Reduction in population size and not a shift in parasite community affect evolution of immune genes in island birds

Mathilde BARTHE[1*], Claire DOUTRELANT[2], Rita COVAS[3-5], Martim MELO[3-5], Juan Carlos ILLERA[6], Marie-Ka TILAK[1], Constance COLOMBIER[1], Thibault LEROY[1,7], Claire LOISEAU[2,3a], Benoit NABHOLZ[1,8a]

[1] ISEM, Univ Montpellier, CNRS, IRD, Montpellier, France

[2] Centre d'Ecologie Fonctionnelle et Evolutive, CNRS, Univ Montpellier, EPHE, IRD, Montpellier, France

[3] CIBIO, Research Centre in Biodiversity and Genetic Resources, InBIO Associated Laboratory, Campus Agrário de Vairão, Vairão, Portugal

[4] MHNC-UP, Natural History and Science Museum of the University of Porto, Porto, Portugal

[5] DST/NRF Centre of Excellence, FitzPatrick Institute, University of Cape Town, Rondebosch, South Africa

[6] Biodiversity Research Institute (CSIC-Oviedo University-Principality of Asturias), Oviedo University, Mieres, Spain

[7]IRHS-UMR1345, Université d'Angers, INRAE, Institut Agro, SFR 4207 QuaSaV, 49071, Beaucouzé, France

[8]Institut universitaire de France, Paris

[*] Corresponding author : mathilde.barthe.pro@gmail.com

[a] Last co-authors

Abstract

Shared ecological conditions encountered by species that colonize islands often lead to the evolution of convergent phenotypes, commonly referred to as “island syndrome”. Reduced immune functions have been previously proposed to be part of the island syndrome, as a consequence of the reduced diversity of pathogens on island ecosystems. According to this hypothesis, immune genes are expected to exhibit genomic signatures of relaxed selection pressure in island species. In this study, we used comparative genomic methods to study immune genes in island species (N = 20) and their mainland relatives (N = 14). We gathered public data as well as generated new data on innate (Toll-Like Receptors, Beta Defensins) and acquired immune genes (Major Histocompatibility Complex classes I and II), but also on hundreds of genes annotated as involved in various immune functions. As a control, we used a set of 97 genes not involved in immune functions, to account for the lower effective population sizes in island species. We used synonymous and non-synonymous variations to estimate the selection pressure acting on immune genes. For the genes evolving under balancing selection, we used simulation to estimate the impact of population size variation. We found a significant effect of drift on immune genes of island species leading to a reduction in genetic diversity and efficacy of selection. However, the intensity of relaxed selection was not significantly different from control genes, except for MHC class II genes. These genes exhibit a significantly higher level of non-synonymous loss of polymorphism than expected assuming only drift and an evolution under frequency dependent selection, possibly due to a reduction of extracellular parasite communities on islands. Overall, our results showed that demographic effects lead to a decrease in the immune functions of island species, but the relaxed selection caused by a reduced parasite pressure may only occur in some immune genes categories.

Keywords: genetic drift, island evolution, immunity, Toll-Like Receptors, Beta-Defensins, major histocompatibility complex, molecular evolution, population genomics

Introduction

Island colonizers face new communities of competitors, predators and parasites in a small area with limited resources, which generally result in high extinction rates of colonizers (Losos and Ricklefs 2009). Oceanic island faunas are characterized by a low species richness, coupled with high population densities for each species (MacArthur and Wilson 1967; Warren et al. 2015) - which translates in communities with, on average, low levels of inter-specific interaction and high levels of intra-specific competition (but see Rando et al. 2010 for an example of character displacement due to competition among island finch species). These shared island characteristics are thought to underlie the evolution of convergent phenotypes, in what is called the 'island syndrome' (Baeckens and Van Damme 2020). Convergence has been documented in multiple traits, such as size modification (dwarfism or gigantism; Lomolino 2005), reduction of dispersal (Bertrand et al. 2014; Waters et al. 2020), shift towards K life history strategies (MacArthur and Wilson 1967; Boyce 1984; Covas 2012), evolution of generalist traits (Blondel 2000; Warren et al. 2015), or changes in colour and acoustic signals (Grant 1965; Doutrelant et al. 2016).

Reduced immune function has also been hypothesized as an island syndrome trait, directly linked to reduced parasite pressure on islands (Wikelski et al. 2004; Matson and Beadell 2010; Lobato et al. 2017). Island parasite communities are i) less diverse (Beadell et al. 2006; Maria et al. 2009; Pérez-Rodríguez et al. 2013; Illera et al. 2015; Loiseau et al. 2017), and ii) could be less virulent due to the expansion of the ecological niche expected by the theory of island biogeography. In fact, island parasites are generally more generalists than their mainland counterparts, which could lead to a reduced virulence due to the trade-off between replication capacity and resistance against host immune defenses (Hochberg and Møller 2001; Garamszegi 2006; Pérez-Rodríguez et al. 2013). Overall, a reduction of parasitic pressure should lead to a weakening of the immune system due to the costs of maintaining efficient immune functions (Lindström et al. 2004; Wikelski et al. 2004; Matson and Beadell 2010). Such reduction may have important implications for the ability of these populations to resist or tolerate novel pathogens. The introduction of avian malaria in the Hawaiian archipelago, and the subsequent extinctions and population declines of many endemic species is the most emblematic example (Van Riper III et al. 1986; Wikelski et al. 2004).

Immunological parameters, such as blood leukocyte concentration, antibodies or other immune proteins (e.g. haptoglobin), hemolysis, and hemagglutination (Lee 2006; Matson and Beadell 2010) may serve as proxies to determine population immune functions. To date, the

majority of studies that focused on island avifauna have found ambiguous results, with either no support for a reduced immune response on island species (Matson 2006; Beadell et al. 2007), or contrasted results, such as a lower humoral component (total immunoglobulins) on islands, but a similar innate component (haptoglobin levels) between island and mainland species (Lobato et al. 2017). The use of immune parameters as proxies of immune function is fraught with difficulties (Lobato et al. 2017). The study of molecular evolution of immune genes therefore represents an alternative strategy to tackle this question. However, it is necessary to distinguish neutral effects, the demographic effects resulting from island colonization, from selective ones, the potential relaxation of selection pressures due to the changes in the pathogen community.

First, the bottleneck experienced by species during island colonization leads to a decrease in genetic variability (Frankham 1997). A reduced genetic diversity at loci involved in immunity should have a direct implication on immune functions (Hawley et al., 2005; Hale & Briskie, 2007 but see Spurgin et al. 2011). Second, small population sizes increase genetic drift, which may counteract the effect of natural selection of low-effect mutations, especially weakly deleterious mutations (Ohta 1992). Several recent studies found a greater load of deleterious mutations in island species (Loire et al. 2013; Robinson et al. 2016; Rogers and Slatkin 2017; Kutschera et al. 2020; Leroy et al. 2021). Finally, it is necessary to differentiate genes involved in the innate versus the acquired immune response. The innate immune response is the first line of defense and is composed of phagocytes, macrophages and dendritic cells. These cells allow non-specific recognition of pathogens (Alberts et al. 2002; Akira 2003). For example, Toll-Like Receptors (TLR; transmembrane proteins) trigger a chain reaction leading to the production of various substances, including antimicrobial peptides such as beta-defensins (BD) that have active properties in pathogen cell lysis (Velová et al. 2018). On the other hand, the acquired immune system allows a specific response, characterized by immune memory. Major Histocompatibility Complex (MHC) genes code for surface glycoproteins that bind to antigenic peptides, and present them to the cells of the immune system; class I and II genes ensure the presentation of a broad spectrum of intra- and extracellular-derived peptides, respectively (Klein 1986). Although all these genes are directly involved in the identification and neutralization of pathogens, previous studies found that they evolve under different selection regimes: TLRs and BDs are under purifying selection which usually results in the selective removal of deleterious alleles and stabilizing selection (van Dijk et al. 2008; Grueber et al. 2014), whereas MHC genes are under balancing selection (Bernatchez and Landry 2003).

