1986; Ohta 2002; van Oosterhout 2009)**.** Rarely, loci may show an elevated level of polymorphism compared to the genome wide average as a result of balancing selection. Such elevated levels of variation occur as a results of a variety of potentially interacting mechanisms, including overdominance (Doherty & Zinkernagel 1975), rare allele advantage (Hill *et al.* 1991) and spatio-temporally fluctuating selection (Slade & McCallum 1992). These mechanism act to maintain genetic variation at a locus as a result of the superior fitness (at a given point in time or space) of individuals that either carry a specific allele or heterozygous combination of alleles (Hedrick 2006; Mitchell-Olds *et al.* 2007). Different selective agents can mediate balancing selection, although most commonly, balancing selection is associated with pathogens (for reviews, see Jeffery & Bangham 2000; Ford 2002; Bernatchez & Landry 2003).

TLRs are membrane-bound glycoproteins of the innate immune system that recognise distinctive pathogen-associated molecular patterns (PAMPs) (Jin & Lee 2008) and trigger an appropriate immune response depending on the pathogen-derived antigen they bind to and the class of that TLR molecule (Roach *et al.* 2005). *TLR15* is the most recently evolved of the TLR multigene family (Brownlie & Allan 2011) and importantly, it appears to be key to the recognition of intracellular parasites, including malarial parasites (Creagh & O'Neill 2006). It was first identified in the cecum of chickens when up-regulated in response to *Salmonella* infection (Higgs *et al.* 2006). The consequences of TLR activation are diverse and tailored to ensure efficient destruction and clearance of invading pathogens (Brownlie & Allan 2011). It has already been recently well-shown that TLRs in a range of vertebrates are under positive (balancing) selection (e.g. Nakajima *et al.* 2008; Areal *et al.* 2011; Grueber *et al.* 2014). Furthermore, selection maintaining variation at these immune genes can result in differential pathogen infection outcome (e.g. Bihl *et al.* 2003; Nerren *et al.* 2009; Boyd *et al.* 2012).

A good host-pathogen system is needed to explore the mechanisms of pathogen mediated balancing selection. Malaria is widely-studied because it is responsible for high impact diseases in humans, livestock and wildlife (Garnham 1980, for review, see Bordes & Morand 2015), including birds (e.g. Bonneaud *et al.* 2011; Fuller *et al.* 2012; Marzal *et al.* 2015). Studies have tested specific predictions about infection intensity and host fitness using avian-malaria models (e.g. Wood *et al.* 2007; Szollosi *et al.* 2011; Ferrer *et al.* 2014). For example, although few studies have evidenced fitness consequences of malaria infections in wild birds (e.g. Marzal *et al.* 2005; Radwan *et al.* 2012; Garamszegi *et al.* 2015), trade-offs between reproduction and defence against *Plasmodium* infections have been shown (Bonneaud *et al.* 2006; Lachish *et al.* 2011; Ferrer *et al.* 2014). However, the evidence of fitness consequences associated with haemoparasites infections in wild bird populations is so limited that these infections are often considered to be relatively benign (for review, see Escalante *et al.* 2004). To fully understand the impact of parasites such as avian malaria on their hosts we need precise information about the many biotic and abiotic factors that determine an individual exposure to malaria, and accurate estimates of subsequent fitness. Unfortunately such complete data on a wild host system is rare.

Here we use a combination of detailed life history, malarial infection and survival data collected over 18 years in the Seychelles warbler (*Acrocephalus sechellensis* - SW*)* population on Cousin to investigate the interactions between TLR variation and fitness in a wild living population. A previous study characterised variation in the toll-like receptor (TLR) gene group in the SW. The *TLR15* locus was the most polymorphic of the seven TLR loci screened, retaining four functional variants despite the recent bottleneck suffered by the SW and showing some signatures of positive selection at the codon level (chapter 4). In this study, we screened individual *TLR15* variation for a cohort of Seychelles warblers from Cousin Island and followed them throughout their lives, over a period of 18 years. We tested for associations between individual *TLR15* variation and malarial infection and survival (juvenile and adult). We controlled for other factors found to influence Seychelles warbler survival in previous studies including MHC diversity and the specific MHC allele *Ase-ua4* (Brouwer *et al.* 2010; Wright 2014), in addition to key ecological factors that may influence individual infection and survival.

Materials and Methods

Study species and sampling

The Seychelles warbler (SW) is a small (ca 12-15 g) insectivorous passerine bird endemic to the Seychelles islands (Safford & Hawkins 2013). Due to anthropogenic effects, by the 1960s the SW was reduced to just one population of ca 26 individuals remaining on the island of Cousin (Collar & Stuart 1985). As a result, the SWs effective population size was dramatically reduced from ca 7000 in the early 1800s to <50 in the contemporary population (Spurgin *et al.* 2014). However, with effective conservation management, the population recovered to its carrying capacity of ca 320 adults on the island by 1982 (Komdeur 1992) and has since remained relatively stable (Brouwer *et al.* 2009; Wright *et al.* 2014). The SW has shown to be an ideal study species for evolutionary, ecological and conservation question (e.g. Komdeur 1992; Richardson *et al.* 2003; van de Crommenacker *et al.* 2012; Wright *et al.* 2014). Since 1997, >96% of the Cousin population have been caught and ringed with a unique combination of colour rings and a metal British Trust for Ornithology (BTO) ring

(Richardson *et al.* 2002). Blood samples (ca 25 µl) are taken via brachial venipuncture, placed in absolute ethanol in a 2 ml screw-top Eppendorf tube and stored at 4^{o}C.

We focused on all SWs born in 1997-2002 (n = 205), that we monitored throughout their lives. Of these all but seven have since died (these seven are excluded from the survival analyses but included in the malarial analyses). An overall population census is carried out on Cousin bi-annually (for each field season) and individuals are intensively monitored with a re-sighting probability of juveniles (assigned as < 1 year of age in that study) at 0.87 ± 0.05 and adults at 0.92 ± 0.03 (Brouwer *et al.* 2006). If an individual is not seen for three consecutive seasons it is declared dead. There are two main breeding seasons for the Seychelles warbler which, for the purpose of this study, are termed 'summer' (April – October) and 'winter' (November – March). Birds are aged at first catch according to their eye-colour and behaviour and a number of age classes come under 'Juvenile', which is defined in this study as anything up to 10 months of age, and 'Adult' is once a bird is > 10 months old (Wright 2014). A 'Juvenile' can include all birds which are chicks, fledglings and old fledglings (all of these classes have grey eyes but different behaviours), and sub-adults (light-brown eyes, anything from 5-10 months old). Eye colour is subject to considerable natural variation and thus approximation is based on long-term data (for examples, see Komdeur 1991; Hammers *et al.* 2012). Upon first catch, a hatch date is estimated based on the ringing date and strict ageing guidelines. Total life-span of a bird can thus be estimated using the hatch date and the date the bird was last observed, though this is clearly a minimal estimate as the time between the date last seen and the actual date of death could vary between birds.

In addition to the population census done each season, bi-annual environmental surveys are taken across the island in order to measure both local and general ecological factors. Territory quality (TQ) is used as an index of food availability and is recorded as an insect count within a specified territory. In this study it applies to the territory of which a bird was born in. TQ also reflects abiotic factors such as weather given that temperature, humidity and rainfall all correlate with insect abundance (Komdeur 1992). Local density (the number of individual Seychelles warblers resident in a territory) is another important measure used here as it has been shown to influence Seychelles warbler survival (Brouwer *et al.* 2006). This previous study showed that while natal local density was not associated

with survival (natal group size), lifetime local density had a negative effect on survival. Birds living in larger groups had lower survival probabilities than those living in small groups. Lifetime local density is the average group size a bird lived in from its second year of life onwards (adulthood) and is defined in this study as 'local density' hereafter.

Molecular procedures

Genomic DNA was extracted using a salt-extraction method (Richardson *et al*. 2001). The sex of each bird was determined by polymerase chain reaction (PCR) as described in Griffiths *et al.* (1998). Only DNA samples that successfully amplified the sex-specific markers were subsequently screened for malaria. Avian malaria (*Haemoproteus* and *Plasmodium* species) was screened using the nested PCR method described by Waldenström *et al.* (2004). All samples were screened at least twice and only samples that amplified twice and were verified as malaria through Sanger sequencing were taken to be positive infections.

TLR15 variation in the Seychelles warbler

Specific primers were designed to target the variable region of exon 1 of *TLR15* previously identified in the SW (Chapter 3). The primers were designed in the program PerlPrimer v1.1.21 (Marshall 2004) The forward primer, TCTCCTGCAAATCCTTAGCC, has a melting temperature (Tm) of 59.94^{o}C and the reverse primer, CTGCTGTGTAGATGAAGTGG, has a Tm of 58.45^{o}C. Together these primers successfully amplified ca 420 bp around the variable region of interest.

PCRs were carried out in 10 µl volume with genomic DNA at a concentration of ca 10 ng/µl. Taq PCR Master Mix was used (Qiagen, UK) which includes: *Taq* DNA Polymerase, QIAGEN PCR Buffer, $MgCl_2$, and ultrapure dNTPs at optimised concentrations. PCRs were carried out using the following conditions: 30 s at 95^{o}C, 30 s at the locus-specific annealing temperature of 58^{o}C, 60 s at 72^{o}C, all repeated for 34 cycles. All PCRs started with an incubation step of 5 mins at 95^{o}C and finished with an incubation step of 5 mins at 72^{o}C. All PCR products were electrophoresed on a 2% agarose gel containing ethidium bromide to determine successful amplification of the expected size fragment. Successful samples were submitted to MWG Operon (Eurofins, Germany) for Sanger sequencing. All unique

sequences were confirmed by repeated sequencing across multiple individuals or (where identified in only one individual) multiple independent PCRs from that individual. Each chromatogram was examined by eye to identify single nucleotide polymorphisms (SNPs) in BioEdit (Hall 1999) via ClustalW codon alignment. Sequences with multiple SNPs had their haplotypes inferred using Bayesian PHASE algorithms (Stephens & Donnelly 2003) in the program DnaSP (Librado & Rozas 2009). However, given the relatively low level of polymorphism observed (six haplotypes), it is possible to manually assign haplotypes to each individual within the cohort based on the presence of given SNPs. Amino acid sequences were translated using Mega v5.1 (Tamura *et al.* 2011).

Tests for signatures of selection

All *TLR15* haplotype frequencies were tested for deviation from Hardy-Weinberg equilibrium using the Markov chain method available in Genepop v.2 (Raymond & Rousset 1995) for Fisher's test of exact probability (Guo & Thompson 1992). F_{IS} values are presented using Robertson and Hill's estimates (1984), which have lower variance under the null hypothesis compared to the alternative Weir and Cockerham's estimate (1984). Should there be a significant deviation, subsequent testing for heterozygote deficiency / excess was also carried out based on the *U*-test, which is more powerful than a probability-test (Raymond & Rousset 1995). Furthermore, individual genotypes that deviate from HW are presented using software HW-Quick check (Kalinowski 2006).

DnaSP was used to calculate basic measures of genetic variation including: number of sequences (N), overall number of segregating sites (S), number of haplotypes (H), haplotype diversity (Hd), nucleotide diversity (π) and ratio of synonymous (dS) to non-synonymous (dN) substitution. Z-tests of selection were carried out in Mega v5.1 (Tamura *et al.* 2007) in order to identify selection based on the average dN/dS across entire sequences (Kryazhimskiy & Plotkin 2008).

Site-specific tests were carried out using the HyPhy package available on DataMonkey (Delport *et al.* 2010). Different models were run to identify individual sites under selection based on dN/dS ratios at each codon. Two models were used: the mixed effects model of evolution that identifies sites under episodic diversifying selection (MEME)

(Murrell *et al.* 2012) and the fast unbiased Bayesian un-approximation model that identifies sites under putative diversifying and purifying selection (FUBAR) (Murrell *et al.* 2013). Both models had the default settings applied. This included a significance level of 0.1 for MEME to classify a site under selection, given that this method has been shown to be more conservative than empirical Bayesian approaches (Murrell *et al.* 2012). As the FUBAR model uses a Monte Carlo routine, it has a Bayes factor / posterior probability set at 0.9 as a minimum value for a site to be deemed under putative selection (Murrell *et al.* 2013).

Association analyses

For all linear model analyses, a model that including only fixed factors of primary interest and that maximised the sample size from our dataset was first tested. Any fixed factors deemed to be significant were then included in more complex models with fixed factors of secondary interest where some of these variables were not available for all 205 birds and thus the sample size was reduced. All variables were checked for significant correlation measures as a precaution beforehand. Microsatellite heterozygosity was included in all models as a control measure, given the results presented in 'Box 1' that there is no association between microsatellite heterozygosity and *TLR15* heterozygosity. This means that neutral variation is independent of variation at functional loci like TLRs and so must be considered when assessing the effects of genetic variation on individual fitness parameters. The variable 'microsatellite heterozygosity' represents standardised multi-locus heterozygosity across 30 neutral markers (Table S1).

Malaria prevalence was investigated in early life (0 – 5 months, i.e. first catch once fledged) and averaged across a bird's life. This excludes when the bird is a chick in the nest as the bird fledges after two weeks and it takes malaria a minimum of two weeks to develop and be able to be detected. Annual prevalence of malaria for the entire population of Cousin (all age classes) was plotted from 1997 to 2014 to show the current trend of malarial infection on the island (Fig S1). We constructed generalised linear mixed models (GLMMs) using the *lme4, lmerTest* and *car* packages in R (Bates *et al.* 2015, R Core Team 2015). Birth year was included as a random factor. For early-life malaria, the response variable was presence or absence of malaria, with a specified binomial error structure and built in Logit

Link function. For lifetime malaria, the response variable was a proportion measure of the number of samples where infection was detected over the total number of samples available for that individual, to which a quasi-binomial error structure was specified. A weighted correction factor was applied to control for the different lifespan's of the birds and thus the number of blood and malaria sample screens available.

TLR15 variation was represented as a predictor variable in two ways: (i) specific genotypes (> 0.05 in frequency) and (ii) Homozygosity / Heterozygosity. Given the only alleles observed at > 0.05 frequency in the population were 'A' 'B' and 'C', and alleles 'B' and 'C' were only found in the heterozygous state, all individuals possessed at least one copy of the 'A' allele. Therefore, homozygosity represents 'AA' individuals whereas heterozygosity represents both 'AB and AC' individuals. If heterozygosity was deemed significant in any of the statistical analyses, 'AB' and 'AC' were then separated out and the two alternate alleles were tested independently. Additional variables were included on the premise that they had previously been found to have some effect on SW fitness and thus needed controlling for: MHC diversity and *Ase-ua4* (Brouwer *et al.* 2010; Wright 2014), territory quality and season-born (Komdeur 1992; van de Crommenacker *et al.* 2011), sex (Richardson *et al.* 2002, 2003) and local density (Brouwer *et al.* 2006).

Survival analyses

Survival analyses were split into juvenile survival (likelihood of surviving into adulthood (>10 months)) and adult survival (overall lifespan). These were carried out using the *survival* and *flexsurv* packages in R. The optimal survival distribution curve for these analyses was obtained from plotting our data against five potential curves and determining the best fit using an Akaike Information Criterion (AIC) analysis. We found that the log-normal curve was the most optimal model distribution from six potential distributions tested, including: exponential, Weibull, Gamma, Gompertz and log-logistic (Log Likelihood = -962.37, df = 2, AIC = 1928.73). Therefore we then ran a generalised linear model based on the variable *surv* which takes into account survival over time fitted to the optimal log-normal curve, with *TLR15* variation as a fixed variable. The same *TLR15* variation factors used in the malaria analyses were used for the survival analyses.

Other variables included in the survival analyses were MHC diversity, presence/ absence of *Ase-ua4,* early-life malarial infection, season-born, territory quality (averaged for that particular season) and local density. Sex was excluded from the analysis as there was no significant difference in survival between sexes in preliminary analysis (U = 5035.50, df = 1, P = 0.68) nor was there any *a priori* reason to include it based on a previous study that established high annual survival in the SW had no difference between the sexes (Brouwer *et al.* 2010). The only variable not included to have previously been shown to influence survival was the presence of extreme weather conditions. We plotted weather variables over the time period of 1997-2014 to show the relatively benign and constant conditions over the years (Fig S2). Furthermore, territory quality correlates with weather variables.

Finally, we investigated variables which could influence the ways in which an individual responds to malaria exposure. There are five potential responses: (i) death (ii) complete resistance (never acquiring infection) (iii) tolerance / maintaining infection (iv) partial resistance (gained but cleared an infection) and (v) susceptibility through re-infection. These categories were not mutually-exclusive and a bird could show a number of these responses within its lifetime as we analysed all blood samples that had been screened for malaria which were available for each bird. We did one-way ANOVA's to look at variables exclusively and post-hoc multiple-comparison Tukey tests were able to decipher which categories were specifically being affected by the variable in question.

Results

Selection tests

In the 205 birds genotyped at the predetermined variable region of TLR15, four non-synonymous substitutions were identified (combining to create six unique alleles) and both haplotype and nucleotide diversity were low (Hd = 0.27, π = 0.001) The Fisher's exact probability test provided equivocal results, suggesting that genotypic frequencies were close to deviating from Hardy-Weinberg (HW) proportions (F_{IS} = 0.05*, P* = 0.08). This deviation was support by the more powerful *U*-test which rejected heterozygote deficiency (F_{IS} = 0.05*, P* = 0.04). This was further confirmed by HW one-tailed global tests, which identified three heterozygous genotypes where more individuals had been observed to possess these

heterozygote genotypes than expected under HW-equilibrium ('AB' P = 0.04, 'AC' P = 0.08 and 'AD' P = 0.03) (Fig 1).

Z-tests of selection failed to detect any signatures of selection at the *TLR15* locus in this cohort of SW, when testing for positive selection based on dN > dS (Z = 1.13, df = 139, *P* > 0.1), and negative selection based on dN < dS (Z = -1.16, df = 139, *P* > 0.1). This confirmed previous results from Chapter 3 where a smaller subset of birds (n = 30) were amplified at the *TLR15* locus to characterise variation. However, site-specific tests for selection which looked at each of the 140 codons individually, identified a single site under putative positive selection within this cohort using the FUBAR model (Posterior Prob dN > dS = 0.94). It is at this specific site (codon 51) where two non-synonymous mutations were observed, producing three different amino acids depending upon the genotype: 'AA' (Asparganine), 'AB' (Isoleucine) and 'AC' (Lysine).

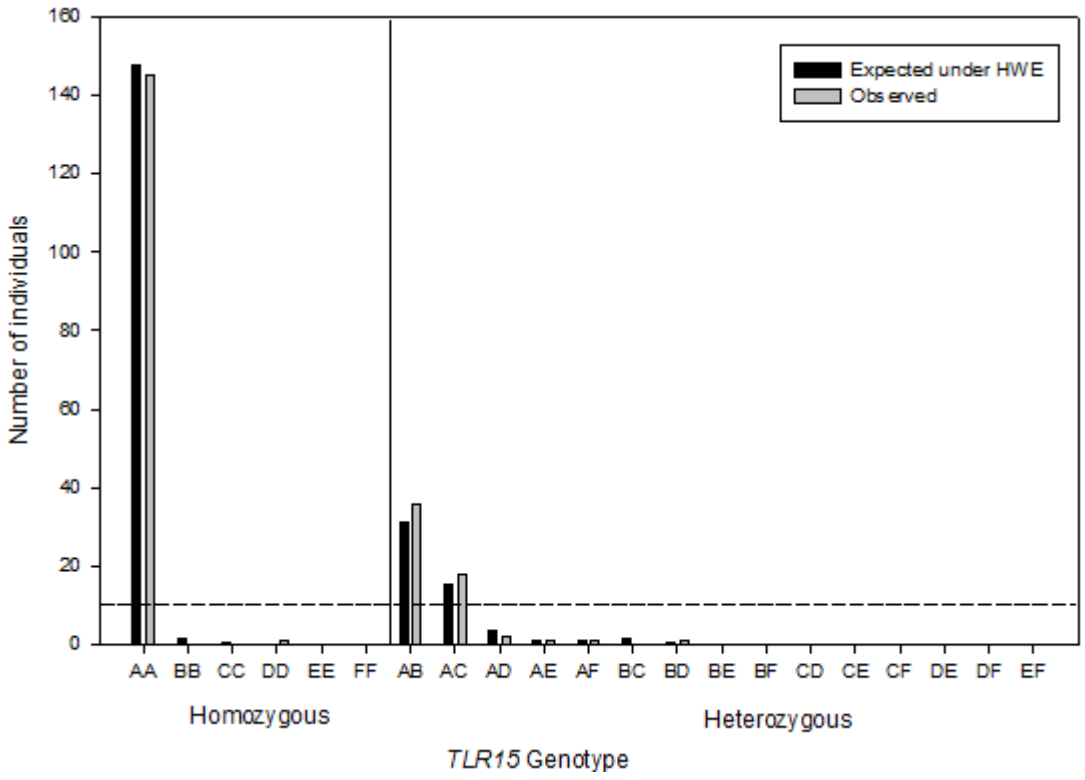


Figure 1. Expected and observed number of individuals with different *TLR15* genotypes, according to Hardy-Weinberg equilibrium (HWE). Dashed line represents the threshold of an allele frequency of 0.05 and thus all genotypes < 0.05 were then excluded from the association analyses.

Box 1: *Microsatellite and TLR15 heterozygosity*

Methods - We examined the ability of microsatellite multi-locus heterozygosity (MLH) to predict TLR heterozygosity of individuals of each species using a GLMM implemented by the 'MCMCglmm' function in the R-package *MCMCglmm* (Hadfield 2010). Individual MHL was calculated for a larger subset of individuals (n = 205) from the Seychelles warbler (SW) population, by taking an individual mean measure of heterozygosity across 30 microsatellite markers (Table S1), and then turning it into a measure of standardised heterozygosity (SH) (Coltman *et al.* 1999) to correct for the fact that not all 205 individuals have been successfully typed at all loci. This is important for standardising for variability of the typed markers to allow meaningful comparisons. It is possible to also use MLH metric internal relatedness (IR), which is a DNA-based measure of an individual's inbreeding coefficient (Amos *et al.* 2001). Another study found that both IR and SH were highly correlated and qualitatively similar (r = -0.929 for 216 individuals, Grueber *et al.* 2015). Typically, IR is more suited to studies with an aim to inform patterns of inbreeding, so we chose to use SH as this has a more centralised focus on evolutionary forces shaping variation in a population, measured as heterozygosity. For the model, we used measures of heterozygosity at the *TLR15 locus* (the most polymorphic TLR locus in the SW) as the response variable. Given this is a proportion measure the model was estimated with a logit link function (specified in the *MCMCglmm* package as the family 'Multinomial2'). SH was the fixed predictor variable and Bird ID was the random factor.

Results- On average, MLH showed no relationship with TLR heterozygosity (posterior mean = -0.140, *P*-MCMC = 0.824); the 95 % credible interval included zero (l – 95% = -1.214, U – 95% = 0.895). Raw data presented in 'Additional Information' (Table 6).

Discussion- We found no relationship between microsatellite SH and TLR heterozygosity in the SW. These results support the argument that microsatellite multi-locus heterozygosity (MLH) is not a good indicator of inter-individual variation in heterozygosity at genic regions, for example, at TLR loci. Microsatellite MLH is often estimated with a relatively small number of markers, however we have used 30 microsatellite markers specifically designed to be polymorphic and reliable for use in the SW (Richardson *et al.* 2000) to overcome issues raised in predicting individual-level genome-wide heterozygosity reliably (Balloux *et al.* 2004).

References

Amos W, Wilmer JW, Fullard K, *et al*. (2001) The influence of parental relatedness on reproductive success. Proc Biol Sci 268:2021–2027.

Balloux F, Amos W, Coulson T (2004) Does heterozygosity estimate inbreeding in real populations? Mol Ecol 13:3021–3031.

Coltman DW, Pilkington JG, Smith JA, Josephine M (1999) Parasite-Mediated Selection against Inbred Soay Sheep in a Free-Living, Island Population. Evolution 53:1259–1267.

Grueber CE, Knafler GJ, King TM, *et al*. (2015) Toll-like receptor diversity in 10 threatened bird species : relationship with microsatellite heterozygosity. Conserv Genet 16:595–611.

Hadfield JD (2010) MCMC methods for multi-response generalized linear mixed models: The MCMCglmm R package. J Stat Softw 33:1–22.

Richardson DL, Jury FL, Dawson DA, *et al*. (2000) Fifty Seychelles warbler (Acrocephalus sechellensis) microsatellite loci polymorphic in Sylviidae species and their cross-species amplification in other passerine birds. Mol Ecol 9:2155–7.

Association analyses

Malaria infection

Early-life malarial infection- 59% of juvenile birds was infected with malaria in early-life. Individuals with the 'AC' TLR genotype were significantly more likely to test positive for early-life malaria infection and 'AA' individuals were significantly less likely to have early-life malaria, given it is the only homozygous genotype included when testing the effects of homozygosity (Figure 1, Table 1). MHC variation and the specific MHC allele *Ase-ua4* had no significant association with early-life malaria and nor was there any significant difference between sexes. However, being born in the winter season and / or being born in a territory of high quality increased the likelihood of infection (Table 1, Table S2).

Table 1. Investigating how TLR variation is associations with early-life malarial infection in the Seychelles warbler using generalised linear mixed models which include both genetic and ecological factors. All models were built with a logit link function and binomial error structure. Codes applied to *P*-values to show significance are as follows: . (<0.1) * (<0.05) ** (<0.01) and *** (< 0.001).

Explanatory variables	N	Significant variables	Estimate	Std. Error	Z	*P*- value	2.5% CI	97.5% CI
TLR15 genotype, microsatellite Hz, sex, season-born	205	*TLR15* genotype- 'AC'	1.231	0.552	2.228	0.026*	0.176	2.375
		Season-born - 'Winter'	1.544	0.369	4.187	0.000***	0.838	2.291
TLR15 Hm / Hz, sex, microsatellite Hz season-born	205	Homozygous (Hm)	-0.597	0.329	-1.815	0.070 .	-1.246	0.047
		Season-born - 'Winter'	1.446	0.358	4.039	0.000***	0.759	1.694
TLR15 Hap B, *TLR15* Hap C, sex, microsatellite Hz season-born	205	*TLR15* haplotype C	1.235	0.549	2.250	0.024*	0.188	2.372
		Season-born - 'Winter'	1.440	0.359	4.008	0.000***	0.750	1.656
TLR15 genotype, season-born, *Ase-ua4,* MHC diversity	111	*TLR15* Genotype- 'AC'	1.253	0.692	1.812	0.070 .	-0.051	2.714
		Season-born - 'Winter'	1.894	0.485	3.905	0.000***	0.983	2.901
TLR Hm / Hz, season-born, *Ase-ua4,* MHC diversity	111	Season-born - 'Winter'	1.776	0.471	3.773	0.000***	0.891	2.751
TLR Hm / Hz sex, season-born, territory quality	65	Territory quality	0.128	0.056	2.293	0.022*	0.042	0.285

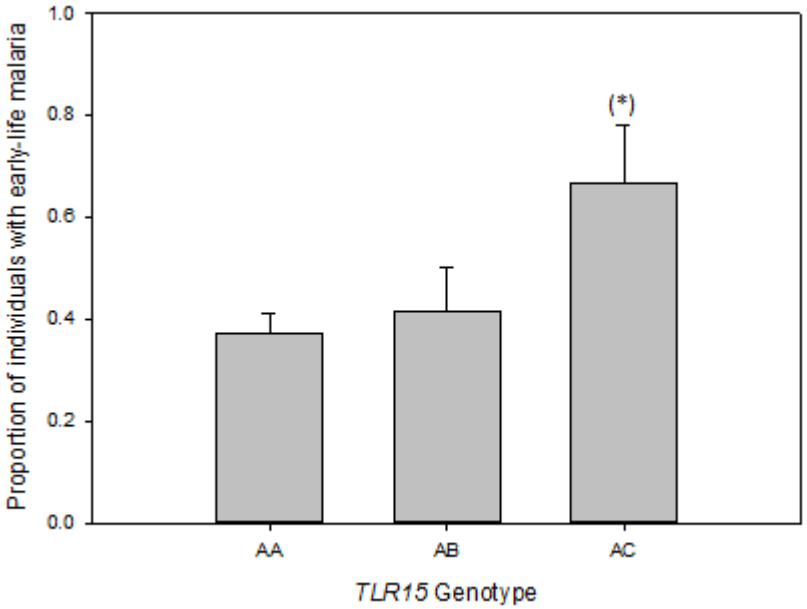


Figure 2. Proportion of individuals with early-life malaria in relation to their genotype at the *TLR15* locus. Only genotypes made up from alleles with frequencies > 5% in the population are shown. * indicates genotypes that significantly influenced life-time malaria at $P < 0.05$.

Lifetime malarial infection- TLR15 variation did not influence average lifetime rate of malarial infection, nor did any other immunogenetic variable (MHC diversity and *Ase-ua4*). However, both sex and early-life exposure to malaria and gaining infection, showed a significant interaction in negatively influencing the likelihood of life-time malarial infection. Male birds that had gained malarial infection in early-life had significantly reduced chances of being re-infected later in life (Figure 2, Table 2 & S3).

Survival

Juvenile survival - 67.8% of juvenile birds survived to adulthood. When investigating the association between *TLR15* variation and juvenile survival, early-life malaria predicted survival in all models (Table 3). *TLR15* variation did not appear to predict survival , but given its significant association with early-life malaria, we also did separate models of juvenile survival excluding early-life malaria. *TLR15* variation still did not significantly predict survival: genotype 'AB' ($z = -0.047$, $P = 0.962$) 'AC' ($z = -0.775$, $P = 0.438$) 'AA / Homozygous' ($z = 0.316$, $P = 0.752$) haplotype 'B' ($z = 0.024$, $P = 0.981$) and 'C' ($z = -0.803$, $P = 0.422$). With all variables included MHC diversity and *Ase-ua4* had no effect on juvenile survival but territory quality almost had a significant effect, considering the model had little power (Table 3 & S4).

Table 2. Investigating if TLR variation is associated with average lifetime malarial infection in the Seychelles warbler, using generalised linear mixed models which include both genetic and ecological fixed factors. All models were built with a quasi-binomial error structure with 'Birth Year' included as a random factor. Codes applied to *P*-values to show significance are as follows: . (<0.1) * (<0.05) ** (<0.01) and *** (< 0.001).

Explanatory variables	N	Significant variables	Estimate	Std. Error	t	*P*-value	2.5% CI	97.5% CI
TLR15 genotype, microsatellite Hz, early-life malaria*sex, local density, sex, season-born	205	Early-life malaria*Sex	-1.112	0.545	-2.039	0.043*	-2.188	-0.046
TLR Hm / Hz, microsatellite Hz, early-life malaria*sex, local density, sex, season-born	205	Early-life malaria*Sex	-1.120	0.545	-2.039	0.043*	-2.188	-0.046
TLR15 Hap B, *TLR15* Hap C, microsatellite Hz, early-life malaria*sex, local density, sex, season-born	205	Early-life malaria*Sex	-1.112	0.545	-2.039	0.043*	-2.188	-0.046
Early-life malaria*sex, *Ase-ua4*, MHC diversity	110	Early-life malaria*Sex	-1.430	0.747	-1.914	0.059 .	-2.908	0.030
Early-life malaria*sex, *Ase-ua4*, MHC diversity, territory quality	65	Early-life malaria*Sex	-2.278	1.103	-2.065	0.044*	-4.544	-0.175
		Early-life malaria	2.285	1.054	2.168	0.035*	0.312	4.527

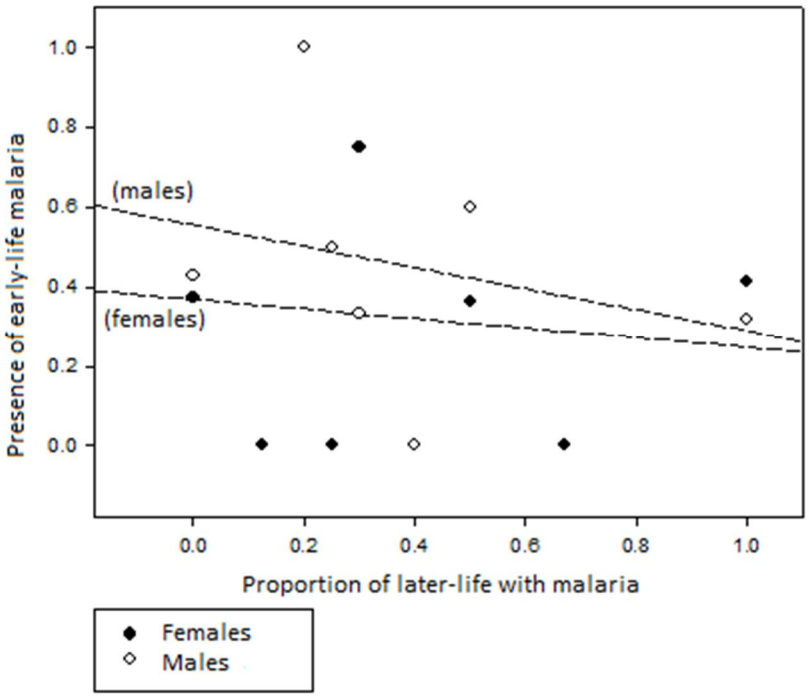


Figure 3. Relationship between being infected with malaria in early-life and being infected overall throughout the lifetime for females and males analysed separately.

Table 3. Investigating TLR variation in relation to juvenile survival in the Seychelles warbler using generalised linear mixed models which include both genetic and ecological fixed factors. All models were built with a logit link function and binomial error structure. Codes applied to *P*-values to show significance are as follows: . (<0.1) * (<0.05) ** (<0.01) and *** (< 0.001).

Explanatory variables	N	Significant variables	Estimate	Std. Error	Z	*P*- value	2.5% CI	97.5% CI
TLR15 genotype, microsatellite Hz, early-life malaria, season-born	205	Early-life malaria	1.501	0.384	3.908	0.000***	0.777	2.293
TLR15 Hm / Hz, microsatellite Hz, early-life malaria, season-born	205	Early-life malaria	1.419	0.369	3.842	0.000***	0.720	2.176
TLR15 Hap B, *TLR15* Hap C, Hm / Hz, microsatellite Hz, early-life malaria, season-born	205	Early-life malaria	1.497	0.380	3.937	0.000***	0.781	2.281
Early-life malaria, *Ase-ua4*, MHC diversity	111	Early-life malaria	1.726	0.450	3.833	0.0001***	0.872	2.648
Early-life malaria, *Ase-ua4,* MHC diversity, territory quality	65	Early-life malaria	3.236	0.864	3.743	0.0002***	1.716	5.238
		Territory quality	-0.058	0.035	-1.667	0.0955 .	-0.131	0.007

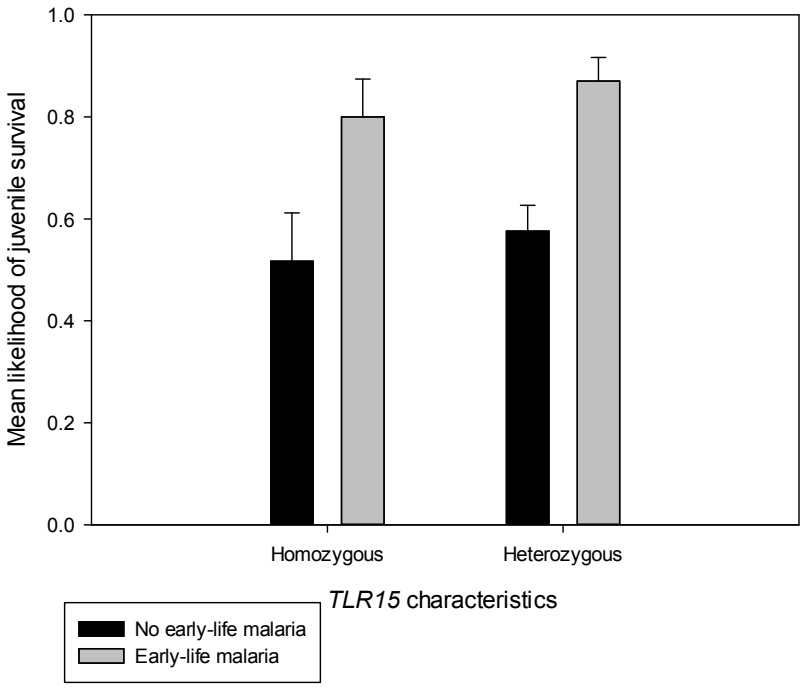


Figure 4. Relationship between individual *TLR15* characteristics and mean likelihood of juvenile survival using the comparison of homozygous ('AA') and heterozygous individuals (including Hz genotypes derived from the three alleles at frequencies >5% ('AB' or 'AC'). These categories are further divided into individuals with and without malaria in early-life.