Recent studies on birds (Gonzalez-Quevedo, et al. 2015a; Gonzalez-Quevedo, et al. 2015b), amphibians (Belasen et al. 2019), and lizards (Santonastaso et al. 2017) found that the demographic history of island populations led to the loss of genetic variation at immune genes involved in pathogen recognition, such as TLRs and MHC. For example, Santonastaso et al., (2017) demonstrated that the polymorphism pattern in MHC genes and microsatellites covary positively with island area in Podarcis lizards, suggesting a dominant role for genetic drift in driving the evolution of the MHC. Gonzalez-Quevedo, et al. (2015a) found a similar pattern comparing TLR and microsatellite polymorphism in the Berthelot pipit, Anthus berthelotii, an endemic species from the Macaronesian archipelago, supporting a predominant role of genetic drift in TLR evolution. However, these studies did not explicitly test the hypothesis of a relaxed selection pressure on islands imposed by an impoverished parasite community. All other things being equal, it is expected that the polymorphism pattern of a coding sequence decreases with population size (Buffalo 2021; Leroy et al. 2021). Therefore, a decrease in polymorphism with population size could not be taken as a proof of a relaxation in the selection pressure.

Here, we study a dataset of 34 bird species (20 insular and 14 mainland species; Figure 1) combining the 24 species of Leroy et al. (2021) and 10 newly generated by targeted-capture sequencing (Table 1). To be able to demonstrate a change in natural selection, a traditional approach is to contrast polymorphism of synonymous sites (Ps) with polymorphism of non-synonymous sites (Pn). Synonymous mutations refer to mutations that do not alter amino acid sequences, whereas non-synonymous mutations do. Assuming no selection on Ps and a population at equilibrium, $Ps = 4*Ne*\mu$ where Ne is the effective population size and μ the mutation rate, and $Pn = 4*Ne*\mu*f$, where f is a function that depends on the distribution of the fitness effect (DFE) of mutations and integrates the probability of an allele to segregate at a given frequency, ranging from 0 to 1 (Sawyer and Hartl 1992). The DFE describes the density probability of the selective advantage of an allele (s); the fixation probability of a non-synonymous allele is therefore dependent on the product Ne*s (Kimura 1962). In this context, the ratio Pn/Ps is typically interpreted as the result of a change in natural selection.

If the average selection coefficient (s) changes due to shifts in the parasitic community, the ratio Pn/Ps could have been expected to increase on the island. However, the fixation probability depends on the product Ne*s, and a variation in Ne is also expected to impact the efficacy of selection and thus the ratio Pn/Ps across the entire transcriptome, particularly in the presence of slightly deleterious mutations (Ohta 1992; Charlesworth and Eyre-Walker 2008; Loire et al. 2013; Leroy et al. 2021). Therefore, we predict a significant effect of drift on island species leading to a genome-wide reduction in genetic diversity and efficacy of

selection, as reported by previous studies. In addition, due to their lower population sizes, island birds compared to continental species exhibit a genome-wide reduction in genetic diversity and efficacy of selection (Kutschera et al. 2020; Leroy et al. 2021). Therefore, we expect a similar reduction in immune genes diversity even without any change in the parasite pressure.

To disentangle the effect of population size from a change in parasite pressure and estimate the impact of demography on the efficacy of selection, we randomly selected protein-coding genes (i.e., control genes) implied in various biological functions (Fijarczyk et al. 2016; Leroy et al. 2021). If a reduced parasite pressure on islands directly impacts the evolution of immunity genes, the genetic diversity of immunity genes is expected to show a larger variation between island and continental species than the control genes. More specifically, for genes under purifying selection, non-synonymous weekly deleterious mutations, normally eliminated under strong selection, would be maintained, leading to an increase of genetic diversity. By contrast, for genes under balancing selection, non-synonymous advantageous mutations, normally maintained in the polymorphism under strong selection, would be eliminated leading to a decrease of genetic diversity (Figure 2).