Adult survival – TLR15 variation had no significant effect on adult survival in the Seychelles warbler (Table S5). When we removed early-life malaria from the model, given its established relationship with *TLR15* variation, we still found that *TLR15* variation had no significant effect on adult survival for genotype 'AB' (t = -0.604, *P* = 0.546) 'AC' (t = 0.117, *P* = 0.907) 'AA' / homozygous (t = 0.154, *P* = 0.878) haplotype 'B' (t = -1.043, *P* = 0.298) and 'C' (t = 0.001, *P* = 0.999). However, with all variables included, individuals with higher MHC diversity had longer adulthood lifespans (Table 4; Fig 5). Furthermore, lifetime malarial infection predicted adult survival in that individuals that continue to show resilience to infection are more likely to live for longer (Table 4). The only other ecological factor to be related to adult survival was local density in that living in a large group in adulthood was negatively related to survival (Table 4; Fig S5). These results were fully supported when an alternative GLMM was ran which was not fitted to a model distribution survival curve and simply used adult survival measured in months as the response variable (Table S6).

Table 4. Investigating TLR variation in relation to adult survival in the Seychelles warbler using generalised linear mixed models of survival over time fitted to a log-normal distribution curve using the surv package in R (R Core Team 2015). Codes applied to *P*-values to show significance are as follows: . (<0.1) * (<0.05) ** (<0.01) and *** (< 0.001).

Explanatory variables	N	Significant variables	Value	Std. Error	Z	*P*- value	2.5% CI	97.5% CI
TLR15 genotype, microsatellite Hz, early-life malaria, life-time malaria, local density	205	Local Density	-5.756	2.370	-2.429	0.016*	-10.400	-1.111
		Life-time malaria	14.738	7.557	1.950	0.053 .	-0.073	29.549
TLR15 hm / hz, Microsatellite Hz, early-life malaria, life-time malaria, local density	205	Local Density	-6.353	2.327	-2.729	0.007 **	-10.914	-1.791
		Life time malaria	16.272	7.418	2.193	0.030 *	1.732	30.812
TLR15 hap B, *TLR15* hap C, microsatellite Hz, early-life malaria, life-time malaria, local density	205	Local Density	-6.385	2.327	-2.743	0.007**	-10.946	-1.823
		Life time malaria	15.388	7.483	2.057	0.041*	0.723	30.054
Life-time malaria, local density, *Aseua4*, MHC diversity	110	Local Density	-0.109	0.056	-1.957	0.0503 .	-0.219	0.000
		Life time malaria	0.352	0.181	1.947	0.0516 .	-0.002	0.706
		MHC diversity	0.107	0.049	2.175	0.0296*	0.011	0.204
Life-time malaria, local density, *Aseua4*, MHC diversity, territory quality	65	Local Density	-0.145	0.074	-1.960	0.0499*	-0.290	-3.623
		Life time malaria	0.747	0.246	3.043	0.0023**	0.266	1.228

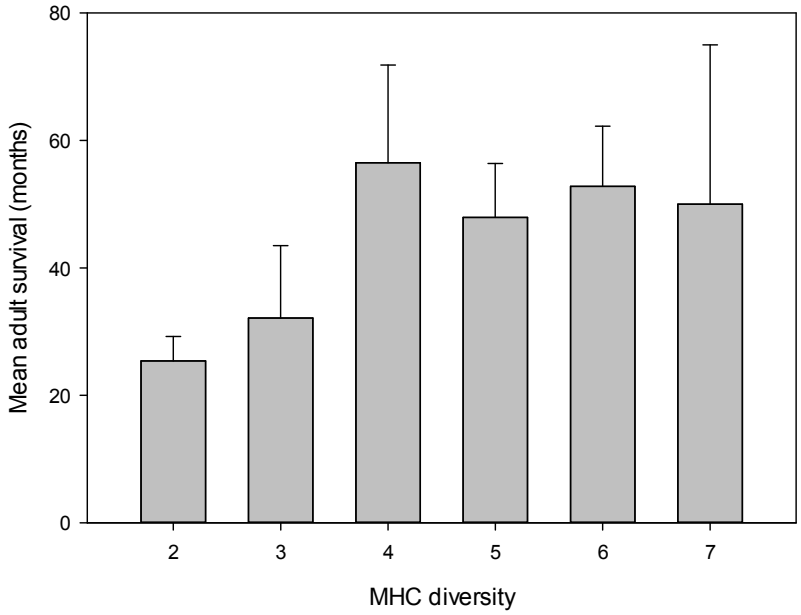


Figure 5. Association between MHC diversity and mean adult survival (months) in a cohort of Seychelles warbler where MHC diversity represents a total count of unique MHC alleles.

Individual response to infection - There is no significant difference between the different categories of response to malarial infection observed in the SW with *TLR15* variation. Having a specific *TLR15* allele had no effect on the patterns of response to infection observed: allele 'A' ($F = 1.496$, df = 4, $P = 0.205$) allele 'B' ($F = 0.351$, df = 4, $P = 0.843$) and allele 'C' ($F = 1.192$, df = 4, $P = 0.315$). However, the presence of MHC allele '*Ase-ua4'* did significantly influence the patterns of malaria infection observed when looking across the different responses ($F = 3.859$, df = 4, $P = 0.006$) (Fig 6). When looking within the different outcomes, it appeared that individuals that do possess the *Ase-ua4* allele had significantly less likelihood of dying as a result of infection and more likely to respond in one of the other categorised ways, such as tolerating infection or clearing the infection ($X^2 = 71.329$, df = 1, $P < 0.001$). Territory quality had no effect on individual outcome of infection ($F = 1.281$, df = 4, $P = 0.283$) but local density did have an effect ($F = 3.112$, df = 4, $P = 0.016$). Individuals who lived in larger residential groups were less likely to gain an infection in the first place ($X^2 = 4.365$, df = 1, $P = 0.037$) (Fig 7).

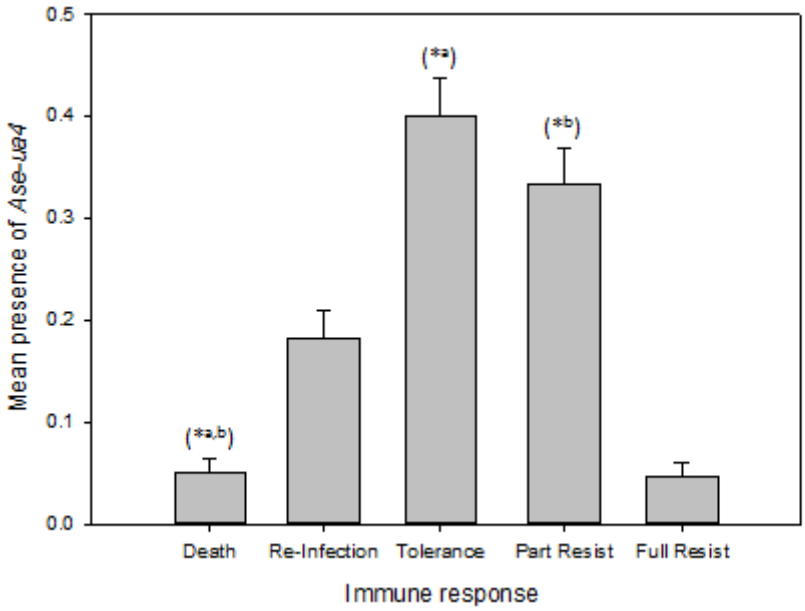


Figure 6. Different outcomes to malarial parasite exposure by Seychelles warblers during their lifetime and its association with whether the bird possesses the MHC allele *Ase-ua4.* Potential outcomes include: death, re-infection, tolerance, partial resistance (clearing an infection) and full resistance (complete avoidance of infection). * indicates outcomes that were significantly different from one another and the letters denote the pairwise relationship.

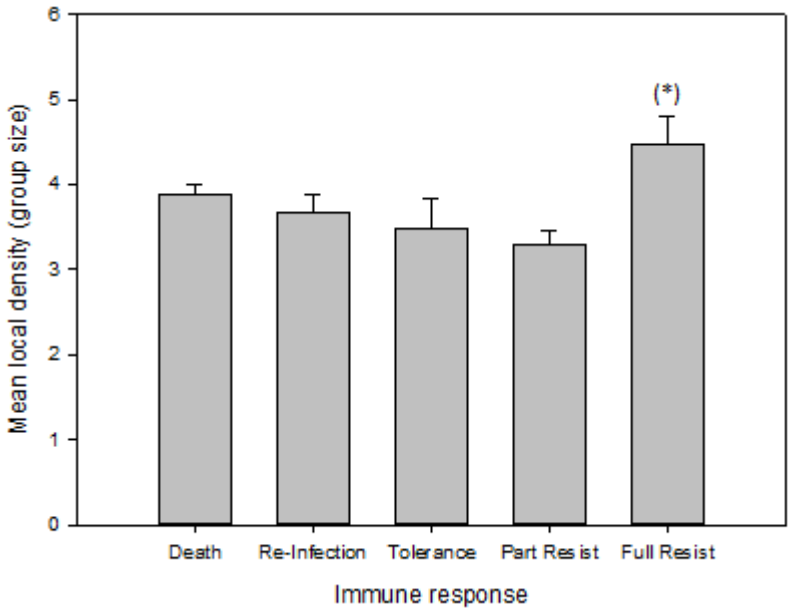


Figure 7. Different outcomes elicited to malarial parasite exposure by Seychelles warblers during their lifetime and its association with mean local density (resident territory group size). Potential outcomes include: death, re-infection, tolerance, partial resistance (clearing an infection) and full resistance (complete avoidance of infection). * indicates genotypes that significantly influenced life-time malaria at $P < 0.05$.

Discussion

We investigated the effect of individual variation at the polymorphic *TLR15* locus on malarial infection and survival within an isolated population of the Seychelles warbler (SW). We found that individuals possessing the specific 'AC' TLR 15 genotype - or arguably the 'C' haplotype as this allele was only observed in a heterozygous state with allele A -significantly influenced the likelihood of being infection with malaria when sampled on the natal territory. Individuals with 'AC' were more likely to have early-life malaria and 'AA' individuals were least likely to have early-life malaria. *Haemosporidian* parasites normally takes ca two weeks to develop into an infection (Garnham 1980), so no chicks were infected in the SW. Consequently all the infected individuals of 0 – 4 months will have at least fledged from the nest. Avian malaria, like malaria in all other vertebrates, consists of a number of stages: (i) a pre-patent stage shortly after transmission when parasites develop in host tissues, (ii) acute phase where parasites are in the blood and parasitaemia increases, thus having negative symptomatic effects on the host, and (iii) the latent / chronic phase when parasitaemia falls (Garnham 1980; Atkinson & van Riper III 1991; Thomas *et al.* 2008). This latter phase can last for years, even for life, and relapses can occur (e.g. Bensch *et al.* 2007; Lachish *et al.* 2011; Asghar *et al.* 2015). Therefore, the individuals that we catch with infection will be in the latent / chronic phase and are essentially already 'survivors' of the infection. Consequently, individuals with the 'AC' genotype are in this chronic phase and thus having the 'C' haplotype has provided a form of resilience against the pathogen. These same individuals that had early-life malaria infection were also less likely to become re-infected as an adult also supports this idea of these individual having a degree of resilience/resistance. Individuals that are homozygous for 'AA' are more likely to have been exposed to malaria and not survived the acute phase, and therefore not be sampled.

These results support the hypothesis of pathogens mediating balancing selection within this bottlenecked population. It is important to note that neutral variation had no effects on disease resistance or survival in our models, but 'adaptive' variation did. Having the specific heterozygous combination of an 'A' and 'C' allele has advantages over being homozygous for either 'A' or 'C' on its own. This is evidence of overdominance, a mechanism of heterozygote advantage (Doherty & Zinkernagel 1975). However, heterozygote advantage is not the only mechanism in effect, as it does not explain why 'AB'

heterozygous individuals do not share the same benefits observed with 'AC' individuals. Balancing selection due to spatio-temporal fluctuations in selection favouring one particular allelic variant over others is consistent with our results as an explanatory mechanism (Robertson & Hill 1984) as is the rare-allele advantage hypothesis if the 'C' allele has only recently emerged in the population gene pool (Slade & McCallum 1992). Therefore, this emphasises how a number of mechanisms can be proposed to explain pathogen-mediated balancing selection and they do act in concert (for excellent review, see (Spurgin & Richardson 2010).

TLR15 variation had no direct association with individual survival in the SW. However malarial infection, which was in part influenced by TLR variation, did appear to affect survival. Consequently we suspect TLR15 variation must have an indirect role on survival through this interaction. Studies on another island endemic passerine species- the Stewart Island Robin *Petroica australis raikura*- found a survival advantage conferred by the presence of a specific *TLR4* allele (Grueber *et al.* 2013). However, this was one of only two TLR genes that were indeed found to be monomorphic in the SW population. This suggests that there is large variation in pathogen-selection regimes on different islands and perhaps, there is a paucity of pathogens on Cousin Island where the SW is a suitable host.

Overall, our results for TLR15 are very much in line with other studies. It appears that the locus is generally highly-conserved and under purifying selection, but shows evidence of positive (balancing) selection at specific sites even if the rate of non-synonymous (dN) substitutions is slow. This was the consensus found when Alcaide & Edwards (2011) examined ten TLR genes in seven phylogenetically-distant avian species. Another m;4eta-analyses has also shown this in-depth by looking at eight different vertebrate species (including human, chimpanzee, macaque, mouse, cow, chicken, western clawed frog and zebrafinch) and showing that all genes in the TLR signalling pathway are highly conserved (Song *et al.* 2012). Only specific sites are under positive selection and they are always sites involved with the extracellular leucine-rich repeat domain responsible for pathogen recognition. Nakajima *et al.* (2008) show the extent of the 'rapid evolution' occurring specifically in this domain of TLRs across primates and has even been shown in cetaceans with the common effect of functional constraint but some codons having made radical changes with parallel evolution between independent lineages (Shen *et al.* 2012).

Mukherjee *et al.* (2009) have further shown how this is an example of local adaptation in humans. They looked at six TLRs in 171 Indian people with high microbial loads and show the large diversity at these loci just compared to European and African populations. Interestingly, they find an excess of rare variants but low tolerance of dN substitutions. We also find an excess of rare heterozygous alleles in the SW and find low tolerance, with the exception of the rare allele proving advantageous (Slade & McCallum 1992). Studies on other innate immune genes have mirrored our findings by finding specific genotypes confer a fitness advantage. Basu *et al.* (2012) had already showed that dN substitutions at the *TLR4* locus influenced blood infection load of *Plasmodium falciparum.* However, further work looking at the Interleukin 12B gene in humans showed an 'AC' genotype increased log-parasitaemia levels specifically ($P = 0.01$). This is what we found in the SW and is an excellent example of how studies on model species can be applied to other taxa, including humans and this research holds much importance in the hope of developing novel adjuvants.

Consistent with previous studies on the SW (Brouwer *et al.* 2010), we did find other immunogenetic variation directly influenced adult survival. Individual MHC diversity was positively related to the lifespan of the bird. Such a relation between MHC diversity and survival has also been shown across a range of vertebrate taxa (for examples, see Wegner *et al.* 2003; Kalbe *et al.* 2009; Sepil *et al.* 2013). Interestingly, the specific MHC allele *Ase-ua4* did not appear to have significant effects on survival in this particular SW cohort (Brouwer *et al.* 2010; Wright 2014). However, we suspect there is an underlying fitness effect present which went undetected due to limited power in our analysis. This underlying fitness effect is related to our analysis into differential pathogen-infection outcomes, of which we showed that the presence of *Ase-ua4* allele did significantly reduce post-malarial infection mortality. In fact, individuals carrying Ase-ua4 were more likely to be able to tolerate the infection.

Another interesting result from this study is the observed differences between sexes with malaria infection (and consequently, survival). We found that males born in the winter season are less likely to be infected later in life because they are more likely to have early-life malaria. We showed that winter-born birds had increased chances of early life infection. This is not surprising, given that these months are hotter and wetter and thus promote the abundance of dipteran vectors, such as mosquitoes. The fact that territory quality (a

measure of local insect availability) was also positively correlated with early-life malaria was consistent with this result. Therefore, it appears that male SW born in the winter are surviving this increased early-life malaria better than the female SW, based on the previously outlined theory of only catching birds in the chronic phase of infection. This could reflect the different gender roles within the SW breeding system (Richardson *et al.* 2002, 2003) and thus be an example of the Immuno-competence handicap hypothesis where different sexes have different levels of investment in immunity and reproduction (Folstad & Karter 1992; for review, see Roberts *et al.* 2004).

It is clear that the role of immunogenetic variation in determining malaria infection and survival in the SW population could explain its maintenance and drive within the population. Although, the environmental factors we included based on previous studies are also important and interact with immunogenetic variables. Local density influences adult survival, which is not surprising given that Cousin is a small island with finite resources and local competition will heavily associate with food (and other resources) availability. It is also in concurrence with previous findings by Brouwer *et al.* (2006). The advantages to helping in this system (e.g. Komdeur 1994; Richardson *et al.* 2003, 2007) and their trade-off with finite resources and territory quality (e.g. Richardson *et al.* 2004; Brouwer *et al.* 2006) are already well-documented. Our novel finding concerning local density was its relationship with differential pathogen-infection outcomes once an individual had been exposed to the malarial parasite. Of all the different possible outcomes, having a larger local density appeared to increase the likelihood of complete resistance, which is when a bird consistently tests negative for infection. This is not what we expected given our findings that larger local density reduces individual survival, which is also well-supported from a previous SW study (Brouwer *et al.* 2006). However, this pattern has been shown in other studies in a range of vertebrates including birds, rodents and primates (Plaut *et al.* 1969; Daviews *et al.* 1991; Marzal *et al.* 2005). Some of these studies have used a 'dilution' effect of vector activity to explain their findings and this was further investigated in a meta-analysis study, which conclusively showed that intensity of infection by mobile parasites or parasites requiring intermediate vector hosts, consistently decreased as host group size increased (Cote & Poulin 1995). However, this would not sufficiently explain this result in the SW given the enormous abundance of vector (mosquito) species. Local density is a mean measure of

the number of individuals in a resident territory and this would include birds within and outside of the breeding group. Therefore, I hypothesise that what we are seeing is less of a local-density effect, and perhaps a reflection on social roles. A higher local density will represent a bigger range of social roles including dominant breeders, helpers and 'other birds'- birds that reside in a territory but have yet to gain a social role. 'Other birds' will be less likely to gain infection due to their isolation and increased activity. This also means they would be less likely to acquire immunity in their 'naïve' state which could have negative consequential effects on survival and not maximise TLRs ability to link innate and adaptive immune defence (for reviews, see Akira *et al.* 2001; Schnare *et al.* 2001). This is on top of not gaining the fitness benefits that come with helping in a social breeding system (Wiley & Rabenold 1984; Griffin & West 2003; Komdeur *et al.* 2014).

In conclusion, it is important that we establish the key factors which influence SW survival and thus shape its evolution. We have focused on innate immunogenetic variation at a relatively polymorphic TLR locus. We have shown that TLR characteristics have a role in resilience to malaria in early-life, which consequently leads to reduced infection in later life and benefits to overall survival. Our results also support previous studies which indicated that the MHC influences survival in this species, and we have shown that this may be because of its interaction with malaria. Finally we have confirmed the importance of specific ecological factors that interact with genetic factors and pathogens as part of an overall evolutionary framework. Elucidating the components of this framework has important conservation implications, particularly for maintaining genetic diversity as part of intensive management of a species (Grueber & Carolyn 2015).

Acknowledgments

We thank Nature Seychelles for facilitating the work on Cousin Island. We would like to thank the Seychelles Bureau of Standards and the Department of Environment for giving permission for sampling and fieldwork.

Data Accession Statement

All sequences used in the study have been published and are available in GenBank (accession numbers: KT203560-KT203565).

References

Akira S, Takeda K, Kaisho T (2001) Toll-like receptors: critical proteins linking innate and acquired immunity. *Nature immunology*, **2**, 675–680.

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

Amos W, Wilmer JW, Fullard K *et al.* (2001) The influence of parental relatedness on reproductive success. *Proceedings. Biological sciences / The Royal Society*, **268**, 2021–2027.

Areal H, Abrantes J, Esteves PJ (2011) Signatures of positive selection in Toll-like receptor (TLR) genes in mammals. *BMC evolutionary biology*, **11**, 368.

Asghar M, Hasselquist D, Hansson B *et al.* (2015) Hidden costs of infection: Chronic malaria accelerates telomere degradation and senescence in wild birds. *Science*, **347**, 436–438.

Atkinson CT, van Riper III C (1991) *Pathogenicity and epizootiology of avian haematozoa: Plasmodium, Leucocytozoon, and Haemoproteus.*

Atkinson CT, Woods KL, Dusek RJ, Sileo LS, Iko WM (1995) Wildlife disease and conservation in Hawaii: Pathogenicity of avian malaria (Plasmodium relictum) in experimentally infected Iiwi (Vestiaria coccinea). *Parasitology*, **111**, S59.

Balloux F, Amos W, Coulson T (2004) Does heterozygosity estimate inbreeding in real populations? *Molecular Ecology*, **13**, 3021–3031.

Barreiro LB, Quintana-Murci L (2010) From evolutionary genetics to human immunology: how selection shapes host defence genes. *Nature Reviews Genetics*, **11**, 17–30.

Basu M, Das T, Ghosh A *et al.* (2012) Gene-gene interaction and functional impact of polymorphisms on innate immune genes in controlling Plasmodium falciparum blood infection level. *PLoS One*, **7**, e46441.

Bates D, Mächler M, Bolker BM, Walker SC (2015) Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, arXiv:1406.5823v1.

Bensch S, Waldenström J, Jonzén N *et al.* (2007) Temporal dynamics and diversity of avian malaria parasites in a single host species. *Journal of Animal Ecology*, **76**, 112–122.

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

Biedrzycka A, Radwan J (2008) Population fragmentation and major histocompatibility complex variation in the spotted suslik, Spermophilus suslicus. *Molecular Ecology*, **17**, 4801–4811.

Bihl F, Salez L, Beaubier M *et al.* (2003) Overexpression of Toll-like receptor 4 amplifies the host response to lipopolysaccharide and provides a survival advantage in transgenic mice. *Journal of Immunology*, **170**, 6141–50.

Bonneaud C, Balenger SL, Russell AF *et al.* (2011) Rapid evolution of disease resistance is accompanied by functional changes in gene expression in a wild bird. *PNAS*, **108**, 7866–7871.

Bonneaud C, Perez-Tris J, Federici P, Chastel O, Sorci G (2006) Major Histocompatibility Alleles Associated with Local Resistance to Malaria in a Passerine. *Evolution*, **60**, 383–389.

Bordes F, Morand S (2015) *Parasite Diversity and Diversification: Evolutionary Ecology Meets Phylogenetics* (S Morand, BR Krasnov, DTJ Littlewood, Eds,). Cambridge University Press.

Boyd AC, Peroval MY, Hammond JA *et al.* (2012) TLR15 is unique to avian and reptilian lineages and recognizes a yeast-derived agonist. *Journal of Immunology*, **189**, 4930–4938.

Brouwer L, Barr I, van de Pol M *et al.* (2010) MHC-dependent survival in a wild population: evidence for hidden genetic benefits gained through extra-pair fertilizations. *Molecular Ecology*, **19**, 3444–3455.

Brouwer L, Richardson DS, Eikenaar C, Komdeur J (2006) The role of group size and environmental factors on survival in a cooperatively breeding tropical passerine. *Journal of Animal Ecology*, **75**, 1321–9.

Brouwer L, Tinbergen JM, Both C *et al.* (2009) Experimental evidence for density-dependent reproduction in a cooperatively breeding passerine. *Ecology*, **90**, 729–741.

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

Collar NJ, Stuart SN (1985) *Threatened birds of Africa and related islands.* tional Council for Bird Preservation, Cambridge.

Coltman DW, Pilkington JG, Smith JA, Josephine M (1999) Parasite-Mediated Selection against Inbred Soay Sheep in a Free-Living, Island Population. *Evolution*, **53**, 1259–1267.

Cote IM, Poulin R (1995) Parasitism and Group-size In Social Animals - A Metaanalysis. *Behavioral Ecology*, **6**, 159–165.

Creagh EM, O'Neill L a J (2006) TLRs, NLRs and RLRs: a trinity of pathogen sensors that co-operate in innate immunity. *Trends in Immunology*, **27**, 352–7.

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

Van de Crommenacker J, Richardson DS, Koltz AM, Hutchings K, Komdeur J (2012) Parasitic infection and oxidative status are associated and vary with breeding activity in the Seychelles warbler. *Proceedings of the Royal Society B*, **279**, 1466–1476.

Daviews CR, Ayres JM, Dye C, Deane LM (1991) Malaria infection rate of Amazonian primates increases with body weight and group size. *British Ecological Society*, **5**, 655–662.

Delport W, Poon AFY, Frost SDW, Kosakovsky Pond SL (2010) Datamonkey 2010: a suite of phylogenetic analysis tools for evolutionary biology. *Bioinformatics (Oxford, England)*, **26**, 2455–7.

Deter J, Charbonnel N, Cosson J-F, Morand S (2007) Regulation of vole populations by the nematode Trichuris arvicolae: insights from modelling. *European Journal of Wildlife Research*, **54**, 60–70.

Doherty PC, Zinkernagel RM (1975) Enhanced immunological surveillance in mice heterozygous at the H-2 gene complex. *Nature*, **256**, 50–52.

Escalante A a., Cornejo OE, Rojas A, Udhayakumar V, Lal A a. (2004) Assessing the effect of natural selection in malaria parasites. *Trends in Parasitology*, **20**, 388–395.

Ferrer ES, García-Navas V, Sanz JJ, Ortego J (2014) Individual genetic diversity and probability of infection by avian malaria parasites in blue tits (Cyanistes caeruleus). *Journal of evolutionary biology*, 1–15.

Flajnik MF, Du Pasquier L (2004) Evolution of innate and adaptive immunity: Can we draw a line? *Trends in Immunology*, **25**, 640–644.

Folstad I, Karter AJ (1992) Parasites, bright males and the Immunocompetence handicap. *The American Naturalist*, **139**, 603–622.

Fonesca DM, LaPointe DA, Fleischer RC (2000) Bottlenecks and multiple introductions: population genetics of the vector of avian malaria in Hawaii. *Molecular Ecology*, **9**, 1803–1814.

Ford MJ (2002) Applications of selective neutrality tests to molecular ecology. *Molecular Ecology*, **11**, 1245–62.

Fuller T, Bensch S, Müller I *et al.* (2012) The ecology of emerging infectious diseases in migratory birds: An assessment of the role of climate change and priorities for future research. *EcoHealth*, **9**, 80–88.

Garamszegi LZ, Zagalska-Neubauer M, Canal D *et al.* (2015) Malaria parasites, immune challenge, MHC variability, and predator avoidance in a passerine bird. *Behavioral Ecology*, **00**, 1–11.

Garnham PCC (1980) *Malaria in its various vertebrate hosts. In "Malaria, Epidemiology, Chemotherapy, Morphology, and Metabolism".* (JP Kreier, Ed,). Academic Press, New York.

Griffin AS, West S a (2003) Kin discrimination and the benefit of helping in cooperatively breeding vertebrates. *Science*, **302**, 634–636.

Griffiths R, Double MC, Orr K, Dawson RJG (1998) A DNA test to sex most birds. *Molecular Ecology*, **7**, 1071–1075.

Grueber CE, Carolyn J (2015) Impacts of early viability selection on management of inbreeding and genetic diversity in conservation. *Molecular Ecology*, **24**, 1645–1653.

Grueber CE, Knafler GJ, King TM *et al.* (2015) Toll-like receptor diversity in 10 threatened bird species : relationship with microsatellite heterozygosity. *Conservation Genetics*, **16**, 595–611.

Grueber CE, Wallis GP, Jamieson IG (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**, 4470–4482.

Grueber CE, Wallis GP, Jamieson IG (2014) Episodic positive selection in the evolution of avian toll-like receptor innate immunity genes. *PloS one*, **9**, e89632.

Guo SW, Thompson EA (1992) Performing the exact test of Hardy-Weinberg proportion for multiple alleles. *Biometrics*, **48**, 361–372.

Hadfield JD (2010) MCMC methods for multi-respoinse generalized linear mixed models: The MCMCglmm R package. *Journal of Statistical Software*, **33**, 1–22.

Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium*, **41**, 95–98.

Hedrick PW (2002) Pathogen resistance and genetic variation at MHC loci. *Evolution*, **56**, 1902–1908.

Hedrick PW (2006) Genetic Polymorphism in Heterogeneous Environments: The Age of Genomics. *Annual Review of Ecology, Evolution, and Systematics*, **37**, 67–93.

Higgs R, Cormican P, Cahalane S *et al.* (2006) Induction of a Novel Chicken Toll-Like Receptor following Salmonella enterica Serovar Typhimurium Infection Induction of a Novel Chicken Toll-Like Receptor following Salmonella enterica Serovar Typhimurium Infection. *Infection and Immunity*, **74**, 1692–1698.

Hill AVS, Allsopp CEM, Kwiatkowski D *et al.* (1991) Common West African HLA antigens are associated with protection from severe malaria. *Nature*, **352**, 595–560.

Hudson PJ (1986) The Effect of a Parasitic Nematode on the Breeding Production of Red Grouse. *Journal of Animal Ecology*, **55**, 85–92.

Jeffery KJM, Bangham CRM (2000) Do infectious diseases drive MHC diversity? *Microbes and Infection*, **2**, 1335–1341.

Jin MS, Lee JO (2008) Structures of the toll-like receptor family and its ligand complexes. *Immunity*, **29**, 182–191.

Kalbe M, Eizaguirre C, Dankert I *et al.* (2009) Lifetime reproductive success is maximized with optimal major histocompatibility complex diversity. *Proceedings of the Royal Society B*, **276**, 925–934.

Kalinowski ST (2006) HW-QUICKCHECK: An easy-to-use computer program for checking genotypes for agreement with Hardy-Weinberg expectations. *Molecular Ecology Notes*, **6**, 974–979.

Kimura M (1986) DNA and the Neutral Theory. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **312**, 343–354.

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

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**, 175–186.

Komdeur J, Burke T, Dudgale H, Richardson DS (2014) *Seychelles warblers: the complexities of the helping paradox.* (JL Koenig, WD Dickinson, Eds,). Cambridge University Press.

Kryazhimskiy S, Plotkin JB (2008) The population genetics of dN/dS. *PLoS genetics*, **4**, e1000304.

Lachish S, Knowles SC, Alves R, Wood MJ, Sheldon BC (2011) Fitness effects of endemic malaria infections in a wild bird population: the importance of ecological structure. *Journal of Animal Ecology*, **80**, 1196–1206.

Lampert KP, Fischer P, Schartl M (2009) Major histocompatibility complex variability in the clonal Amazon molly, Poecilia formosa: is copy number less important than genotype? *Molecular Ecology*, **18**, 1124–1136.

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

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

Librado P, Rozas J (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics*, **25**, 1451–2.

Marshall OJ (2004) PerlPrimer: cross-platform, graphical primer design for standard, bisuplahte and real-time PCR. *Bioinformatics*, **20**, 2471–2472.

Marzal A, García-Longoria L, Cárdenas Callirgos JM, Sehgal RN (2015) Invasive avian malaria as an emerging parasitic disease in native birds of Peru. *Biological Invasions*, **17**, 39–45.

Marzal A, De Lope F, Navarro C, Møller AP (2005) Malarial parasites decrease reproductive success: An experimental study in a passerine bird. *Oecologia*, **142**, 541–545.

McDonald B a, Linde C (2002) Pathogen population genetics, evolutionary potential, and durable resistance. *Annual Review of Phytopathology*, **40**, 349–379.

Mitchell-Olds T, Willis JH, Goldstein DB (2007) Which evolutionary processes influence natural genetic variation for phenotypic traits? *Nature Reviews Genetics*, **8**, 845–856.

Mukherjee S, Sarkar-roy N, Wagener DK, Majumder PP (2009) Signatures of natural selection are not uniform across genes of innate immune system , but purifying. *PNAS*, **106**, 7073–7078.

Murrell B, Moola S, Mabona A *et al.* (2013) FUBAR: a fast, unconstrained bayesian approximation for inferring selection. *Molecular Biology and Evolutioniology and evolution*, **30**, 1196–205.

Murrell B, Wertheim JO, Moola S *et al.* (2012) Detecting individual sites subject to episodic diversifying selection. *PLoS genetics*, **8**, e1002764.

Nakajima T, Ohtani H, Satta Y *et al.* (2008) Natural selection in the TLR-related genes in the course of primate evolution. *Immunogenetics*, **60**, 727–35.

Nerren JR, Swaggerty CL, MacKinnon KM *et al.* (2009) Differential mRNA expression of the avian-specific toll-like receptor 15 between heterophils from Salmonella-susceptible and resistant chickens. *Immunogenetics*, **61**, 71–7.

Newman D, Pilson D (1997) Increased Probability of Extinction Due to Decreased Genetic Effective Population Size : Experimental Populations of Clarkia pulchella. *Evolution*, **51**, 354–362.

Nowak M a, Tarczy-Hornoch K, Austyn JM (1992) The optimal number of major histocompatibility complex molecules in an individual. *Proceedings of the National Academy of Sciences of the United States of America*, **89**, 10896–10899.

Ohta T (2002) Near-neutrality in evolution of genes and gene regulation. *Proceedings of the National Academy of Sciences USA*, **99**, 16134–16137.

Van Oosterhout C (2009) A new theory of MHC evolution: beyond selection on the immune genes. *Proceedings. Biological sciences / The Royal Society*, **276**, 657–65.

Pedersen AB, Greives TJ (2008) The interaction of parasites and resources cause crashes in a wild mouse population. *Journal of Animal Ecology*, **77**, 370–377.

Plaut SM, Ader R, Friedman SB, Ritterson AL (1969) Social Factors and Resistance to Malaria in the Mouse. *Psychomatic Medicine*, 536–552.

Radwan J, Zagalska-Neubauer M, Cichon M *et al.* (2012) MHC diversity, malaria and lifetime reproductive success in collared flycatchers. *Molecular Ecology*, **21**, 2469–2479.

Raymond M, Rousset F (1995) GENEPOP (Version 1.2): Population Genetics Software for Exact Tests and Ecumenicism. *Heredity*, **86**, 248–249.

Redpath SM, Mougeot F, Leckie FM, Elston D a, Hudson PJ (2006) Testing the role of parasites in driving the cyclic population dynamics of a gamebird. *Ecology Letters*, **9**, 410–418.

Richardson DS, Burke T, Komdeur J (2002) Direct benefits and the evolution of female-biased cooperative breeding in Seychelles warblers. *Evolution*, **56**, 2313–21.

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

Richardson DS, Burke T, Komdeur J (2007) Grandparent helpers: The adaptive significance of older, postdominant helpers in the Seychelles warbler. *Evolution*, **61**, 2790–2800.

Richardson DS, Jury FL, Blaakmeer K, Komdeur J, Burke T (2001) Parentage assignment and extra-group paternity in a cooperative breeder: the Seychelles warbler (Acrocephalus sechellensis). *Molecular Ecology*, **10**, 2263–73.

Richardson DL, Jury FL, Dawson DA *et al.* (2000) Fifty Seychelles warbler (Acrocephalus sechellensis) microsatellite loci polymorphic in Sylviidae species and their cross-species amplification in other passerine birds. *Molecular Ecology*, **9**, 2155–7.

Richardson DS, Komdeur J, Burke T (2004) Inbreeding in the Seychelles Warbler : Environment-Dependent Maternal Effects. *Evolution*, **58**, 2037–2048.

Richardson DS, Komdeur J, Burke T, Lusk C (2003b) Altruism and infidelity among warblers. *Nature*, **422**, 580–581.

Van Riper III C, van Riper SG, Goff ML, Laird M (1986) The epizootiology and ecological significance of malaria in Hawaiian land birds. *Ecological Monographs*, **56**, 327–344.

Roach JC, Glusman G, Rowen L *et al.* (2005) The evolution of vertebrate Toll-like receptors. *Proceedings of the National Academy of Sciences USA*, **102**, 9577–9582.

Roberts ML, Buchanan KL, Evans MR (2004) Testing the immunocompetence handicap hypothesis: A review of the evidence. *Animal Behaviour*, **68**, 227–239.

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

Safford R, Hawkins F (2013) *The Birds of Africa, Volume 8: The Malagasy Region*.

Schnare M, Barton GM, Holt a C *et al.* (2001) Toll-like receptors control activation of adaptive immune responses. *Nature Immunology*, **2**, 947–50.