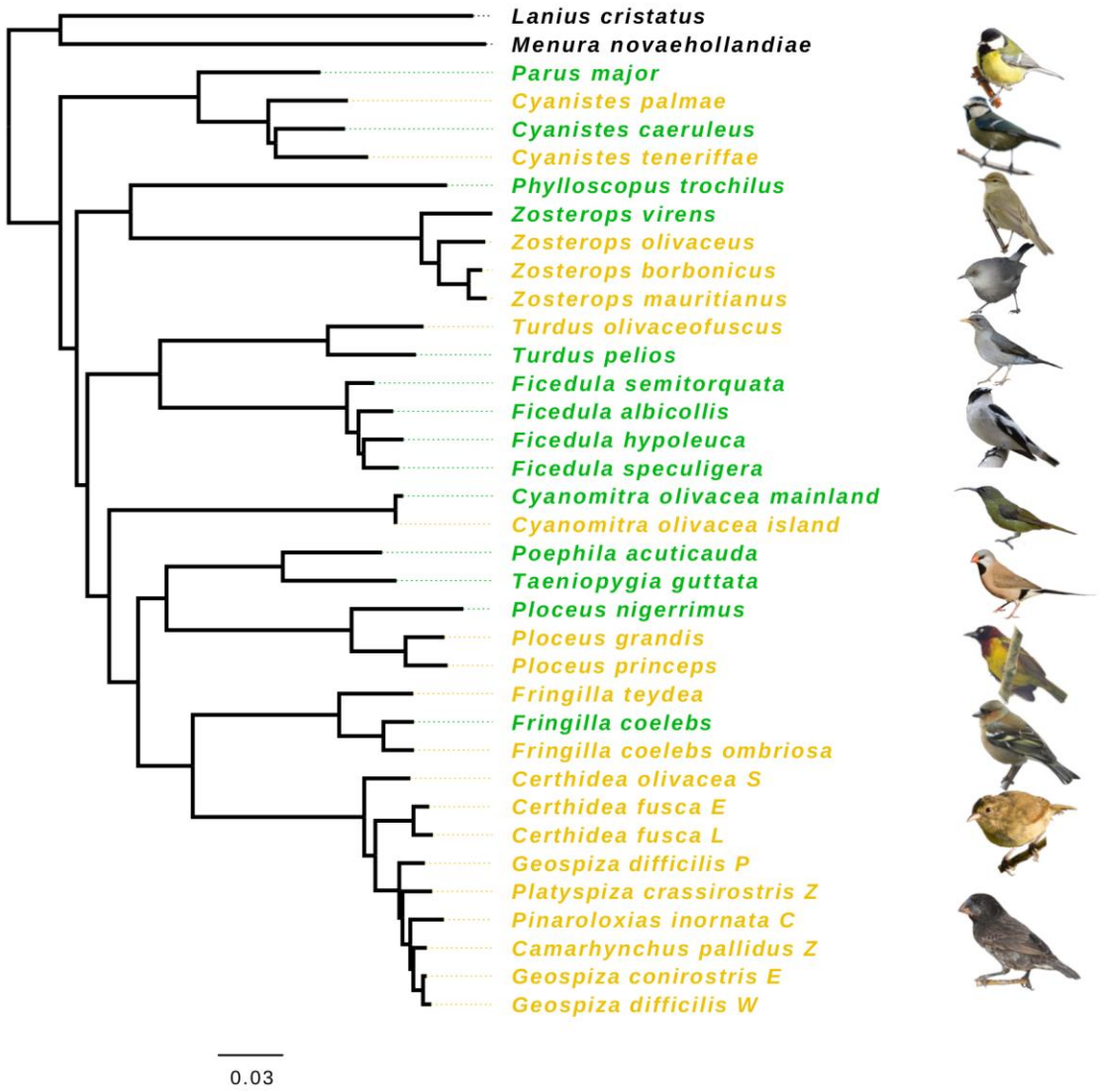


Figure 1: Phylogeny based on mitochondrial genes of species from the dataset reconstructed by maximum likelihood method (IQTREE model GTR+Gamma). Species names in yellow indicate island species, and in green, mainland species. Ultrafast bootstrap values are provided in the supplementary methods. Some relationships are poorly supported. Bird representations are not to scale. Photos from top to bottom : *P. major, C. caeruleus, P. trochilus, Z. borbonicus, T. pelios, F. albicollis, C. olivacea, P. acuticauda, P. grandis, F. coelebs, C. fusca, G. conirostris*. Photo credits: A. Chudý, F. Desmoulins, E. Giacone, G. Lasley, Lianaj, Y. Lyubchenko, B. Nabholz, J.D. Reynolds, K. Samodurov, A. Sarkisyan, Wimvz, Birdpics, T. Aronson, G. Lasley, P. Vos (iNaturalist.org); M. Gabrielli (*Zosterops borbonicus*).

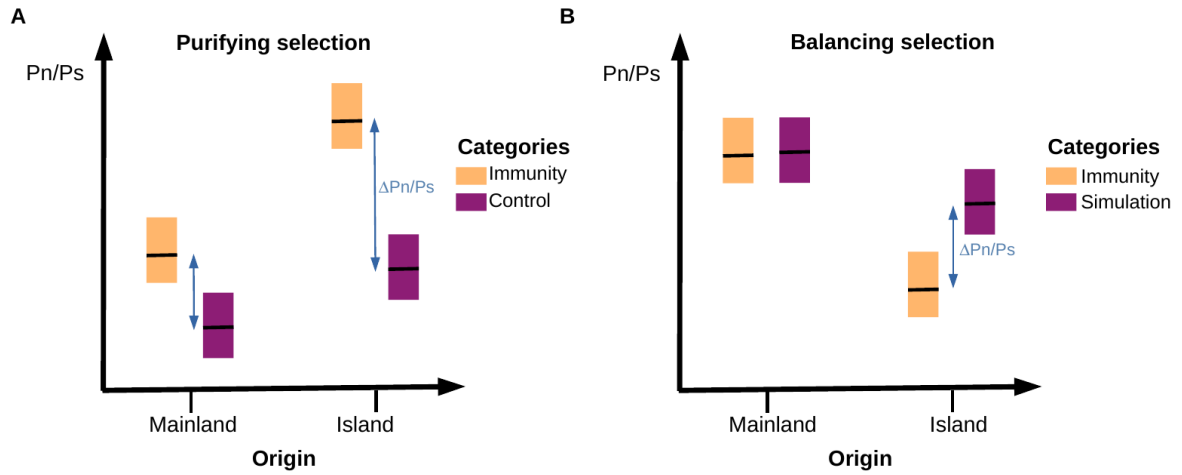


Figure 2: Conceptual diagram showing the expected results under the hypothesis of a relaxation in the selection pressure of the immunity genes in island species due to a change in the parasitic community. A) Genes evolving under purifying selection where control genes are randomly selected protein-coding genes. B) Genes evolving under balancing selection where controls are obtained from SLiM simulations of genes evolving under the same balancing selection but different population size. Under the hypothesis of a relaxed selection as a consequence of the reduced diversity of pathogens on island ecosystems, the difference in Pn/Ps between categories (ΔPn/Ps) is expected to be different between species' origin, leading to a statistical interaction between gene categories and origin.

Methods

Dataset

Alignments of Coding DNA Sequences (CDS) of individuals from 24 species were obtained from Leroy et al., (2021). In addition, data for ten other species (six and four from islands and mainland, respectively) were newly generated for this study by targeted-capture sequencing. Blood samples and subsequent DNA extractions were performed by different research teams. The complete dataset consisted of 34 bird species (20 and 14 insular and mainland species respectively; Table 1; Figure 1). We filtered alignments in order to retain only files containing a minimum of five diploid individuals per site (Table 1).

Sequence enrichment was performed using MYBaits Custom Target Capture Kit targeting 21 immune genes: 10 Toll-Like receptors (TLR), 9 Beta Defensins (BD), 2 Major Histocompatibility Complex (MHC) and 97 control genes. We followed the manufacturer's protocol (Rohland and Reich 2012). Illumina high-throughput sequencing using a paired-end 150 bp strategy was performed by Novogene (Cambridge, UK).

196 Table 1: List of species and sampling localities, along with the type of data obtained and the
197 number of individuals (N).

Species	Origin	Island/Country	N	Reference genome	Reference	Type of data
Cyanistes teneriffae palmae	Island	La Palma	15			
Cyanistes teneriffae teneriffae	Island	Tenerife	15	*Cyanistes caeruleus*	This study	Capture
Cyanistes caeruleus	Mainland	France	15			
Parus major	Mainland	Europe	10	*Parus major*	(Corcoran et al. 2017)	Whole genome
Phylloscopus trochilus	Mainland	Europe	9	*Phylloscopus trochilus*	(Lundberg et al. 2017)	Whole genome
Zosterops virens	Mainland	South Africa	7			
Zosterops olivaceus	Island	Réunion	15	*Zosterops borbonicus*	(Leroy et al. 2021)	Whole genome
Zosterops mauritianus	Island	Mauritius	9			
Zosterops borbonicus	Island	Réunion	25			
Ficedula semitorquata	Mainland	Europe	20	*Ficedula albicolis*	(Ellegren et al. 2012)	Whole genome
Ficedula albicollis	Mainland	Europe	20			
Ficedula speculigera	Mainland	Nord Africa	20			
Ficedula hypoleuca	Mainland	Europe	20			
Turdus olivaceofuscus	Island	São Tomé	15	*Turdus pelios*	This study	Capture
Turdus pelios	Mainland	Gabon	15			
Cyanomitra olivacea	Island	Príncipe	15	*Cyanomitra olivacea*	This study	Capture
Cyanomitra olivacea	Mainland	Gabon	15			
Ploceus grandis	Island	São Tomé	14	*Ploceus cuculatus*	This study	Capture
Ploceus princeps	Island	Príncipe	15			
Ploceus nigerrimus	Mainland	Cameroon Gabon	15			
Poephila acuticauda acuticauda	Mainland	Australia	10	*Taeniopygia guttata*	(Singhal et al. 2015)	Whole genome
Taeniopygia guttata castanotis	Mainland	Australia	19			
Fringilla teydea	Island	Tenerife	10	*Fringilla coelebs*	(Leroy et al. 2021)	Whole genome
Fringilla coelebs palmae	Island	La Palma	15			
Fringilla coelebs	Mainland	Spain	9			
Certhidea olivacea	Island	Santiago (Galápagos)	5			
Certhidea fusca	Island	San Cristobal (Galápagos)	10			
Certhidea fusca	Island	Espanola (Galápagos)	10			
Geospiza difficilis	Island	Pinta(Galápagos)	10	*Geopsiza fortis*	(Lamichhaney et al. 2015)	Whole genome
Platyspiza crassirostris	Island	Santa Cruz (Galápagos)	5			
Pinaroloxias inornata	Island	Coco (Galápagos)	8			
Camarhynchus pallidus	Island	Santa Cruz (Galápagos)	5			
Geospiza difficilis	Island	Wolf (Galápagos)	8			
Geospiza conirostris	Island	Espanola (Galápagos)	10			

198

Newly generated draft genome sequence

We generated whole genome sequences at moderate coverage (~40X) for Turdus pelios, Ploceus cucullatus and Cyanomitra olivacea (from Gabon). Library preparation from blood DNA samples and Illumina high-throughput sequencing using a paired-end 150 bp strategy were performed at Novogene (Cambridge, UK). Raw reads were cleaned using FastP (vers. 0.20.0; Chen et al. 2018). Genomes assemblies were performed using SOAPdenovo (vers. 2.04) and Gapcloser (v1.10) (Luo et al. 2012) with parameters "-d 1 -D 2" and a kmers size of 33. Protein annotation was performed by homology detection using genBlastG (She et al. 2011; http://genome.sfu.ca/genblast/download.html) and the transcriptome of the collared flycatcher (Ficedula albicollis; assembly FicAlb1.5; Ellegren et al. 2012) as reference.