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

Shen T, Xu S, Wang X *et al.* (2012) Adaptive evolution and functional constraint at TLR4 during the secondary aquatic adaptation and diversification of cetaceans. *BMC evolutionary biology*, **12**, 39.

Slade RW, McCallum HI (1992) Overdominant vs. Frequency-Dependent Selection at MHC Loci. *Genetics*, **132**, 861–862.

Song X, Jin P, Qin S, Chen L, Ma F (2012) The evolution and origin of animal Toll-like receptor signaling pathway revealed by network-level molecular evolutionary analyses. *PLoS One*, **7**, e51657.

Spielman D, Brook BW, Frankham R (2004) Most species are not driven to extinction before genetic factors impact them. *Proceedings of the National Academy of Sciences of the United States of America*, **101**, 15261–15264.

Spurgin LG, Richardson DS (2010) How pathogens drive genetic diversity: MHC, mechanisms and misunderstandings. *Proceedings of the Royal Society B*, **277**, 979–988.

Spurgin LG, Wright DJ, van der Velde M *et al.* (2014) Museum DNA reveals the demographic history of the endangered Seychelles warbler. *Evolutionary Applications*, 1–10.

Stahl E a, Bishop JG (2000) Plant – pathogen arms races at the molecular level. *Current Opinion in Plant Biology*, **3**, 299–304.

Stephens M, Donnelly P (2003) A comparison of bayesian methods for haplotype reconstruction from population genotype data. *American Journal of Human Genetics*, **73**, 1162–9.

Szollosi E, Cichon M, Eens M *et al.* (2011) Determinants of distribution and prevalence of avian malaria in blue tit populations across Europe: separating host and parasite effects. *Journal of Evolutionary Biology*, **24**, 2014–2024.

Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Molecular Biology and Evolution*, **24**, 1596–9.

Thomas NJ, Hunter DB, Atkinson CT (2008) *Infectious Diseases of Wild Birds.* John Wiley & Sons.

Waldenström J, Bensch S, Hasselquist D, Östman O (2004) A new nested Polymerase Chain Reaction method very efficient in detecting Plasmodium and Haemoproteus infections from avian blood. *Journal of Parasitology*, **90**, 191–194.

Wegner KM, Reusch TBH, Kalbe M (2003) Multiple parasites are driving major histocompatibility complex polymorphism in the wild. *Journal of Evolutionary Biology*, **16**, 224–232.

Weir BS, Cockerham CC (1984) Estimating F-Statistics for the Analysis of Population Structure. *Evolution*, **38**, 1358–1370.

Wiley RH, Rabenold KN (1984) The evolution of cooperative breeding by delayed reciprocity and queuing for favourable social positions. *Evolution*, **38**, 609–621.

Wood MJ, Cosgrove CL, Wilkin TA *et al.* (2007) Within-population variation in prevalence and lineage distribution of avian malaria in blue tits, Cyanistes caeruleus. *Molecular Ecology*, **16**, 3263–3273.

Wright DJ (2014a) Evolutionary and conservation genetics of the Seychelles warbler (Acrocephalus sechellensis). University of East Anglia.

Wright DJ (2014b) Evolutionary and conservation genetics of the Seychelles warbler (Acrocephalus sechellensis). *Thesis*.

Wright DJ, Spurgin LG, Collar NJ *et al.* (2014) The impact of translocations on neutral and functional genetic diversity within and among populations of the Seychelles warbler. *Molecular Ecology*, **23**, 2165–2177.

Supplementary material

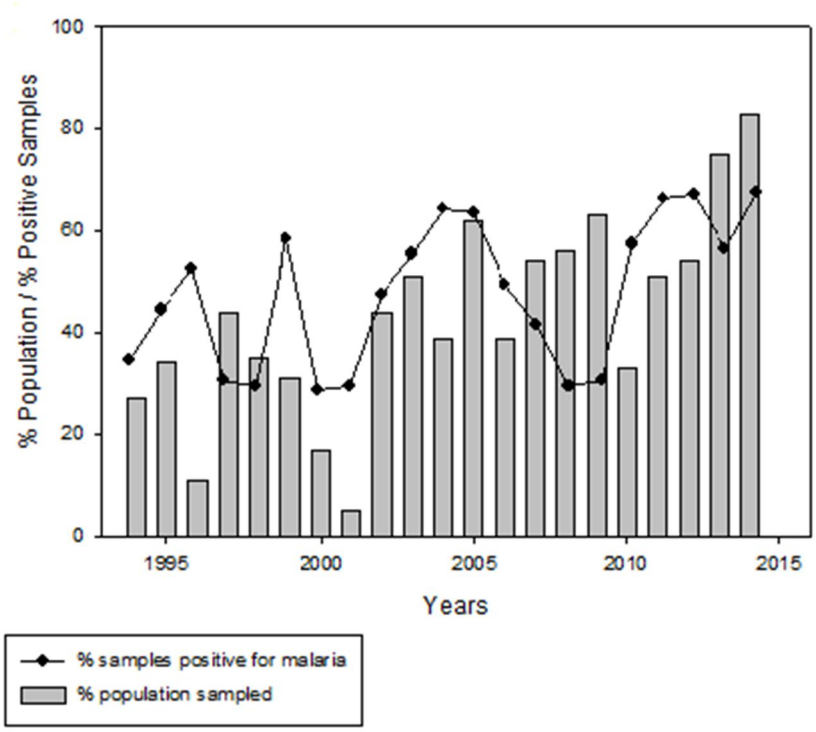


Figure S1. Annual malarial prevalence in the Cousin Island population of Seychelles warblers including individuals of all age classes and all sexes.

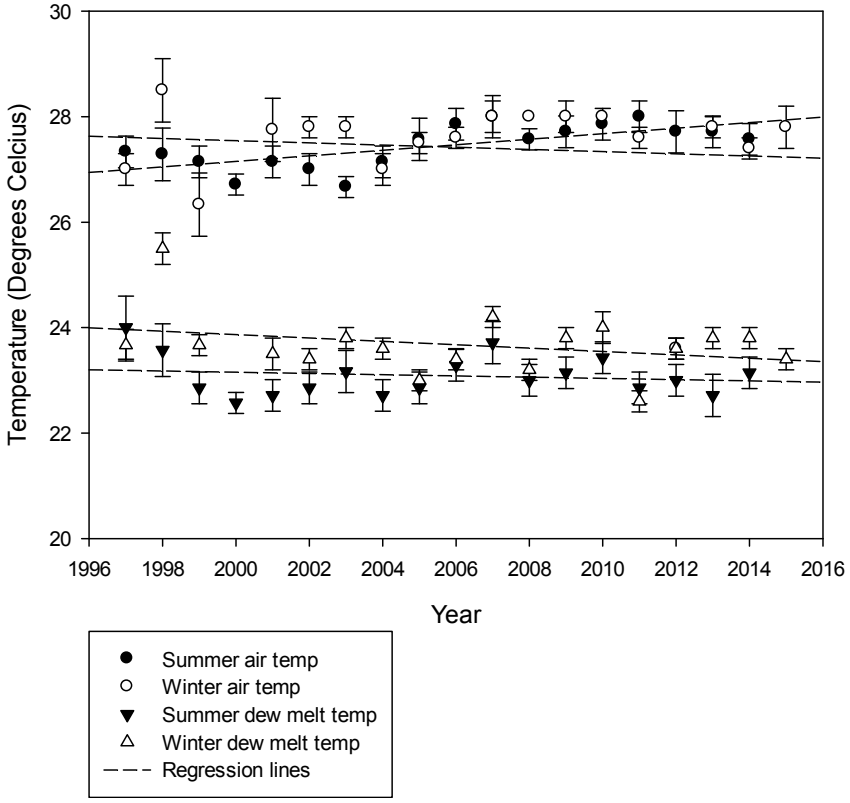


Figure S2. Annual mean air temperature and dew melting temperature (measure of moisture in the air) from 1997 to 2015 for both April-October and November-March breeding seasons for the Seychelles warbler on Cousin Island (ca 2km from Praslin Island, the source of this weather data).

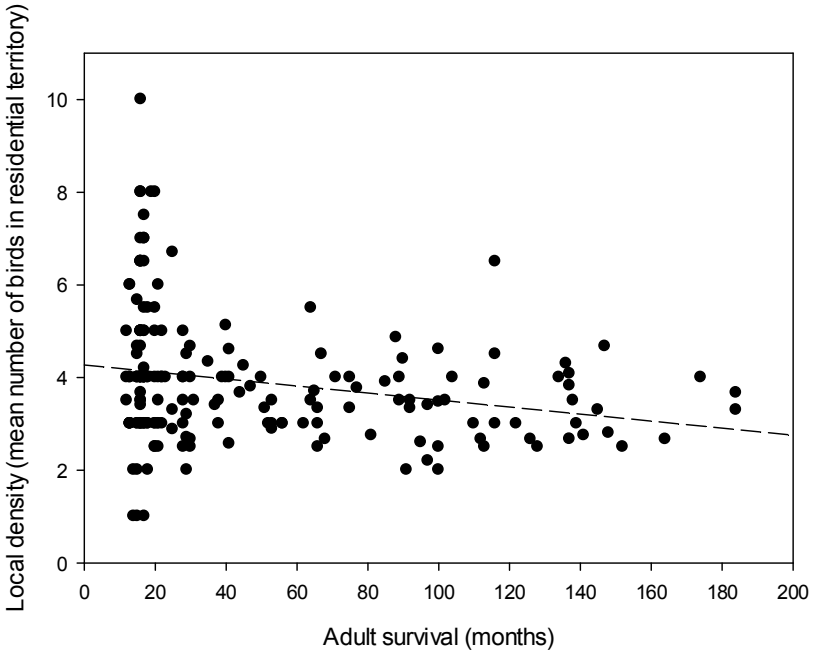


Figure S3. Local density influencing adult survival (months) within a cohort of Seychelles warblers.

Table S1. Microsatellite loci and Toll-like receptor loci genotyped in this study, with total number of individuals (*N*), number of heterozygous and homozygous individuals (N-Het and N-Hom, respectively), number of alleles (A), observed and expected heterozygosity (H_O and H_E, respectively).

Table S2. Investigating TLR variation associations with early-life malarial infection in the Seychelles warbler, using generalised linear mixed models which include both genetic and ecological fixed factors. All models were built with a logit link function and binomial error structure. Codes applied to *P*-values to show significance are as follows: . (<0.1) * (<0.05) ** (<0.01) and *** (< 0.001).

Table S3. Investigating TLR variation associations with average lifetime malarial infection in the Seychelles warbler, using generalised linear mixed models which include both genetic and ecological fixed factors. All models were built with a quasi-binomial error structure and included birth year as a random factor. Codes applied to *P*-values to show significance are as follows: . (<0.1) * (<0.05) ** (<0.01) and *** (< 0.001).

Table S4. Investigating TLR variation associations with juvenile survival in the Seychelles warbler, using generalised linear mixed models which include both genetic and ecological fixed factors. All models were built with a logit link function and binomial error structure. Codes applied to *P*-values to show significance are as follows: . (<0.1) * (<0.05) ** (<0.01) and *** (< 0.001).

Table S5. Investigating TLR variation associations with adult survival in the Seychelles warbler, using generalised linear mixed models which include both genetic and ecological fixed factors. All models included birth year as a random factor. Codes applied to *P*-values to show significance are as follows: . (<0.1) * (<0.05) ** (<0.01) and *** (< 0.001).

Table S6. Investigating TLR variation associations with adult survival in the Seychelles warbler, using generalised linear mixed models of survival over time fitted to a log-normal distribution curve using the surv package in R (R Core Team 2015). Codes applied to *P*-values to show significance are as follows: . (<0.1) * (<0.05) ** (<0.01) and *** (< 0.001).

Table S1.

Marker Locus	N	N-Het	N-Hom	A	H_O	H_E
10	205	95	110	3	0.463	0.474
13	205	97	108	3	0.473	0.521
18	205	100	105	4	0.488	0.485
25	205	136	69	6	0.663	0.719
27	205	130	75	6	0.634	0.656
35	205	124	81	3	0.605	0.608
37	205	81	124	3	0.395	0.436
4	205	88	117	2	0.429	0.411
42	205	58	147	2	0.283	0.277
48	194	52	142	4	0.268	0.664
56	205	73	132	3	0.356	0.403
58	205	152	53	5	0.742	0.702
6	205	141	64	4	0.688	0.694
9	116	16	100	4	0.138	0.396
Ase-11	205	99	106	3	0.483	0.473
Ase-16	205	148	57	7	0.722	0.749
Ase-19	205	105	100	2	0.512	0.489
Ase-22	205	78	127	2	0.381	0.375
Ase-3	205	87	118	3	0.424	0.444
Ase-38	205	82	113	2	0.445	0.479
Ase-53	205	100	105	4	0.488	0.559
Ase-55-cest	205	78	127	2	0.381	0.407
Ase-61	204	97	107	5	0.476	0.518
Ase-64	204	134	70	3	0.657	0.632
Ase-7	205	96	109	2	0.468	0.460
Calex-08-gga	205	42	163	2	0.205	0.207
Cuu4-gga5	205	85	120	2	0.415	0.474
Pdoμ6	205	119	86	3	0.581	0.564
PmaTGA	205	101	104	3	0.493	0.499
Pte24-cest	205	7	198	2	0.034	0.034
Mean Microsatellites	**205**	**93**	**108**	**3.3**	**0.460**	**0.494**
TLR15	**205**	**59**	**146**	**6**	**0.288**	**0.270**

Table S2.

Explanatory variables	N	Significant variables	Estimate	Std. Error	Z	*P*- value	2.5% CI	97.5% CI
TLR15 genotype, microsatellite Hz, sex, season-born	205	*TLR15* genotype- 'AB'	0.186	0.404	0.461	0.645	-0.620	0.975
		TLR15 genotype- 'AC'	1.231	0.552	2.228	0.026*	0.176	2.375
		Microsatellite Hz	0.153	0.711	0.215	0.830	-1.241	1.561
		Sex	0.238	0.314	0.757	0.449	-0.375	0.861
		Season born	1.544	0.369	4.187	0.000***	0.838	2.291
TLR15 Hm / Hz, microsatellite Hz, sex, season-born	205	*TLR15* Hm / Ht	-0.597	0.329	-1.815	0.070 .	-1.246	0.047
		Microsatellite Hz	0.266	0.677	0.393	0.694	-1.057	1.610
		Sex	0.132	0.305	0.432	0.666	-0.465	0.735
		Season born	1.446	0.358	4.039	0.000***	0.759	2.169
TLR15 hap B, *TLR15* hap C, microsatellite Hz, sex, season-born	110	Haplotype B	0.282	0.393	0.718	0.473	-0.499	1.050
		Haplotype C	1.235	0.549	2.250	0.024*	0.188	2.372
		Microsatellite Hz	0.142	0.693	0.205	0.838	-1.216	1.513
		Sex	0.173	0.306	0.566	0.572	-0.425	0.778
		Season born	1.440	0.359	4.008	0.000***	0.750	2.166
TLR15 genotype, season-born, *Ase-ua4,* MHC diversity	110	*TLR15* genotype- 'AB'	-0.096	0.571	-0.168	0.867	-1.236	1.026
		TLR15 genotype- 'AC'	1.253	0.692	1.812	0.070.	-0.051	2.714
		Season born	1.894	0.485	3.905	0.000***	0.983	2.901
		Ase-ua4	-1.306	0.798	-1.636	0.102	-3.021	0.189
		MHC diversity	0.012	0.145	0.081	0.935	-0.274	0.300
TLR15 hap C, season-born, *Ase-ua4*, MHC diversity	110	Haplotype C	1.159	0.670	1.730	0.084 .	-0.103	2.579
		Season born	1.785	0.469	3.807	0.000***	0.902	2.755
		Ase-ua4	-0.882	0.720	-1.225	0.221	-1.236	0.493
		MHC diversity	0.001	0.142	0.006	0.995	-0.051	0.282
TLR15 hap C, season-born, sex, territory quality	65	Haplotype C	0.425	0.969	0.439	0.661	-1.510	2.424
		Season born	0.995	0.669	1.487	0.137	-0.304	2.340
		Sex	-0.324	0.592	-0.547	0.585	-1.515	0.829
		Territory quality	0.131	0.056	2.361	0.018*	0.046	0.288
TLR15 genotype, season-born, sex, territory quality	65	*TLR15* genotype- 'AB'	-0.057	0.896	-0.064	0.949	-1.897	1.718
		TLR15 genotype- 'AC'	0.521	0.992	0.525	0.599	-1.454	2.556
		Season born	1.014	0.700	1.448	0.148	-0.347	2.420
		Sex	-0.471	0.605	-0.777	0.437	-1.693	0.703
		Territory quality	0.129	0.056	2.323	0.020*	0.043	0.285

Table S3.