Capture data processing

Reads from targeted-capture sequencing were cleaned with FastP (vers. 0.20.0; Chen et al. 2018). Reads of each individual were mapped respectively to the nearest available reference genomes using bwa mem (vers. 0.7.17; (Li 2013); Table 1), with default parameters. Samtools (vers. 1.3.1; Li et al. 2009) and Picard (vers. 1.4.2; Picard Toolkit 2019) were used to convert the mapping files, order and index reads according to their position on the chromosomes (or scaffolds) of the reference genomes or on the draft genomes generated in this study for Ploceus, Cyanomitra and Turdus. Duplicate reads were marked using MarkDuplicates (vers. 1.140; Picard Toolkit 2019). SNP calling was performed with Freebayes (vers. 1.3.1; Garrison and Marth 2012). Freebayes output file (VCF file) was converted to a fasta file by filtering out sites with a minimum quality of 40 and a sequencing depth between 10 and 1000X (sites outside these thresholds were treated as missing data, i.e., 'N'). CDS were then extracted from the alignments using the coordinates of the annotations (gff files). CDS were aligned using MACSE (vers. 2.03; Ranwez et al. 2011) to prevent frameshift mutation errors and GNU-parallel (Tange 2018) was used to parallelise the computation.

Selection and identification of immune and control genes

We defined several groups of immune genes to compare with the control genes. The control group consisted of 97 protein-coding genes randomly selected in the genome of Zosterops borbonicus (Leroy et al. 2019). These control genes allowed the estimation of the average selection pressure that a gene, not involved in the immune response, undergoes in the genome. These genes are single copy (absence of paralogue) and have a variable GC content representative of the whole transcriptome.

For the immune genes, we selected three sets of genes from i) a limited set of genes (Core Group) where functions are unambiguously related to immunity, and ii) two larger sets of genes (Database-group & Sma3s-group), obtained through an automatic annotation pipeline.

The Core Group included MHC class I and class II genes, 10 Toll-Like Receptors (TLRs; Velová et al. 2018) and 9 Beta Defensins (BD; Chapman et al. 2016). The Database group included genes identified by Immunome Knowledge Base (Ortutay and Vihinen 2009, http://structure.bmc.lu.se/idbase/IKB/; last access 04/02/2020) and InnateDB (Breuer et al. 2013, http://www.innatedb.com ; last access 04/02/2020). We also added a set of genes for which the genetic ontology indicated a role in immune functions. To do so, we used the chicken (Gallus gallus) annotation (assembly GRCg6a downloaded from Ensembl database in March 2020; https://www.ensembl.org/). We identified genes with the terms "immun*" or "pathogen*" in their Gene Ontology identifiers description (directory obtained from http://geneontology.org/). This set included 2605 genes considered to be involved in immunity, although some may be only indirectly involved in immunity or have a small impact on immune functions. Finally, the third set of genes (Sma3s-group) has been built up through the Sma3s-group program (vers. 2; Munoz-Mérida et al. 2014). This program annotated sequences in order to be associated with biological functions through gene ontology identifiers. The annotation of the genome of F. albicollis allowed us to identify 3136 genes associated with the genetic ontology "immune system processes". Like for the Database group, this set may include genes with various functions in the immune response. It should be noted that Sma3s-group and Database-group are not mutually exclusive, and some genes are present in both groups. An analysis was performed to identify and exclude genes under balancing selection from Database-group and Sma3s-group sets, due to the potentially antagonistic responses of these genes (see Detection of genes under balancing selection in Supplementary Methods).

Data Analysis

Use of control genes and simulation

We assumed that control genes mostly evolve under purifying selection since balancing selection is generally rare (Andrés et al. 2009; Fijarczyk and Babik 2015). Therefore, the control gene set should provide a good estimation of the impact of genetic drift due to island colonization and the effects of smaller population size of species on islands compared to the mainland species (Leroy et al. 2021). However, MHC genes are known to evolve under balancing selection (Hughes and Nei 1988; Takahata and Nei 1990; Apanius et al. 1997). In the absence of a control gene set evolving under balancing selection, we used simulations to

estimate the impact of demographic changes on polymorphism patterns under this selection regime. SLiM (vers. 3.3.2; Haller and Messer 2017) was used to simulate two populations of 270,000 and 110,000 individuals representing mainland and island effective population size respectively. These sizes correspond to the mean population sizes estimated from our polymorphism data set (see Supplementary Methods). All the details of the simulation parameters, calculations of non-synonymous polymorphism rate (PN) and synonymous polymorphism rate (PS) of simulated sequences, as well as SLiM command lines are provided in Supplementary Methods (Simulation of control genes under balancing selection).

Polymorphism analyses

Synonymous (Ps) and non-synonymous (Pn) nucleotide diversity were estimated from seq_stat_coding written from the Bio++ library (Available as Supplementary data; Guéguen et al. 2013). The mean Pn/Ps was computed as the sum of Pn over the sum of Ps (Wolf et al. 2009). Ps of concatenated sequences of control genes were estimated for each species of our dataset. For the whole-genome sequence species, we compared the Pn/Ps and Ps estimated obtained using the 97 control genes with the values from Leroy et al., (2021; ~5000 genes used in their study). Pn/Ps and Ps correlations showed a R^2 of 0.6 and 0.95 respectively (Figure S2). Thus, the 97 control genes used in our study were representative of a larger set. This allowed us to identify Phylloscopus trochilus as an outlier. Unlike for all other species (e.g. Fringilla coelebs, Figure S3), synonymous polymorphism level was very dependent on the number missing data tolerated in P. trochilus alignments (Figure S3). As such, we excluded P. trochilus from further analysis.

The mean Pn/Ps, calculated from the concatenated sequences of genes from the same gene class (control genes; BD; TLR; MHC I; MHC II; Database-group; Sma3s-group), was estimated for each bird species. Alternative transcripts were identified based on the genomic position in the GFF file. If several transcripts were available, one transcript was randomly selected. Pn/Ps estimates based on less than four polymorphic sites were excluded from the analysis, as were those with no polymorphic non-synonymous sites.

Statistical analyses

To estimate the impact of demographic history on genome-wide polymorphism of island species and the potentially reduced constraints on their immunity genes, we computed the ratio of non-synonymous genetic diversity over synonymous genetic diversity (Pn/Ps). We distinguished the part due to the island or mainland origin of species from the one due to the gene category (i.e., immunity versus control genes). A linear mixed model was performed, using the Pn/Ps ratio as dependant variable and, as explanatory variables, the mainland or

insular origin of species as well as the category of genes (packages lme4 and lmerTest (Bates et al. 2012; Kuznetsova et al. 2017)). In order to take into account the phylogenetic effect, the taxonomic rank "family" was included as a random effect in the model. Five linear mixed models were defined i) null model, ii) model with only the origin parameter, iii) model with only the gene category parameter, iv) model using both origin and gene category parameters, and finally v) model including those two parameters and the interaction effect. In some cases, the phylogenetic effect was difficult to estimate because the number of species per family was reduced to one. In that case, we chose to reduce the number of families by grouping Turdidae with Muscicapidae, Nectariniidae, and Estrildidae with Ploceidae and Fringillidae within Thraupidae. The results obtained with these family groupings were similar to the original model (Table S1), except when stated. The categories Database-group and Sma3s-group were tested separately from the Core group because they contained hundreds of genes annotated using the automatic pipeline that were only available for species with genome wide data. Database-group and Sma3s-group were not analysed simultaneously because they contained a partially overlapping set of genes. Finally, genes evolving under purifying selection and genes evolving under balancing selection were also analysed separately. Model selection was based on two methods. First, we use the difference in corrected Akaike Information Criterion (ΔAICc) calculated using the qpcR package (Spiess and Spiess 2018). Second, a model simplification using an ANOVA between models was also performed.

We also tested an alternative model using the difference between Pn/Ps of immunity genes and control genes (⊿Pn/Ps) as dependent variable, and species origin as explanatory variable. Under the hypothesis of a relaxation in selection pressure on islands due to a change in the parasite community, we expect the ⊿Pn/Ps to be higher on island species compared to the mainland ones and, therefore, the species origin (i.e., mainland or island) to be significant. In this model, we used the Phylogenetic Generalized Least Squares model (PGLS; implemented in the "nlme" packages; Pinheiro et al. 2017). This model assumes that the covariance between species follows a Brownian motion evolution process along the phylogeny (implemented using the "corBrownian" function from the ''ape'' package; Paradis and Schliep 2019). The species phylogeny was estimated using mitochondrial genes and a maximum likelihood inference implemented in IQTREE (model GTR+Gamma and ultrafast bootstrap; Nguyen et al. 2014; median of 11,134 bp analysed per species). The phylogeny with the bootstrap support is provided as supplementary material.

All the statistical analyses were performed using R (R core team 2018), and dplyr package (Wickham 2016). Graphical representations were done using ggplot2, ggrepel, ggpubr and ggpmisc (Kassambara 2018; Slowikowski et al. 2018; Wickham et al. 2019; Aphalo 2020).

Results

For the ten species (N = 150) for which we generated new data by targeted capture sequencing, an average of 3.3 millions paired-ends reads per individual was generated (Table S1). After mapping, genotyping and cleaning, we analysed 112.5 control and 16.4 immunity genes on average per species, out of the 141 targeted genes (120 control and 21 immunity related genes; Table S3). For the species with whole-genome sequences, we analysed 133 control and 20 immunity genes on average per species, out of the 141 targeted genes, and 904 and 785 genes on average in the Database-group and Sma3s-group respectively (Table S4).

Immunity genes evolving under purifying selection

We first focused on a restricted set of genes unambiguously involved in immunity function, namely the BD and TLR genes. At control genes, insular species had, on average, higher Pn/Ps ratios than the mainland ones (0.12 and 0.2 respectively).

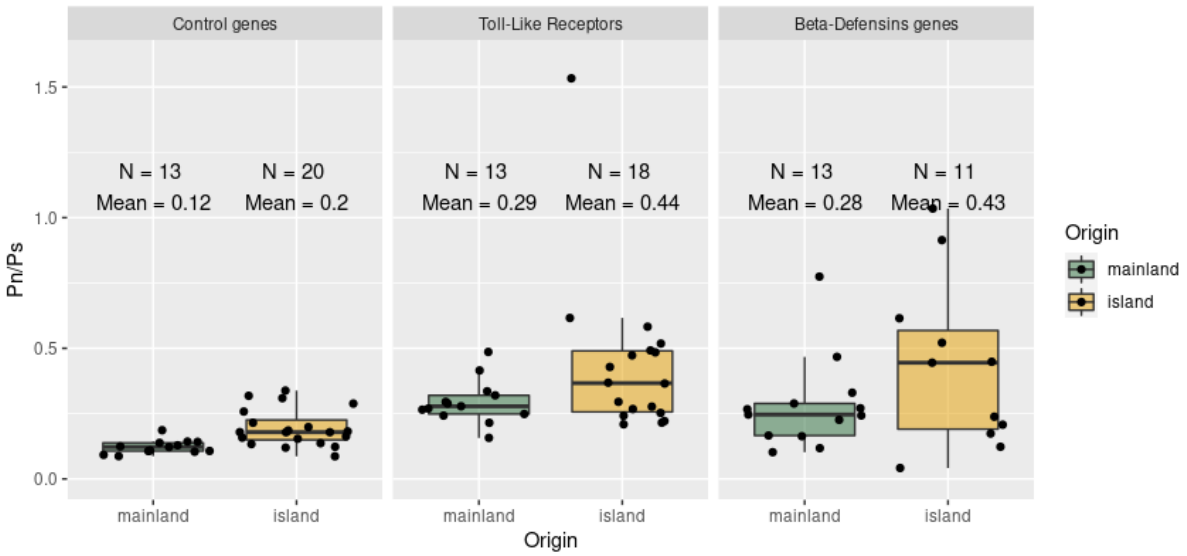


Figure 3: Boxplot of Pn/Ps according to species origin (mainland in green and insular in orange) for different gene categories under purifying selection. The number of species (N), and the mean Pn/Ps are shown for each modality.

Model selection based on AICc identified two models as similarly performant at explaining variation in Pn/Ps across species (ΔAICc < 2; Table 2). The first one only includes the gene category. The second one includes the origin (i.e., mainland or island) and gene category without interaction (Table 2). A model selection approach based on simplification with ANOVA identified the latter as the best (Table 2, $p < 0.05$). In this model, island origin of species is

associated with a greater Pn/Ps (0.12 vs. 0.09; Table 3; $p < 0.01$). Gene categories corresponding to TLRs and BDs showed a significantly higher Pn/Ps than control genes (Table 3; $p < 0.001$). In all cases, the best models have no interaction between origin and gene categories invalidating the hypothesis of a reduced parasite communities on island (Figure 2).

Table 2: Statistical model explaining Pn/Ps variation of Toll-Like Receptors, Beta-Defensins genes, and control genes. The p-values of ANOVA test between simpler models are not reported if a more complex model is significant.

Model	Model selection by AIC			ANOVA test			
	AICc	ΔAICc	Likelihood	Model 1	2	3	4
Pn/Ps~ 1+ category +origin+ category *origin	-1.93	8.95	0.01		0.69		
Pn/Ps~ 1+ category +origin	-10.87	0.00	1.00			0.05	1.73E-04
Pn/Ps~1+ category	-10.82	0.05	0.97				
Pn/Ps~1+ origin	-2.29	8.58	0.01				
Pn/Ps~1	-3.79	7.09	0.03				

Table 3: Summary of the best statistical model selected using AICc explaining variation in Pn/Ps in control genes, Toll-Like receptors and Beta-Defensins genes under purifying selection with origin, gene category parameters. * indicates significances : * < 0.05; ** < 0.01; *** < 0.001.