Explanatory variables	N	Significant variables	Estimate	Std. Error	t	*P*-value	2.5% CI	97.5% CI
TLR15 genotype, Microsatellite Hz, early-life malaria, *TLR15* genotype* early-life malaria, sex, sex* early-life malaria, local density, season-born	191	*TLR15* genotype- 'AB'	-0.158	0.440	-0.360	0.719	-1.037	0.697
		TLR15 genotype- 'AC'	0.549	0.913	0.601	0.549	-1.286	2.488
		Microsatellite Hz	0.297	0.573	0.518	0.605	-0.828	1.425
		Early-life malaria	0.508	0.439	1.158	0.249	-0.351	1.375
		'AB' * early-life malaria	-0.871	0.864	-1.008	0.315	-2.725	0.750
		'AC'* early-life malaria	-0.840	1.053	-0.797	0.426	-2.988	1.289
		Sex	0.530	0.349	1.519	0.131	-0.150	1.222
		Sex(Male)* early-life malaria	-1.112	0.545	-2.039	0.043*	-2.188	-0.046
		Local density	-0.140	0.130	-1.077	0.283	-0.404	0.110
		Season-born	0.422	0.331	1.273	0.205	-0.229	1.074
TLR15 hm / hz, Microsatellite Hz, early-life malaria, *TLR15* genotype* early-life malaria, sex, sex* early-life malaria, local density, season-born	191	*TLR15* Hm / Hz	-0.215	0.385	-0.558	0.578	-0.974	0.543
		Microsatellite Hz	0.297	0.573	0.518	0.605	-0.828	1.425
		Early-life malaria	0.508	0.439	1.158	0.249	-0.351	1.375
		'AB'* early-life malaria	-0.871	0.864	-1.008	0.315	-2.725	0.750
		'AC'* early-life malaria	-0.805	1.057	-0.761	0.448	-2.988	1.289
		Sex	0.530	0.349	1.519	0.131	-0.150	1.222
		Sex(M)* early-life malaria	-1.112	0.545	-2.039	0.043*	-2.188	-0.047
		Local density	-0.140	0.130	-1.077	0.283	-0.404	0.110
		Season-born	0.422	0.331	1.273	0.205	-0.229	1.074
TLR15 hap B, *TLR15* hap C, Microsatellite Hz, early-life malaria, *TLR15* genotype* early-life malaria, sex, sex* early-life malaria, local density, season-born	191	Haplotype B	-0.411	0.344	-1.193	0.234	-1.106	0.252
		Haplotype C	0.549	0.913	0.601	0.549	-1.286	2.488
		Microsatellite Hz	0.297	0.573	0.518	0.605	-0.828	1.425
		Early-life malaria	0.508	0.439	1.158	0.249	-0.351	1.375
		'AB'* early-life malaria	-0.871	0.864	-1.008	0.315	-2.725	0.750
		'AC'* early-life malaria	-0.805	1.057	-0.761	0.448	-2.988	1.289
		Sex	0.530	0.349	1.519	0.131	-0.150	1.222
		Sex(M)* early-life malaria	-1.112	0.545	-2.039	0.043*	-2.188	-0.046
		Local density	-0.140	0.130	-1.077	0.283	-0.404	0.110
		Season-born	0.422	0.331	1.273	0.205	-0.229	1.074
Early-life malaria, sex, early-life malaria *sex, *Ase-ua4,* MHC diversity	103	Early-life malaria	0.782	0.505	1.548	0.125	-0.198	1.794
		Sex	0.144	0.572	0.252	0.802	-0.988	1.272
		Early-life malaria*sex	-1.430	0.747	-1.914	0.059 .	-2.908	0.030
		Ase-ua4	0.441	0.528	0.835	0.406	-0.603	1.489
		MHC diversity	-0.105	0.138	-0.761	0.449	-0.376	0.167
Early-life malaria *sex, *Ase-ua4,* MHC diversity, territory quality	59	Early-life malaria	2.285	1.054	2.168	0.035*	0.312	4.527
		Sex	0.707	0.891	0.793	0.431	-1.001	-2.571
		Early-life malaria*sex	-2.278	1.103	-2.065	0.044*	-4.544	-0.175
		Ase-ua4	0.190	0.795	0.239	0.812	-1.435	1.743
		MHC diversity	-0.250	0.201	-1.244	0.219	-0.663	0.136
		Territory quality	-0.043	0.041	-1.032	0.307	-0.132	0.036

Table S4.

Explanatory variables	N	Significant variables	Estimate	Std. Error	Z	*P*- value	2.5% CI	97.5% CI
TLR15 genotype, Microsatellite Hz, early-life malaria, season-born	205	*TLR15* genotype-'AB'	-0.085	0.426	-0.200	0.842	-0.909	0.772
		TLR15 genotype-'AC'	-0.904	0.573	-1.579	0.114	-2.034	0.241
		Microsatellite Hz	0.778	0.740	1.050	0.294	-0.666	2.251
		Early-life malaria	1.501	0.384	3.908	0.000***	0.777	2.293
		Season born	0.368	0.432	0.852	0.394	-0.462	1.250
TLR15 Hm / Hz, Microsatellite Hz, early-life malaria, season-born	205	*TLR15* hm / hz	0.311	0.353	0.881	0.378	-0.386	1.003
		Microsatellite Hz	0.701	0.716	0.979	0.327	-0.698	2.123
		Early-life malaria	1.419	0.369	3.842	0.000***	0.720	2.176
		Season born	0.370	0.424	0.872	0.383	-0.443	1.235
TLR15 hap B, *TLR15* hap C, Microsatellite Hz, early-life malaria, season-born	205	Haplotype B	-0.076	0.423	-0.179	0.858	-0.893	0.776
		Haplotype C	-0.920	0.571	-1.610	0.107	-2.046	0.223
		Microsatellite Hz	0.831	0.731	1.137	0.256	-0.594	2.287
		Early-life malaria	1.497	0.380	3.937	0.000***	0.781	2.281
		Season born	0.379	0.427	0.889	0.374	-0.438	1.250
Early-life malaria, *Ase-ua4,* MHC diversity	110	Early-life malaria	1.726	0.450	3.833	0.000***	0.872	2.648
		Ase-ua4	0.320	0.714	0.448	0.654	-1.051	1.807
		MHC diversity	0.213	0.145	1.472	0.141	-0.068	0.503
Early-life malaria, *Ase-ua4,* MHC diversity, territory quality	65	Early-life malaria	3.236	0.864	3.743	0.000***	1.717	5.239
		Ase-ua4	0.091	1.035	0.088	0.930	-1.904	2.302
		MHC diversity	0.0176	0.212	0.083	0.934	-0.404	0.438
		Territory quality	-0.058	0.035	-1.667	0.096 .	-0.131	0.007

Table S5.

Explanatory variables	N	Significant variables	Estimate	Std. Error	t	*P*- value	2.5% CI	97.5% CI
TLR15 genotype, Microsatellite Hz, early-life malaria, lifetime malaria, local density	205	*TLR15* genotype- 'AB'	-4.891	8.237	-0.594	0.553	-21.035	11.254
		TLR15 genotype- 'AC'	-0.582	10.875	-0.054	0.957	-21.896	20.732
		Microsatellite Hz	-4.473	14.496	-0.309	0.758	-32.884	23.937
		Early-life malaria	6.641	6.656	0.998	0.320	-6.404	19.685
		Lifetime malaria	14.738	7.557	1.950	0.053	-0.0731	29.549
		Local density	-5.756	2.370	-2.429	0.016*	-10.400	-1.111
TLR15 Hm / Hz, Microsatellite Hz, early-life malaria, lifetime malaria, local density	205	*TLR15* Hm / Hz	1.617	6.789	0.238	0.812	-11.689	14.923
		Microsatellite Hz	-3.215	14.105	-0.228	0.820	-30.860	24.430
		Early-life malaria	5.939	6.514	0.912	0.363	-6.828	18.706
		Lifetime malaria	16.272	7.418	2.193	0.030	1.732	30.812
		Local density	-6.353	2.327	-2.729	0.007**	-10.914	-1.791
TLR15 hap B, *TLR15* hap C, Microsatellite Hz, early-life malaria, lifetime malaria, local density	205	Haplotype B	-8.707	8.230	-1.058	0.291	-24.836	7.423
		Haplotype C	-1.665	10.872	-0.153	0.878	-22.975	19.644
		Microsatellite Hz	-4.600	14.253	-0.323	0.747	-32.535	23.335
		Early-life malaria	5.928	6.581	0.901	0.369	0.723	30.054
		Lifetime malaria	15.388	7.483	2.057	0.041*	0.723	30.054
		Local density	-6.385	2.327	-2.743	0.007**	-10.946	-1.823
Lifetime malaria, local density, *Ase-ua4,* MHC diversity	110	Lifetime malaria	15.401	10.045	1.533	0.129	-3.690	35.289
		Local density	-4.518	3.187	-1.418	0.160	-10.997	1.042
		Ase-ua4	17.331	13.608	1.274	0.206	-8.471	44.262
		MHC diversity	6.500	2.732	2.379	0.019*	1.165	11.790
Lifetime malaria, local density, *Ase-ua4,* MHC diversity	65	Lifetime malaria	32.011	14.896	2.149	0.036*	-3.690	35.289
		Local density	-6.742	4.492	-1.501	0.139	-10.997	1.042
		Ase-ua4	20.794	19.545	1.064	0.292	-8.471	44.262
		MHC diversity	7.482	3.932	1.903	0.063 .	1.165	11.790
		Territory quality	-0.313	0.657	-0.477	0.635	-18.545	10.088

Table S6.

Explanatory variables	N	Significant variables	Estimate	Std. Error	Z	*P*- value	2.5% CI	97.5% CI
TLR15 genotype, Microsatellite Hz, early-life malaria, lifetime malaria, local density	205	*TLR15* genotype- 'AB'	0.266	0.786	0.338	0.735	-1.274	1.806
		TLR15 genotype- 'AC'	0.286	0.796	0.359	0.720	-1.274	1.846
		Microsatellite Hz	0.017	0.263	0.065	0.948	-0.499	0.533
		Early-life malaria	0.131	0.121	1.085	0.278	-0.106	0.368
		Lifetime malaria	0.313	0.137	2.277	0.023*	0.043	0.582
		Local density	-0.125	0.043	-2.895	0.004**	-0.209	-0.040
TLR15 Hm / Hz, Microsatellite Hz, early-life malaria, lifetime malaria, local density	205	*TLR15* Hm / Hz	-0.027	0.124	-0.215	0.830	-0.271	0.217
		Microsatellite Hz	0.010	0.259	0.038	0.970	-0.497	0.516
		Early-life malaria	0.122	0.119	1.025	0.305	-0.112	0.356
		Lifetime malaria	0.321	0.136	2.359	0.018*	0.054	0.587
		Local density	-0.136	0.043	-3.194	0.001**	-0.220	-0.053
TLR15 hap B, *TLR15* hap C, Microsatellite Hz, early-life malaria, lifetime malaria, local density	205	Haplotype B	-0.121	0.151	-0.803	0.422	-0.416	0.174
		Haplotype C	-0.051	0.199	-0.255	0.799	-0.441	0.339
		Microsatellite Hz	-0.005	0.261	-0.020	0.984	-0.516	0.506
		Early-life malaria	0.125	0.120	1.036	0.300	-0.111	0.361
		Lifetime malaria	0.310	0.137	2.263	0.024*	0.041	0.578
		Local density	-0.137	0.043	-3.212	0.001**	-0.220	-0.053
Lifetime malaria, local density, *Ase-ua4,* MHC diversity	110	Lifetime malaria	0.352	0.181	1.947	0.052 .	-0.002	0.706
		Local density	-0.109	0.056	-1.957	0.050 .	-0.219	0.000
		Ase-ua4	0.195	0.245	0.798	0.425	-0.284	0.674
		MHC diversity	0.107	0.049	2.175	0.030*	0.011	0.204
Lifetime malaria, local density, *Ase-ua4,* MHC diversity	65	Lifetime malaria	0.747	0.246	3.043	0.002**	0.266	1.228
		Local density	-0.145	0.074	-1.960	0.049 *	-0.290	0.000
		Ase-ua4	0.285	0.322	0.884	0.377	-0.347	0.916
		MHC diversity	0.109	0.065	1.685	0.092 .	-0.018	0.236
		Territory quality	-0.006	0.011	-0.587	0.557	-0.028	0.015

Chapter 6: General Discussion

Genetics will disappear as a separate science because,

in the 21st century, everything in biology will become gene-based,

and every biologist will be a geneticist (Sydney Brenner).

In this thesis I have highlighted the importance of exploring evolutionary forces and how they shape genetic variation within natural populations. I explored how different groups of immune genes, all with pivotal roles in innate immune defence from pathogens, may have evolved in response to demographic and selective drivers in a bottlenecked island population of the Seychelles warbler. Here in this chapter, I discuss my findings in a collective context and outline potential avenues for future research.

6.1 Comparative evolution of different immune genes

Understanding the relative roles of demographic and selective forces in shaping genetic variation has become a central focus of evolutionary biology (Lande 1976). In particular, we need to understand how variation is driven, maintained and eroded at functional loci that are linked to individual fitness in terms of natural selection (Darwin 1859; Klein 1986; Takahata *et al.* 1992)**.** This is important in bottlenecked or fragmented populations where a demographic event has reduced diversity across the genome and can have severe genetic consequences. Immune genes are ideal candidates for this type of study because of the vital role they play in combating pathogens in wild populations (Anderson & May 1978; May & Anderson 1983). All living things will encounter familiar and / or novel pathogens continuously throughout their lifetime and the immune gene repertoire is critical in determining how an individual responds to initial and subsequent exposures to infections. Variation at these immune loci should theoretically enable an individual to better combat any given pathogen, but also a greater variety of pathogens and thus improve individual fitness. At the population level within and among individual variation will enhance long-term viability and persistence (O'Brien & Evermann 1988). However, immune loci can differ

considerably in their structure, function and mechanisms of mutation, and this ultimately influences how their evolution is shaped by different deterministic forces. In this thesis, I examined two gene groups from the vertebrate innate immune system: β-defensins (specifically in avian β-defensins (AvBDs), Chapter 2) and Toll-like receptors (TLRs, Chapters 3 - 5) in the Seychelles warbler (SW); a recently bottlenecked species endemic to the Seychelles archipelago.

My results showed considerable similarities and differences in the evolution of these two very different innate immune gene groups. Firstly, I found that higher levels of polymorphism had been maintained at the TLRs in comparison to the AvBDs in the face of the recent bottleneck event in the SW population. AvBDs were initially chosen as a candidate group because of (i) their simple structure, (ii) their direct function and (iii) *a priori* hypothesis based on previous literature demonstrating an association between nucleotide variation and pathogen infection outcomes within an individual. However, little variation was observed across the six loci amplified in the SW. Furthermore, amplifying the shortest locus- *AvBD7*- in museum specimens of the SW pre-dating the bottleneck did not detect the greater level of variation that I had expected. This posed new considerations in explaining the role of AvBDs within this relatively pathogen-depauperate system.

In contrast to the AvBDs, five out of seven TLR loci had several polymorphisms, of which two loci *TLR3* and *TLR15* showed signatures of positive (balancing) selection. TLRs are structured similarly to the Major Histocompatability Complex (MHC). However the MHC has a complex evolutionary history involving repeat gene duplications and so an individual can possess multiple MHC loci. This presents a number of logistical challenges when doing allele-specific studies because the alleles have highly similar sequences (Bach 1976). The MHC is also burdened with many other problems, ironically with a main technical issue being that for the very reason it is attractive as a candidate gene group, its complexity can confound its study. Since most organisms are exposed to enormous numbers of pathogens, this in turn is characterised by a highly complex MHC. Therefore, when studies focus on a single exon of an MHC locus and its relationship with individual pathogens, the results are often mixed (for reviews, see Bernatchez & Landry 2003; Garrigan & Hedrick 2003)**.** It is unlikely to be possible to fully characterise the MHC and pathogen load in most study systems. The only feasible approach would be to focus on a simpler more tractable study system (Richardson

& Westerdahl 2003; Miller & Lambert 2004; Ejsmond & Radwan 2009), but this arguably could still have no application to more complex systems.

Like the MHC, the TLR molecule is a receptor molecule and has specific domains with specific functions. There is a cystolic domain which contains the Toll / Interleukin-1 receptor interaction in order to form the adaptor molecules needed to initiate a primary immune response (Vogel *et al.* 2003). Nakajima *et al.* (2008) show that TLRs across 25 different primate species are highly conserved at this domain, but are rapidly evolving in another region. This region is the extracellular domain made up of leucine-rich repeats that is responsible for recognising pathogens and this can be highly variable between different TLR molecules depending on what class of pathogen ligand they bind to and auto-regulate (Kawai & Akira 2006). This mirrors the peptide-binding region of the MHC and a number of studies show strong evidence of positive (balancing) selection at specific sites encoding for this domain (e.g. Alcaide & Edwards 2011; Areal *et al.* 2011; Grueber *et al.* 2014). There is a common effect of functional constraint, which Shen *et al.* (2012) show nicely across different cetacean species that dN / dS averaged across sequences is < 1, but there are radical changes at specific sites and parallel evolution between independent lineages that are indicative of balancing selection. It is this variation which permits specific TLR activation to be tailored to ensure efficient immune defence against invading pathogens (for review,, see Brownlie & Allan 2011).

There was no evidence of gene conversion at the TLR loci and so there were no problems with the recombination rate exceeding the mutation rate at these loci (Ohta 1995). When looking across the *Acrocephalus* genus at multiple species, results suggested that at least one recombination event has occurred in each of the four polymorphic TLR genes across the *Acrocephalus* warblers (*TLR1LB, TLR3, TLR4* and *TLR15*) (minimum number of recombination events identified between specific sites = 2, 1, 2, 2 respectively). It could have been expected to find evidence of gene conversion at the TLR loci, given that these alleles are duplicates (Roach *et al.* 2005). However, this was not the case and sequences were dissimilar enough to identify different alleles (Chapter 3). I used several TLR loci from different classes and this may partly explain why I did not detect evidence of gene conversion. It is still unclear why the structure of these classes within the TLR multigene family is so different and what role gene conversion played in their evolution. However, it is

evident that gene conversion plays a much greater role in the evolution of the MHC than at TLR loci.

My findings are concurrent with previous studies in that point mutation is the main source of genetic variation at TLRs (Barreiro *et al.* 2009; Alcaide & Edwards 2011) (and AvBDs (Hellgren & Ekblom 2010; Hellgren 2014)). Interestingly, point mutations are more likely to be deleterious compared to the entire sections of DNA copied by gene conversion events. The latter is more likely to be functional because it is being directly copied from one variant to another, and yet we found no evidence of GC at AvBD or TLR loci in the Seychelles warbler. This could simply be due to the fact that for the MHC, the genes are all very closely physically positioned and this allows for gene conversion to take place. However, it can also be argued that this is a consequence of different evolutionary forces at play.

I used a combination of approaches in order to identify the role of selection in the evolution of AvBDs and TLRs in the SW and tried to delineate the effects of selection from that of drift. Firstly, I used traditional population genetics statistical methods (Chapters 2 & 3) including haplotype-specific and PAML-based site-specific tests. These were based on either allele frequencies or dN / dS ratios. Within polymorphic TLR loci, even at loci with multiple non-synonymous variants (*TLR3* and *TLR15*), I failed to detect any overall signatures of selection using haplotype-specific whole sequence-based tests in the SW. This remained the case when considering sequences from several other *Acrocephalus* species. However, at the codon-level I did identify a number of sites under both positive (balancing) and negative (purifying) selection within the SW and across the *Acrocephalus* genus at all five polymorphic TLR loci. Next, I used a novel approach of designing forward-in-time computer simulations to delineate demographic effects from the effects of selection in order to get a better resolution of selective pressures when working with relatively low levels of variation at our candidate genes (Chapter 4). This was feasible because I had estimates for important parameters like effective population size (Spurgin *et al.* 2014) and mutation rate. My simulations suggested that weak balancing selection appears to have acted in the recent past on the five TLR genes. Finally, I conducted association analyses to determine whether the TLR variation observed at the most polymorphic of TLR loci caused differential effects in individual fitness (Chapter 5). I found a specific *TLR15* allele that increases resilience to a specific *Haemosporidian* strain (GRW1) in addition to giving the individual acquired

immunity to prevent future infections later in life. These results suggest that pathogen-mediated selection may at least partly explain the fact that some variation remains in the bottlenecked SW population.