Model	*Parameters*			*Estimate*	*P.value*	
	Origin		*Category*			
Origin and Gene category	*Intercept*	*mainland*	*Control genes*	*0.09*	*2.91E-02*	*
		island		*0.12*	*5.97E-03*	**
			Toll-Like Receptors	*0.21*	*3.77E-05*	***
			Beta-Defensins genes	*0.20*	*2.44E-04*	***
Gene category	*Intercept*		*Control genes*	*0.16*	*4.34E-04*	***
			Toll-Like Receptors	*0.21*	*5.95E-05*	***
			Beta-Defensins genes	*0.19*	*4.53E-04*	***

For larger sets of genes, identified using an automatic pipeline and gene annotation, the model including only origin was identified as the best model explaining Pn/Ps (model selection based on AICc and simplification with ANOVA; Table S5, S7). Island was associated with a higher Pn/Ps of 0.05 ($p < 0.001$; Table S6, S8, Figure 4). For genes of the Sma3s-group, the category

parameter was also identified by simplification with ANOVA, associated with a reduction of the Pn/Ps of about 0.02 compared to control genes ($p < 0.05$; Table S9).

The alternative statistical approach using the difference between Pn/Ps of immunity genes and control genes (⊿Pn/Ps) as dependent variable, and species origin as explanatory variable under a PGLS framework lead to similar results. Island was never associated to a statistically higher ⊿Pn/Ps (table S2) providing no support for an increased relaxed selection of immune genes in island species.

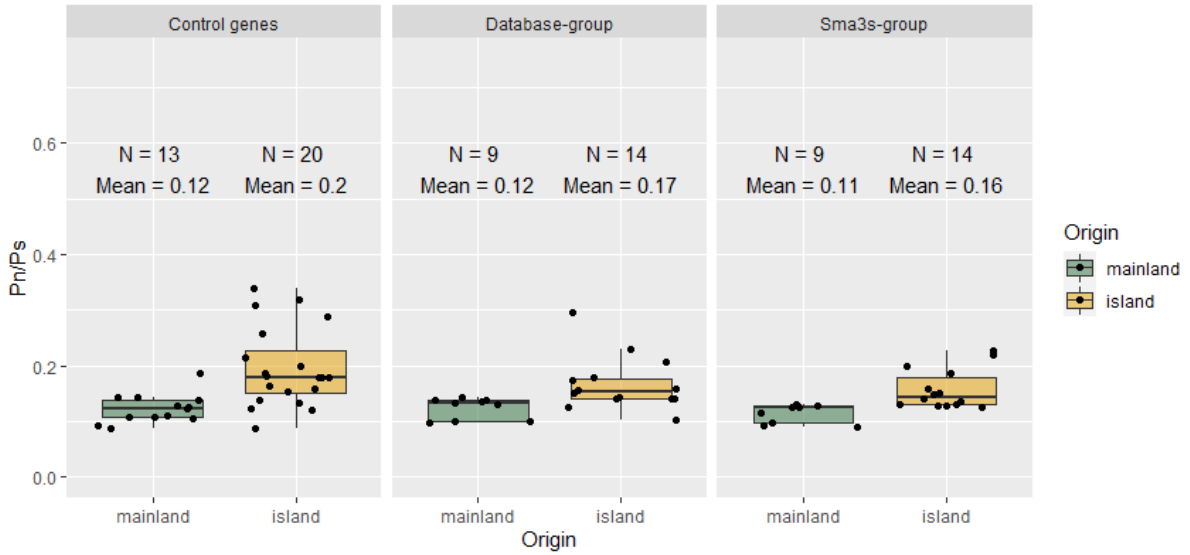
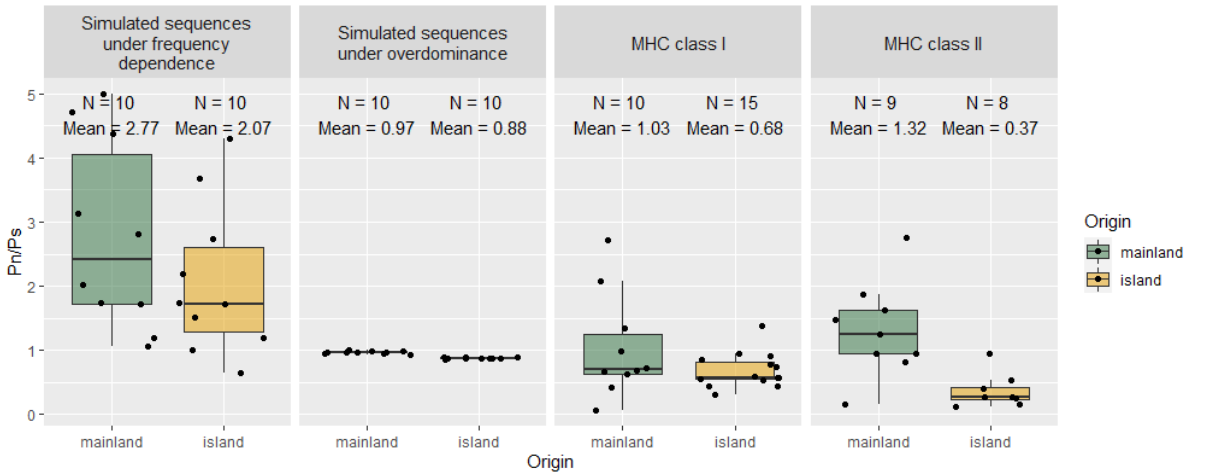


Figure 4: Boxplot of Pn/Ps according to species origin (mainland in green and insular in orange) for different gene categories under purifying selection. The number of individuals (N), and the mean Pn/Ps are shown for each modality.

Genes under balancing selection

First, we estimated the effect of population size variation on the Pn/Ps of the genes evolving under balancing selection by simulating sequences under frequency dependent or overdominance selection using SLiM (see Methods and Supplementary Methods). The simulation under frequency dependence selection revealed an average Pn/Ps equal to 2.07 for island species and 2.77 for mainland species (Figure 5). Under overdominance, simulated sequences from island and mainland populations respectively have an average Pn/Ps equal to 0.88 and 0.97, but the variance between simulated species was very small (Figure 5).



394

395 Figure 5: Boxplot of Pn/Ps according to species origin (mainland in green and insular in
396 orange) for different gene categories under balancing selection. The number of species (N),
397 and the mean Pn/Ps are shown for each modality. The control groups correspond to the results
398 obtained from simulated sequence via SLiM (see Methods and Supplementary Methods
399 Simulation of control genes under balancing selection).

400 Using simulations under frequency dependence selection, model selection identifies the two
401 models as equivalent, first the model with origin and category parameters and the full model
402 (Table 4). However, the full model is not significantly different from the model with origin and
403 category using the method by simplification with ANOVA (Table 4).

404 Using simulations under the overdominance, model selection identifies the model with origin
405 as the best, contrary to the method by simplification with ANOVA which identified the full model
406 therefore including significant interaction between origin and genes category (Table 4). This
407 interaction effect is significant for the MHC II ($p < 0.01$, Table S12) but not for MHC I. As
408 expected, island species have a significantly lower Pn/Ps in MHC genes compared to
409 mainland species ($p < 0.5$; except for the full model based on control genes evolving under
410 overdominance Table S12).

411 Table 4: Statistical model explaining Pn/Ps variation of genes under balancing selection (i.e
412 MHC class I and II), and simulated sequences under i) frequency dependence or ii)
413 overdominance. The p.values of ANOVA test between simpler models are not reported if a
414 more complex model is significant.