By combining several different approaches I was able to compare my results and gain a more holistic view on how evolutionary forces were shaping contemporary variation. I was able to overcome the various limitations and difficulties imposed by using any one of these approaches exclusively. For example, testing for deviations from Hardy-Weinberg equilibrium or any test based on entire sequences can be confounded by null alleles, genotyping errors, population sub-structure and gene flow (migration). These can all cause false-positive results (for review, see Vasemagi & Primmer 2005). Tests based on allele frequency distributions such as Tajima's D test of neutrality, test for either selection or a change in population size, as it can be hard to disentangle the two from one another. Therefore, they can make strong inferences on population demographics but are not considered to be able to give robust inferences of selection (Zhai *et al.* 2009). Additionally, they often cannot delineate current from past selection given their large evolutionary timeframe, with the exception of some tests including Hardy-Weinberg deviation. Tests based on dN / dS ratios give limited evidence that a certain DNA variant or polymorphism *currently* has a direct phenotypic or fitness consequence by using a shorter evolutionary timeframe. Consequently by using different methods, I have been able to span these different time frames using: (i) sequence-based tests to examine long evolutionary past, (ii) novel forward-in-time computer simulations to examine evolution over the span of the bottleneck, and (iii) association analyses to look at selection occurring presently.

By applying multiple approaches to characterise immunogenetic variation in the SW, they mutually suggest that while weak selection is acting on TLR loci, drift is the overriding force shaping such variation. I was able to compare and contrast this pattern to a previous study that assessed population-level variation at TLRs in another island endemic bottlenecked species, the New Zealand robin *Petroica australis raikura* (Grueber *et al.* 2013). This study paralleled our results in finding drift was the overriding force shaping variation at different TLR loci. Grueber *et al.* (2013) expanded on this study by identifying a specific *TLR4* allele which conferred a survival advantage to individuals. We have previously identified a specific MHC allele which also increases individual survival in the SW (Brouwer

et al. 2010) and in this thesis, I show evidence of a specific *TLR15* allele indirectly conferring a survival advantage by providing resilience to disease (Chapter 5). My overall finding of dominating drift and consequential loss of adaptive genetic variation from a bottleneck is concordant with a number of excellent existing studies (e.g. Willi *et al.* 2007; Garcia de Leaniz *et al.* 2007; Ardia *et al.* 2011). Likewise, my conclusion that there are strong effects of drift in shaping variation post-bottleneck in the contemporary population is well-supported in the literature (Miller & Lambert 2004; Miller *et al.* 2010; Sutton *et al.* 2011; Grueber *et al.* 2013).

Although drift is an important evolutionary force in our system, the main focus of this thesis was to investigate whether selection could maintain specific (functional) variation in a bottlenecked avian population. AvBDs encode for anti-microbial peptides, which recognise a broad spectrum of pathogens and directly attack the pathogen cell wall (Hollox & Armour 2008; van Dijk *et al.* 2008; Derache *et al.* 2012)**.** Conversely, TLRs recognise specific PAMPs such as lipopolysaccharides, lipids, DNA and RNA fragments (Takeuchi *et al.* 2002; Boyd *et al.* 2007, 2012; Keestra *et al.* 2010). Consequently, TLR molecules are considerably complex in structure with distinct regions serving different roles. A number of studies have shown that sites encoding the peptide-binding region of the MHC are strongly selected for when sites encoding the 'stalk' are highly conserved (Aguilar *et al.* 2013; Sutton *et al.* 2013; Scherman *et al.* 2014). The exon that I screened in the SW encodes for the leucine-rich repeat region responsible for binding to PAMPs and within this receptor region of the molecule, my results inferred a number of sites to be structural in their role and sites directly associated with pathogen-ligand binding. These codon-specific findings have been mirrored in other studies in TLRs (Alcaide & Edwards 2011; Areal *et al.* 2011; Fornůsková *et al.* 2013; Grueber *et al.* 2014) and they all contribute to better understanding how selection operates to promote the variation responsible for optimising the TLRs function in immune defence.

It is widely-accepted that pathogens are the predominant selective pressure acting on immune genes and driving their diversity (for review, see Spurgin & Richardson 2010). Pathogens are not necessarily considered to be natural groups *per se,* but in fact a phylogenetically and antigenically diverse group of interacting organisms at both cellular and intracellular levels within the host (Nizet 2006)**.** Therefore, immunity itself will occur

differently with each pathogen or array of pathogens invoked, depending on determinants of cell surfaces (e.g. the structure of extracellular domains and their receptors), among other factors including demography, environment and gene-gene interactions (for example, the interaction of different immune gene families within the immune repertoire). By considering this polygenic nature of disease, identifying selection at specific sites directly responsible for these determinants gives us a much better idea of the role of selection in shaping patterns of variation in pathogen resistance.

Pathogen-mediated selection is just one type of selective force behind balancing selection that could be shaping immune gene variation. The MHC has already been linked to other forces, such as kin recognition and mate choice (Manning *et al.* 1992; Reusch *et al.* 2001; Bernatchez & Landry 2003; Richardson *et al.* 2005; Brouwer *et al.* 2010; Huchard *et al.* 2013). While there has been considerable work using TLRs in functional variation studies, it is still unknown whether they may be involved in mechanisms other than the direct immune response. More research is needed to elucidate the role of different selection pressures other than pathogen-mediated selection in shaping variation at these loci in natural populations.

6.2 An evolutionary conservation case study

To understand the long-term viability of a population or species, both the genetics and the ecology must be considered within an evolutionary framework in order to maximise use of the data for informing conservation bodies. In chapters 2 and 3, I establish that there is a relatively low level of immunogenetic variation in the bottlenecked SW. By comparing variation at the same loci across several *Acrocephalus* species with varying demographic histories, I found evidence of considerably more variation existing in mainland migratory species than in island bottlenecked species. This suggests that being from an insular endemic population has genetic consequences, particularly if that population has undergone a bottleneck. The SW's recent bottleneck has been well-documented over the last 150 years (Oustalet 1878; Crook 1960; Collar & Stuart 1985). Effective population size estimates indicate that the SW had a widespread distribution historically, covering the inner granitic islands of the Seychelles (Chapter 1, Figure 6). Furthermore, historic populations will

have dispersed between different islands and so gene flow would have had a role in shaping genetic variation, unlike the contemporary population. If the established 'rule' that a population's effective size is approximately one tenth of its census size is true (Frankham 2010) the SW may once have had a population of > 10 000. Given these figures, it appears that the bottleneck occurred relatively recently and over a longer time span than originally predicted (Spurgin *et al.* 2014). Thus, purging of genetic load is unlikely to have occurred in such a short time (Crnokrak & Barrett 2002). This needs to be considered when assessing variation in the contemporary population.

In this thesis, chapters 2 and 3 show that levels of functional variation are relatively low across two different immune gene groups. I proceeded to do a pre- and post-bottleneck comparison at a specific locus, *AvBD7*, to gain a better idea of how patterns of variation may have changed as a consequence of the bottleneck. It was possible to amplify this short β-defensin in 15 museum samples (< 120 bp) and only two polymorphisms (one synonymous and one non-synonymous) were observed in comparison to the single-nucleotide polymorphism (synonymous) observed in the post-bottleneck contemporary population. This could arguably be direct evidence of a loss of functional variation as a result of the bottleneck. However, when dealing with such low numbers of variants it is next to impossible to statistically assess this. Unfortunately, the DNA was low-quality and fairly degraded. This prevented us from screening the more polymorphic TLR genes. For chapter 5, I designed primers to amplify the key variable region within the *TLR15* sequence, the most polymorphic loci in this study. Unfortunately, this region was ca 400 bp and given that this was at the upper limit of the microsatellite fragment sizes that were genotyped (Wright 2014), amplifying this region in the museum samples was also not an option. The problem was further enhanced by the limited volume of museum DNA available from the specimens of which they were extracted (Spurgin *et al.* 2014). To attempt to design primers and carry out repeated genotyping to counter the high error rates that come with degraded DNA (Arandjelovic *et al.* 2009)**,** was not possible given the insufficient volumes of DNA sample available.

Extending on this work, it would be ideal to characterise functional immunogenetic variation within the different translocated populations, which have all been sourced from the Cousin population examined in this thesis. Translocations can cause bottlenecks

themselves due to small founder sizes and incomplete sampling from the source population (Jamieson 2011). Previous research on the SW has shown that the earlier translocations of just 29 founder individuals each to the islands of Aride and Cousine (in 1988 and 1991 respectively) resulted in limited but detectable genetic divergence (Wright 2014). Interestingly, the same study also found that diversity within the multiple Seychelles warbler populations was temporally stable, thus suggesting that drift has had minimal effect in further eroding genetic variation in the translocation populations since their establishment. This infers that the populations grow rapidly after translocation with little reproductive skew (Nei *et al.* 1975).

It is often the case that many translocations have an absence of follow-up studies on the source population from which the founding individuals were removed (Pertoldi *et al.* 2007). Yet in the SW system, this has been specifically assessed and the impact of the removal of these individuals on the genetic status of the source population has been shown to have no detrimental long-term effect (Wright 2014). My studies use individuals from the contemporary population after three translocations had already taken place (the latest being to Denis Island in 2004) and I can conclude with sufficient evidence that drift has had overriding effects in the SW Cousin population. What positive (balancing) selection I do detect is relatively low and my simulation forecast models predict that this selection will have little effect on the long-term variation at TLR loci (Figure 3, Chapter 4), unlike the bottleneck event which had huge negative effects on overall diversity (Figures 1 & 2, Chapter 4). Many studies which show neutral processes to be the main forces in shaping genetic variation in small populations, compare the levels of variation at both neutral markers and functional (or 'critical') markers representing adaptive variation (Kimura 1986; Alcaide 2010; Sutton *et al.* 2011; Agudo *et al.* 2012).

By carrying out association analyses, it is useful to see what factors directly influence fitness parameters such as individual survival, malarial prevalence and how individuals respond to malarial infection once exposed. My models focus on immunogenetic variation and its association with individual fitness, but we control for ecological factors. I found that a specific *TLR15* allele confers for resilience against malarial infection in early-life, which consequently results in acquired immunity for preventing secondary infections in later life. Given the significant role of lifetime malarial infection on adult survival, this means that this

allele has indirect survival advantages too. Other immunogenetic variables proved functional, with MHC diversity having direct effects on adult survival and the MHC allele *Ase-ua4* significantly influencing whether an individual dies or not once infected with GRW1. Unfortunately, some association analyses are limited when working with endangered wild species. It is not possible to look at mRNA expression, which has been done for TLRs in a number of studies investigating the direct relationship between TLR variation and immune response in model species *in vivo* and *in vitro*, such as in the chicken (*Gallus* species) (e.g. Higgs *et al.* 2006; Nerren *et al.* 2009, 2010) and in mice (e.g. Rehli 2002; Bihl *et al.* 2003). However, we can look at differential outcomes in response to individuals being infected and relatively assess the different patterns in relation to genetic variation, as we did in chapter 5. Although chapters 2-4 explore the genetic composition of the SW population and infer the evolutionary and demographic processes responsible, chapter 5 is informative for conservation biology from both a scientific knowledge and practical perspective. With the chapters combined, I hope to provide conservation stakeholders with novel and useful data that will hopefully be of relevance to other bottlenecked or fragmented populations of conservation concern, where a similar level of monitoring of the population in its natural state can permit in depth molecular ecological study.

6.3 Directions for future research

The research presented in this thesis characterises two families of immune genes in the SW and uses a combination of different approaches to infer the evolutionary forces which have shaped the variation observed within this population. To add to this, it is important to compare and contrast neutral variation with functional variation. This has already been done in the SW with regards to MHC diversity (Hansson & Richardson 2005). Although the MHC markers and microsatellites were not directly compared in this study, relatively, they both showed the same picture of genetic variation in the SW in that much variation had been lost across the genome as a consequence of the recent bottleneck. By comparing the SW to two other *Acrocephalus* species, they further showed that genetic variation in the SW was half to one third of that of its congeners, which has also been shown in a number of other studies (Komdeur *et al.* 1998; Richardson & Westerdahl 2003). It would be useful to

quantify patterns of neutral polymorphism in comparison to functional polymorphism in order to better understand the (potential) selection / drift dynamics at candidate loci.

By modelling microsatellite standardised heterozygosity with *TLR15* heterozygosity, I was able to determine an absence of any significant association between the two measures. This is informative in that it proves microsatellites are not sufficient in explaining variation in a natural population. However, this does not resolve the issue that we cannot say whether the polymorphism statistics for *TLR15* and other immune genes are different than those we would expect for anonymous loci. However, whilst microsatellite studies have been developed for the Seychelles warbler (Richardson *et al.* 2000) you cannot directly compare allele numbers / allelic richness between microsatellites and AvBDs / TLRs because the microsatellite markers designed for the SW were specifically chosen to be polymorphic, which presents a bias (Richardson *et al.* 2000). In order to gain a true neutral reference, I would need to screen another nuclear locus to assess how much non-functional synonymous variation exists in each population in all of our species and ideally, in the same individuals. Also, I would ideally need to screen more than one nuclear locus and then find a way distinguish whether our 'signatures' of selection is positive selection or whether it is just relaxed purifying selection / reduced efficiency of purifying selection, which can be expected in a bottlenecked population (Hughes 2007).

We amplified AvBD and TLR loci in a handful of individuals from other *Acrocephalus* species to simply assess the relative variation at these immune loci across the genus. This approach was used to increase our power to detect selection when using population genetic statistical tests by comparing patterns of selection across a set of ecologically-distinct species. We could improve our approach by obtaining more samples from these other *Acrocephalus* species and to increase our number of individuals screened in order to fully understand the evolutionary processes in operation at specific loci of interest. Associations between individual TLR genotypes and specific pathogens in wild populations, is yet to be explored. Only one blood parasite has been identified in the SW to date and no evidence has been found of any gastro-intestinal parasites (Hutchings 2009). However, it would be interesting to screen for other pathogens such as bacteria and viruses. The AvBDs have been shown to directly attack bacterial pathogens via the amphipathic properties of their encoded anti-microbial peptides. Therefore, assessing the relationship between bacterial

infection and AvBD loci, or viral infection and TLR loci, would provide further understanding on pathogen-mediated balancing selection as a mechanism for maintaining variation in this bottlenecked population.

Advancement for this research would be to screen the pathogen fauna that exist within the different SW populations and how exposure to different suites of pathogens results in different pathogen-selection regimes. Failing to incorporate the complexity of the immune system with the polygenic nature of many pathogen infections limits our ability to test hypotheses about the possible role of selection in shaping patterns of variation in pathogen resistance and/ or susceptibility. In a short period time (< 25 years), there have already been big differences that have emerged with regards to disease resistance. Two out of four translocated populations have eradicated GRW1 (Fairfield *et al. in prep*, for details on translocations see Komdeur 1994; Richardson *et al.* 2006; Wright *et al.* 2014). This is despite the fact that all new populations included a proportion of founders with GRW1 infection (Hutchings 2009). The most recently translocated population to Frégate Island in 2011, which is ten times the size of Cousin, has a much greater diversity of flora and fauna diversity on the island compared to Cousin and other islands holding translocated populations. In addition to considering the pathogen, there is a need in this field to incorporate study on the vectors responsible for pathogen transmission. For example, it has been well-shown that the intermediate dipteran vector host plays a vital role in the co-evolutionary arms race between pathogen and host (for review, see Bordes & Morand 2015). Therefore, I would be keen to assess vector abundance, species diversity and explore individuals at a molecular level to fully understand how the intermediate host fits in with overall pathogen-mediated selection within a community of different hosts and different pathogens.

Formal analyses for detecting evidence of natural selection acting on the parasite population are relatively new. For example, analyses studying the diversity observed in genes encoding antigens, especially those in the merozoite and sporozoite, and attributing that diversity to the action of natural selection imposed by the host immune system (Garamszegi *et al.* 2015; Marzal *et al.* 2015; Pigeault *et al.* 2015). A study has already looked at the genetic diversity of malarial parasite lineages in the great reed warbler *Acrocephalus arundinaceus* (Bensch *et al.* 2007; Westerdahl *et al.* 2012). These studies emphasise how the

knowledge of extrinsic parameters such as vector distribution and alternative hosts are needed to fully understand patterns of infection. Overall, assessing pathogen pressures across SW populations across multiple years and a long time scale, may contribute to our understanding of how pathogen mediated selective pressures fluctuate over time and shape genetic variation in natural populations.

The MHC has long been a paradigm for the study of functional variation and (for review, see Bernatchez & Landry 2003). However it is clear that we need to consider other immune gene groups if we are to fully understand these processes to (Acevedo-Whitehouse & Cunningham 2006). The candidate-gene approach can successfully examine genes based on *a priori* hypotheses and establish functionality of variation by using a bottom-up approach (Fitzpatrick *et al.* 2005). Research is now increasing, particularly in TLRs, and there remain a number of other immune gene groups to be explored; particularly from the innate immune system, which is still relatively understudied (Kaiser 2007).

There are many innate multigene cytokine families, especially the chemokines and their receptors, and the TNF/TNFR super-families. All of these cytokine families are under selective pressure (for review, see Hill 2001). Preliminary evidence shows that non-MHC cytokine gene variants such as Interleukin-1, Interleukin-4, cytotoxic T lymphocyte-associated molecule-4 and natural-resistance-associated macrophage protein 1 are all relevant to disease resistance / susceptibility (for examples, see Walley & Cookson 1996; Donner *et al.* 1997; Bellamy *et al.* 1998; Nicoll *et al.* 2000). Killer-cell Immunoglobulin-like receptors (KIRs) have also been shown to be highly polymorphic (Lindenstrøm *et al.* 2004). Chicken-killer immunoglobulin-like receptors (CHIRs) are especially appealing candidates for their extremely high degree of polymorphism with single nucleotide substitutions generating different CHIRs at a fast evolutionary rate (Nikolaidis *et al.* 2005)**.** Natural killer-cell receptors share many features with the MHC because they are both large dense clusters of loci with high levels of polymorphisms, maintained by resistance to infection (Trowsdale 2001)**.** Conceptually, these are all valid and worthy candidate loci for study into functional variation. There is a persistent need for broader research on traditional vertebrate models which can be transferred to wild populations. Better yet, if there is an opportunity to conduct this research in the wild, such knowledge would enable broader understanding of

the levels at which natural selection can act on immunity and thus better inform conservation biology.