	Model	Model selection by AIC			ANOVA test			
		AICc	ΔAICc	Likelihood	1	2	3	4
Frequency dependence	Pn/Ps~1+ category +origin+ category *origin	164.96	1.23	0.54		0.26		

	Pn/Ps~1+ category +origin	163.73	0.00	1.00	1.09E-02	4.28E-03
	Pn/Ps~1+ category	166.72	3.00	0.22		
	Pn/Ps~1+ origin	168.75	5.02	0.08		
	Pn/Ps~1	171.37	7.64	0.02		
Overdominance	Pn/Ps~1+ category +origin+ category *origin	100.90	5.48	0.06	0.01	
	Pn/Ps~1+ category +origin	103.37	7.96	0.02		
	Pn/Ps~1+ category	107.30	11.89	0.00		
	Pn/Ps~1+ origin	95.42	0.00	1.00		
	Pn/Ps~1	99.68	4.27	0.12		

Discussion

On oceanic islands, the depauperate parasite community is expected to lead to a relaxation of selection on the immune system. In this study, we found support for such an effect, but only on MHC class II genes and under a specific simulation model (i.e, overdominance), which evolves under balancing selection. No effect was detected for MHC class I genes nor for innate immunity genes (TLRs and BDs), evolving under purifying selection. On these gene sets, increased drift effects on island populations limit the efficacy of selection in accordance with the nearly-neutral theory (Ohta 1992). The ability to distinguish between the selective and nearly-neutral processes (relaxed selection due to environmental change vs. drift) could only be achieved by our approach of using random genes (i.e., “control genes") to estimate the genome-wide effect of potential variation in effective population size between populations.

Effects of effective population size variation

Our results support the nearly-neutral theory of evolution for those genes under purifying selection, whereby strong genetic drift acting on small island populations reduces the efficacy of natural selection, leading to an increase in non-synonymous nucleotide diversity compared to the mostly neutral, synonymous nucleotide diversity (i.e., Pn/Ps) (Ohta 1992). This is materialized by a genome-wide increase in frequency of weakly deleterious mutations (Loire et al. 2013; Robinson et al. 2016; Rogers and Slatkin 2017; Kutschera et al. 2020; Leroy et al. 2021).

For genes evolving under balancing selection, we performed simulations under the hypotheses of overdominance (heterozygote advantage) or frequency dependence (rare-allele advantage). Our results showed reduced Pn/Ps for smaller population size (Figure 5, S4, S5). This simulation confirmed our expectations (fig. 5) that a reduction in the efficacy of selection results in a decrease in the frequency of non-synonymous polymorphism, as, under normal circumstances, selection maintains those mutations at intermediate frequencies. It also matches what we obtained in the empirical results, where both MHC classes I and II had a reduced Pn/Ps in island birds. This result supports that the fitness effect of having non-synonymous polymorphisms segregating at high frequencies is not strong enough to counteract entirely the effect of genetic drift on islands, therefore extending the nearly-neutral theory to the overdominance type of selection.

Effects of selection on immunity genes

For immune genes under purifying selection, no effect was detected on insular species, beyond what could be attributed to genetic drift. This is in line with the result of Gonzalez-Quevedo et al. (2015b) and Grueber et al. (2013) who found that TLR genetic diversity was mostly influenced by genetic drift. At first sight, this result seems not in line with the fact that island parasite communities are less diverse (Beadell et al. 2006; Maria et al. 2009; Pérez-Rodríguez et al. 2013; Loiseau et al. 2017 but see Illera et al. 2015). However, a reduced pathogens number has also been found to be associated with a higher prevalence in birds and reptiles from the Macaronesian archipelago (Illera Cobo and Perera 2020). Therefore, these two patterns, i.e. a less diverse pathogen's community on islands with a higher prevalence, could still imply a strong selection pressure on immune genes.

In contrast, for immune genes evolving under balancing selection, MHC class II genes presented a reduction in non-synonymous polymorphism larger than the effects of drift alone, when simulated sequences under overdominance are used as control. This was the only case where a role for relaxed selection pressures in the molecular evolution of immune genes could be invoked.

Our results are in accordance with the hypothesis of Lee (2006), which proposes that innate and acquired immunity may exhibit distinct responses to changes in pressures due to different costs and benefits. However, our result contrasts with the study of Santonastaso et al. (2017) that identified no change in selection pressures on MHC II genes in a lizard species, concluding that their evolution was mostly governed by drift. Similarly, Agudo et al. (2011)

found also a prominent role for genetic drift over selection in the evolution of MHC II genes in the Egyption vulture (*Neophron percnopterus*).

Our results rely on simulations that may be affected by the choice of the parameter values. First, we performed simulations using a fixed effective population size (Ne) estimated from the polymorphism data. Using others values of Ne had a weak impact on the relative difference between island and mainland species for the overdominance type of selection, but had a more noticeable impact for the frequency dependent type of selection (Figure S4, S5). Secondly, we simulated two types of selection, namely overdominance (Doherty and Zinkernagel 1975) and frequency dependence (Slade and McCallum 1992), but it has been argued that the maintenance of MHC polymorphism could be the result of fluctuating selection (Hill 1991). Additionally, recombination and gene conversion has also been put forward as a mechanism responsible for generating diversity (Spurgin et al. 2011). Therefore, our results for the MHC II, which is based on the relative difference between Pn/Ps of island and mainland species comparing empirical and simulated data, should be taken cautiously as their significance can be dependent on the specific parameters that we used, although we did our best to select a realistic range of parameters.

The observed difference between MHC class I and II could be explained by their different pathogen targets: MHC class I genes are primarily involved in the recognition intracellular pathogens (Kappes and Strominger 1988), while MHC class II genes are directly involved in the recognition of extracellular pathogens (Bjorkman and Parham 1990). These differences could lead to variable selection pressures depending on the extracellular versus intracellular parasite communities present on islands. In addition, the relaxed selection pressures on MHC II genes from adaptive immunity is in line with a reduction in acquired immunity parameters observed by Lobato et al. (2017) that used partly the same sets of species.

As a perspective of our work, we should mention that there is an extensive variation in the number of MHC gene copies across the avian phylogeny (Minias et al. 2019; O'Connor et al. 2020). Particularly, it was recently discovered that Passerines have a very dynamic evolution of duplication/loss events compared to other birds (Minias et al. 2019). Here, we used the two copies of MHC gene I and II currently annotated in the collared flycatcher genome as target sequences for our targeted-capture sequencing. The recent improvement of genome assembly, thanks to the development of long-reads technology (Peona et al. 2018; Peona et al. 2021), will certainly help to precisely annotate all MHC copies and to study the whole repertoire of MHC genes.

Consequences of drift effect and selection on immunity

The potential relaxation of the natural selection acting on immune genes in island species is expected to reduce immune functions and increase susceptibility of island populations to pathogens. This is true even if this relaxation is only the consequence of a reduction in the effective population size and not caused by a reduction of the pressure exerted by the parasitic community. This is in line with the results of Hawley et al. (2005) and Belasen et al. (2019) who showed that a decrease in diversity of immune loci (MHC II or through immune proxy) was associated with a reduction in immune functions. It should be noted that even if migration rate is reduced on islands, sedentary and endemic island species are not completely free from the exposure of exogen pathogens through migratory birds (Levin et al. 2013).

As a final remark, we would like to stress that more research is still needed (i) to ascertain both selection pressures on innate and adaptive immune responses and the load of deleterious mutations due to drift, also identified by an increasing body of work (Loire et al. 2013; Robinson et al. 2016; Rogers and Slatkin 2017; Kutschera et al. 2020; Leroy et al. 2021), and (ii) to describe island parasite communities. To date, most of the studies investigated intracellular parasite communities on islands, and more specifically haemosporidian parasites, avian pox and coccidian parasites (Illera et al. 2008; Ishtiaq et al. 2010; Cornuault et al. 2012; Silva-Iturriza et al. 2012; Pérez-Rodríguez et al. 2013; Illera et al. 2015; Martinez et al. 2015; Loiseau et al. 2017; Padilla et al. 2017), whereas very few evaluated the extracellular parasite diversity, such as helminths (Nieberding et al. 2006) but see the review of Illera Cobo and Perera (2020) for reptiles. Metabarcoding of parasites is a new technique to evaluate at the same time both communities of intracellular and extracellular parasites (Bourret et al. 2021) and might be therefore a promising approach to evaluate their communities in island and mainland populations.