It would be beneficial to investigate the interactions between different immune genes and to study how those interactions impact upon individual fitness in order to better understand the role of adaptive genetic variation in small populations. The publication of an *Acrocephalus* genome would allow access to a wealth of genetic data that would greatly enhance our research from the designing of locus-specific primers to a better resolution. Genomic technologies now offer unprecedented opportunities and with the exponential advancement of their speed and affordability, whole genomes are quickly overtaking the use of conformational techniques previously used to explore the structure and function of genes like the MHC (Thomas & Klaper 2004; Avise 2010; Babik 2010; Warren *et al.* 2010). When constructing phylogenies in Chapters 2 and 3 based on the variation characterised at AvBDs and TLRs respectively, a number of nodes remained unresolved. While this is likely to have been a power-issue (limited evidence of shared polymorphism between species), this remains problematic when wanting to infer the role of selection over a longer period of evolutionary time. Phylogenies of other genes in the genome would greatly help to address this problem and make it clearer whether, for example, observed neutral and functional polymorphism is due to recent species divergence.

The SW is an invaluable model for asking important evolutionary- questions, given that it has been intensively monitored and studied for over 25 years. There is a wealth of accurate fitness and life-history data, environmental monitoring and more than 5000 blood samples collected longitudinally from over 6000 birds (for some examples, see Komdeur 1991; van de Crommenacker *et al.* 2011; Barrett *et al.* 2013; Spurgin *et al.* 2014). The island ecology is relatively benign and the absence of predators means that there is a relatively high annual survival rate of 0.61 and 0.85 for juvenile and adults (Brouwer *et al.* 2006) and accurate fitness data available for each individual within the population. By having a re-sighting probability of 0.95 (Brouwer *et al.* 2006), it presents the rare opportunity of being able to study a natural 'laboratory' population when typically, extensive molecular ecology studies in wild populations prove to be scarce.

I hope that the content of this thesis may prove to be of use in its wider applications to conservation biodiversity and emphasise the need to include and progress research into evolutionary conservation. This thesis' research provides novel information about multiple gene families within a natural population and uses a combination of approaches to try to infer the evolutionary processes responsible for shaping variation at these gene families. Knowing how such variation is shaped has important conservation implications in being able to assess population / species adaptive potential, epidemic risks and to predict responses to future novel challenges. In the case of this thesis, the focus lies with response to challenges of a pathogenic nature, at a time when novel pathogens are increasingly emerging in natural populations. Consequently, these sorts of studies are integral to better understanding disease dynamics and the long-term viability of populations or species of conservation concern.

References

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

Agudo R, Carrete M, Alcaide M *et al.* (2012) Genetic diversity at neutral and adaptive loci determines individual fitness in a long-lived territorial bird. *Proceedings of the Royal Society B*, **279**, 3241–3249.

De Aguilar JR, Schut E, Merino S *et al.* (2013) MHC class II B diversity in blue tits: a preliminary study. *Ecology and Evolution*, **3**, 1878–1889.

Alcaide M (2010) On the relative roles of selection and genetic drift in shaping MHC variation. *Molecular Ecology*, **19**, 3842–4.

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

Anderson RM, May RM (1978) Regulation and Stability of Host-Parasite Population Interactions. *Journal of Animal Ecology*, **47**, 219–247.

Arandjelovic M, Guschanski K, Schubert G *et al.* (2009) Two-step multiplex polymerase chain reaction improves the speed and accuracy of genotyping using DNA from noninvasive and museum samples. *Molecular Ecology Resources*, **9**, 28–36.

Ardia DR, Parmentier HK, Vogel L a. (2011) The role of constraints and limitation in driving individual variation in immune response. *Functional Ecology*, **25**, 61–73.

Areal H, Abrantes J, Esteves PJ (2011) Signatures of positive selection in Toll-like receptor (TLR) genes in mammals. *BMC evolutionary biology*, **11**, 368.

Avise JC (2010) Perspective: Conservation genetics enters the genomics era. *Conservation Genetics*, **11**, 665–669.

Babik W (2010) Methods for MHC genotyping in non-model vertebrates. *Molecular Ecology Resources*, **10**, 237–251.

Bach FH (1976) *Differential function of MHC LD and SD determinants: Leukocyte Membrane Determinants Regulating Immune Response.* (V Eijsvoogel, D Roos, W Zeijkemake, Eds,). Academic Press, New York.

Barreiro LB, Ben-Ali M, Quach H *et al.* (2009) Evolutionary dynamics of human Toll-like receptors and their different contributions to host defense. *PLoS Genetics*, **5**, e1000562.

Barrett ELB, Burke T a, Hammers M, Komdeur J, Richardson DS (2013) Telomere length and dynamics predict mortality in a wild longitudinal study. *Molecular Ecology*, **22**, 249–59.

Bellamy R, Ruwende C, Corrah T *et al.* (1998) Variations in the NRAMP1 gene and susceptibility to Tuberculosis in West Africans. *New England Journal of Medicine*, 640–644.

Bensch S, Waldenström J, Jonzén N *et al.* (2007) Temporal dynamics and diversity of avian malaria parasites in a single host species. *Journal of Animal Ecology*, **76**, 112–122.

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

Bihl F, Salez L, Beaubier M *et al.* (2003) Overexpression of Toll-like receptor 4 amplifies the host response to lipopolysaccharide and provides a survival advantage in transgenic mice. *Journal of Immunology*, **170**, 6141–50.

Bordes F, Morand S (2015) *Parasite Diversity and Diversification: Evolutionary Ecology Meets Phylogenetics* (S Morand, BR Krasnov, DTJ Littlewood, Eds,). Cambridge University Press.

Boyd AC, Peroval MY, Hammond JA *et al.* (2012) *TLR15* is unique to avian and reptilian lineages and recognizes a yeast-derived agonist. *Journal of Immunology*, **189**, 4930–4938.

Boyd A, Philbin VJ, Smith AL (2007) Conserved and distinct aspects of the avian Toll-like receptor (TLR) system: implications for transmission and control of bird-borne zoonoses. *Biochemical Society Transactions*, **35**, 1504–1507.

Brouwer L, Barr I, van de Pol M *et al.* (2010) MHC-dependent survival in a wild population: evidence for hidden genetic benefits gained through extra-pair fertilizations. *Molecular Ecology*, **19**, 3444–3455.

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

Collar NJ, Stuart SN (1985) *Threatened birds of Africa and related islands.* tional Council for Bird Preservation, Cambridge.

Crnokrak P, Barrett SCH (2002) Perspective: purging the genetic load: a review of the experimental evidence. *Evolution; international journal of organic evolution*, **56**, 2347–2358.

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

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

Darwin C (1859) *On the origin of species by means of natural selection.* John Murray.

Derache C, Meudal H, Aucagne V *et al.* (2012) Initial insights into structure-activity relationships of avian defensins. *Journal of Biological Chemistry*, **287**, 7746–7755.

Van Dijk A, Veldhuizen EJ, Haagsman HP (2008) Avian defensins. *Vetinary Immunology & Immunopathology*, **124**, 1–18.

Donner H, Rau H, Walfish PG *et al.* (1997) CTLA4 alanine-17 confers genetic susceptibility to Graves' disease and to type 1 diabetes mellitus. *Journal of Clinical Endocrinology and Metabolism*, **82**, 143–146.

Ejsmond MJ, Radwan J (2009) MHC diversity in bottlenecked populations: a simulation model. *Conservation Genetics*, **12**, 129–137.

Fitzpatrick MJ, Ben-Shahar Y, Smid HM *et al.* (2005) Candidate genes for behavioural ecology. *Trends in Ecology & Evolution*, **20**, 96–104.

Fornůsková A, Vinkler M, Pagès M *et al.* (2013) Contrasted evolutionary histories of two Toll-like receptors (*TLR4* and Tlr7) in wild rodents (MURINAE). *BMC evolutionary biology*, **13**, 1–17.

Frankham R (2010) Challenges and opportunities of genetic approaches to biological conservation. *Biological Conservation*, **143**, 1919–1927.

Garamszegi LZ, Zagalska-Neubauer M, Canal D *et al.* (2015) Malaria parasites, immune challenge, MHC variability, and predator avoidance in a passerine bird. *Behavioral Ecology*, **00**, 1–11.

Garcia de Leaniz C, Fleming I a, Einum S *et al.* (2007) A critical review of adaptive genetic variation in Atlantic salmon: implications for conservation. *Biological reviews of the Cambridge Philosophical Society*, **82**, 173–211.

Garrigan D, Hedrick PW (2003) Perspective: detecting adaptive molecular polymorphism: lessons from the MHC. *Evolution*, **57**, 1707–22.

Grueber CE, Wallis GP, Jamieson IG (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**, 4470–4482.

Grueber CE, Wallis GP, Jamieson IG (2014) Episodic positive selection in the evolution of avian toll-like receptor innate immunity genes. *PloS one*, **9**, e89632.

Hansson B, Richardson DS (2005) Genetic variation in two endangered *Acrocephalus* species compared to a widespread congener: estimates based on functional and random loci. *Animal Conservation*, **8**, 83–90.

Hellgren O (2014) Allelic variation at innate immune genes (avian β-defensins), with in a natural population of great tits. *Journal of Avian Biology*, **46**, 113–118.

Hellgren O, Ekblom R (2010) Evolution of a cluster of innate immune genes (beta-defensins) along the ancestral lines of chicken and zebra finch. *Immunome Research*, **6**, 3.

Higgs R, Cormican P, Cahalane S *et al.* (2006) Induction of a Novel Chicken Toll-Like Receptor following Salmonella enterica Serovar Typhimurium Infection Induction of a Novel Chicken Toll-Like Receptor following Salmonella enterica Serovar Typhimurium Infection. *Infection and Immunity*, **74**, 1692–1698.

Hill AVS (2001) Immunogenetics and genomics. *The Lancet*, **357**, 2037–2041.

Hollox EJ, Armour JA (2008) Directional and balancing selection in human beta-defensins. *BMC Evolutionary Biology*, **8**, 113.

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

Hughes AL (2007) Looking for Darwin in all the wrong places: the misguided quest for positive selection at the nucleotide sequence level. *Heredity*, **99**, 364–373.

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

Jamieson IG (2011) Founder effects, inbreeding, and loss of genetic diversity in four avian reintroduction programs. *Conservation Biology*, **25**, 115–123.

Kaiser P (2007) The avian immune genome--a glass half-full or half-empty? *Cytogenetic Genome Research*, **117**, 221–230.

Kawai T, Akira S (2006) TLR signaling. *Cell death and differentiation*, **13**, 816–825.

Keestra M, de Zoete MR, Bouwman LI, van Putten JPM (2010) Chicken *TLR21* is an innate CpG DNA receptor distinct from mammalian TLR9. *Journal of Immunology*, **185**, 460–7.

Kimura M (1986) DNA and the Neutral Theory. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **312**, 343–354.

Klein J (1986) *Natural History of the Major Histocompatibility Complex.* John Wiley & Sons, New York.

Komdeur J (1991) *Cooperative breeding in the Seychelles warbler*. University of Cambridge.

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

Komdeur J, Kappe A, Zande L Van De (1998) Influence of population isolation on genetic variation and demography in Seychelles warblers: a field experiment. *Animal Conservation*, **1**, 203–212.

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

Lindenstrøm T, Secombes CJ, Buchmann K (2004) Expression of immune response genes in rainbow trout skin induced by Gyrodactylus derjavini infections. *Vetinary Immunology & Immunopathology*, **97**, 137–148.

Manning CJ, Wakeland EK, Potts WK (1992) Communal nesting patterns in mice implicate MHC genes in kin recognition. *Nature*, **356**, 133–135.

Marzal A, García-Longoria L, Cárdenas Callirgos JM, Sehgal RN (2015) Invasive avian malaria as an emerging parasitic disease in native birds of Peru. *Biological Invasions*, **17**, 39–45.

Masatoshi N, Takeo M, Chakraborty R (1975) The Bottleneck Effect and Genetic Variability in Populations. *Evolution*, **29**, 1–10.

May RM, Anderson RM (1983) Epidemiology and genetics in the coevolution of parasites and hosts. *Proceedings of the Royal Society B: Biological Sciences*, **219**, 281–313.

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

Nakajima T, Ohtani H, Satta Y *et al.* (2008) Natural selection in the TLR-related genes in the course of primate evolution. *Immunogenetics*, **60**, 727–35.

Nerren JR, He H, Genovese K, Kogut MH (2010) Expression of the avian-specific toll-like receptor 15 in chicken heterophils is mediated by Gram-negative and Gram-positive bacteria, but not TLR agonists. *Veterinary Immunology and Immunopathology*, **136**, 151–156.

Nerren JR, Swaggerty CL, MacKinnon KM *et al.* (2009) Differential mRNA expression of the avian-specific toll-like receptor 15 between heterophils from Salmonella-susceptible and resistant chickens. *Immunogenetics*, **61**, 71–7.

Nicoll JAR, Mrak RE, Graham DI *et al.* (2000) Association of Interleukin-1 gene polymorphisms with Alzheimer's disease. *Annual Review of Neurology*, **47**, 365–368.

Nikolaidis N, Klein J, Masatoshi N (2005) Origin and evolution of the Ig-like domains present in mammalian leukocyte receptors: insights from chicken, frog, and fish homologues. *Immunogenetics*, **57**, 151–157.

Nizet V (2006) Antimicrobial peptide resistance mechanisms of human bacterial pathogens. *Current Issues in Molecular Biology*, **8**, 11–26.

O'Brien SJ, Evermann JF (1988) Interactive influence of infectious disease and genetic diversity in natural populations. *Trends in Ecology & Evolution*, **3**, 254–9.

Ohta T (1995) Gene Conversion vs Point Mutation in Generating Variability at the Antigen Recognition Site of Major Histocompatibility Complex Loci. *Journal of Molecular Evolution*, **41**, 115–119.

Oustalet E (1878) Etude sur la faune ornithologique des iles Seychelles. *Bulletin de La Société Philomáthique*, **2**, 161–206.

Pertoldi C, Bijlsma R, Loeschcke V (2007) Conservation genetics in a globally changing environment: Present problems, paradoxes and future challenges. *Biodiversity and Conservation*, **16**, 4147–4163.

Pigeault R, Vézilier J, Cornet S *et al.* (2015) Avian malaria : a new lease of life for an old experimental model to study the evolutionary ecology of Plasmodium. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **370**, 20140300.

Rehli M (2002) Of mice and men: species variations of Toll-like receptor expression. *Trends in Immunology*, **23**, 375–8.

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

Richardson DS, Bristol R, Shah NJ (2006) Translocation of the Seychelles warbler *Acrocephalus sechellensis* to establish a new population on Denis Island, Seychelles. *Conservation Evidence*, **3**, 54–57.

Richardson DL, Jury FL, Dawson DA *et al.* (2000) Fifty Seychelles warbler (*Acrocephalus sechellensis*) microsatellite loci polymorphic in Sylviidae species and their cross-species amplification in other passerine birds. *Molecular Ecology*, **9**, 2155–7.

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

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

Roach JC, Glusman G, Rowen L *et al.* (2005) The evolution of vertebrate Toll-like receptors. *Proceedings of the National Academy of Sciences USA*, **102**, 9577–9582.

Scherman K, Råberg L, Westerdahl H (2014) Positive selection on MHC class II DRB and DQB genes in the bank vole (Myodes glareolus). *Journal of Molecular Evolution*, **78**, 293–305.

Shen T, Xu S, Wang X *et al.* (2012) Adaptive evolution and functional constraint at *TLR4* during the secondary aquatic adaptation and diversification of cetaceans. *BMC evolutionary biology*, **12**, 39.

Spurgin LG, Richardson DS (2010) How pathogens drive genetic diversity: MHC, mechanisms and misunderstandings. *Proceedings of the Royal Society B*, **277**, 979–988.

Spurgin LG, Wright DJ, van der Velde M *et al.* (2014) Museum DNA reveals the demographic history of the endangered Seychelles warbler. *Evolutionary Applications*, 1–10.

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

Sutton JT, Robertson BC, Grueber CE, Stanton JAL, Jamieson IG (2013) Characterization of MHC class II B polymorphism in bottlenecked New Zealand saddlebacks reveals low levels of genetic diversity. *Immunogenetics*, **65**, 619–633.

Takahata N, Satta Y, Klein J (1992) Polymorphism and balancing selection at major histocompatibility complex oci. *Genetics*, **130TSK92**, 925–938.

Takeuchi O, Sato S, Horiuchi T *et al.* (2002) Cutting Edge: Role of Toll-Like Receptor 1 in Mediating Immune Response to Microbial Lipoproteins. *The Journal of Immunology*, **169**, 10–14.

Thomas MA, Klaper R (2004) Genomics for the ecological toolbox. *Trends in Ecology & Evolution*, **19**, 439–445.

Trowsdale J (2001) Genetic and Functional Relationships Review between MHC and NK Receptor Genes. *Immunity*, **15**, 363–374.

Vasemagi A, Primmer CR (2005) Challenges for identifying functionally important genetic variation: the promise of combining complementary research strategies. *Molecular Ecology*, **14**, 3623–3642.

Vogel SN, Fitzgerald K a, Fenton MJ (2003) TLRs: differential adapter utilization by toll-like receptors mediates TLR-specific patterns of gene expression. *Molecular interventions*, **3**, 466–477.

Walley AJ, Cookson WO (1996) Investigation of an interleukin-4 promoter polymorphism for associations with asthma and atopy. *Journal of Medical Genetics*, **33**, 689–692.

Warren WC, Clayton DF, Ellegren H *et al.* (2010) The genome of a songbird. *Nature*, **464**, 757–762.

Westerdahl H, Asghar M, Hasselquist D, Bensch S (2012) Quantitative disease resistance: to better understand parasite-mediated selection on major histocompatibility complex. *Proceedings of the Royal Society B*, **279**, 577–584.

Willi Y, Van Buskirk J, Schmid B, Fischer M (2007) Genetic isolation of fragmented populations is exacerbated by drift and selection. *Journal of Evolutionary Biology*, **20**, 534–42.

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

Wright DJ, Shah NJ, Richardson DS (2014) Translocation of the Seychelles warbler *Acrocephalus sechellensis* to establish a new population on Frégate Island , Seychelles. *Conservation Evidence*, **11**, 20–24.

Zhai W, Nielsen R, Slatkin M (2009) An investigation of the statistical power of neutrality tests based on comparative and population genetic data. *Molecular Biology and Evolution*, **26**, 273–83.