Conclusion

Our comparative population genomics study has investigated the combined effects of drift and selection on immunity genes from island and mainland passerines. The study of synonymous and non-synonymous polymorphism of these genes confirmed that island species, with smaller population sizes than their mainland counterparts, were more impacted by drift, which induces a load of weakly deleterious mutations in their genome. Indeed most of the genes studied here involved in the immune response do not show a statistically different pattern from control genes. Only MHC II genes, involved in the recognition of extracellular pathogens, showed a reduction in their non-synonymous polymorphism in island species. This response, which may be attributed to reduced selection pressures on these genes, could be associated with the suspected reduced parasitic communities on islands. The increased load of deleterious mutations as well as the potential relaxed selection pressures on MHC II support

538 the reduced immune functions of island species, which could be added to the list of other
539 convergent responses of the island syndrome.

540 *Data availability*

541 Datasets, scripts, supplementary figures and texts are available on figshare :
542 https://figshare.com/s/ab7004cc2f4415b4058f. The reads newly generated for this study have
543 been deposited in the NCBI Sequence Read Archive under the bioproject PRJNA724656.

544 *Acknowledgments*

545 In Gabon, we thank the Director and the guides of the Lekedi Park, Marie Charpentier for her
546 help in organizing the expedition, and Alexandre Vaz for field assistance and outreach work.
547 In São Tomé and Príncipe, we thank the Directorate of the Environment and the Department
548 for Nature Conservation, its directors—Arlindo Carvalho and Victor Bonfim—Guilhermino, the
549 Association Monte Pico, its president Luis Mário, and its members. Philippe Perret, Octávio
550 Veiga, Bikegila, and Yelli provided invaluable assistance in the field. Permissions for fieldwork
551 were given by the authorities of São Tomé and Príncipe and Gabon (CENAREST authorization
552 No. AR0053/12/MENESTFPRSCJS/ CENAREST/CG/CST/CSAR). Permits for the Canary
553 Islands were provided by the Regional Government (Ref.: 2012/0710), and the Cabildo of La
554 Palma and Tenerife. In Montpellier, we thank the blue tit team
555 (https://oreme.org/observation/ecopop/mesanges/) for the capture of the individuals used in
556 this study. The analyses benefited from the Montpellier Bioinformatics Biodiversity (MBB)
557 platform services. This research was conducted in the scope of the international twin-lab “LIA
558 – Biodiversity and Evolution” between CIBIO (Portugal) and ISEM and CEFE-CNRS (France).

559 *Funding information*

560 This research was funded by the Labex CeMEB (project ISLAND IMMUNITY) for BN, the ANR
561 (BirdIslandGenomic project, ANR-14-CE02-0002) for MB, BN, CL and CD, the National
562 Geographic Society (Grant/Award Number:W251-12), the British Ecological Society
563 (Grant/Award Number: 369/4558), the portuguese Foundation for Science and Technology
564 under the PTDC/BIA-EVL/29390/2017 “DEEP” Research Project for MM, RC and CL and the
565 Spanish Ministry of Science, Innovation and Universities, and the European Regional
566 Development Fund (Ref.: PGC2018-097575-B-I00) for JCI.

567

568 **References**

569 Agudo R, Alcaide M, Rico C, Lemus JA, Blanco G, Hiraldo F, Donázar JA. 2011. Major

histocompatibility complex variation in insular populations of the Egyptian vulture: inferences about the roles of genetic drift and selection. *Mol. Ecol.* 20:2329–2340.

Akira S. 2003. Toll-like receptor signaling. *J. Biol. Chem.* 278:38105–38108.

Alberts B, Johnson A, Lewis J, Raff M, Roberts K, Walter P. 2002. Innate immunity. In: Molecular Biology of the Cell. 4th edition. Garland Science.

Andrés AM, Hubisz MJ, Indap A, Torgerson DG, Degenhardt JD, Boyko AR, Gutenkunst RN, White TJ, Green ED, Bustamante CD. 2009. Targets of balancing selection in the human genome. *Mol. Biol. Evol.* 26:2755–2764.

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

Aphalo PJ. 2020. ggpmisc: Miscellaneous Extensions to "ggplot2"(R package version 0.3. 6).

Baeckens S, Van Damme R. 2020. The island syndrome. *Curr. Biol.* 30:R338–R339.

Bates D, Maechler M, Bolker B, Walker S, Christensen RHB, Singmann H, Dai B, Scheipl F, Grothendieck G. 2012. Package 'lme4.' *CRAN R Found. Stat. Comput. Vienna Austria*.

Beadell JS, Atkins C, Cashion E, Jonker M, Fleischer RC. 2007. Immunological change in a parasite-impoverished environment: divergent signals from four island taxa. *PLoS One* 2:e896.

Beadell JS, Ishtiaq F, Covas R, Melo M, Warren BH, Atkinson CT, Bensch S, Graves GR, Jhala YV, Peirce MA. 2006. Global phylogeographic limits of Hawaii's avian malaria. *Proc. R. Soc. Lond. B Biol. Sci.* 273:2935–2944.

Belasen AM, Bletz MC, Leite D da S, Toledo LF, James TY. 2019. Long-term habitat fragmentation is associated with reduced MHC IIB diversity and increased infections in amphibian hosts. *Front. Ecol. Evol.* 6:236.

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

Bertrand J a. M, Bourgeois YXC, Delahaie B, Duval T, García-Jiménez R, Cornuault J, Heeb P, Milá B, Pujol B, Thébaud C. 2014. Extremely reduced dispersal and gene flow in an island bird. *Heredity* 112:190–196.

Bjorkman PJ, Parham P. 1990. Structure, function, and diversity of class I major histocompatibility complex molecules. *Annu. Rev. Biochem.* 59:253–288.

Blondel J. 2000. Evolution and ecology of birds on islands: trends and prospects. *Vie MilieuLife Environ.*:205–220.

Bourret V, Gutiérrez López R, Melo M, Loiseau C. 2021. Metabarcoding options to study eukaryotic endoparasites of birds. *Ecol. Evol.* 11:10821–10833.

Boyce MS. 1984. Restitution of gamma-and k-selection as a model of density-dependent natural selection. *Annu. Rev. Ecol. Syst.* 15:427–447.

Breuer K, Foroushani AK, Laird MR, Chen C, Sribnaia A, Lo R, Winsor GL, Hancock RE, Brinkman FS, Lynn DJ. 2013. InnateDB: systems biology of innate immunity and beyond—recent updates and continuing curation. *Nucleic Acids Res.* 41:D1228–D1233.

Buffalo V. 2021. Quantifying the relationship between genetic diversity and population size suggests natural selection cannot explain Lewontin's paradox.Sella G, editor. *eLife* 10:e67509.

Chapman JR, Hellgren O, Helin AS, Kraus RH, Cromie RL, Waldenström J. 2016. The evolution of innate immune genes: purifying and balancing selection on β-defensins in waterfowl. *Mol. Biol. Evol.* 33:3075–3087.

Charlesworth J, Eyre-Walker A. 2008. The McDonald–Kreitman test and slightly deleterious mutations. *Mol. Biol. Evol.* 25:1007–1015.

Chen S, Zhou Y, Chen Y, Gu J. 2018. fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* 34:i884–i890.

Corcoran P, Gossmann TI, Barton HJ, Slate J, Zeng K. 2017. Determinants of the Efficacy of Natural Selection on Coding and Noncoding Variability in Two Passerine Species. *Genome Biol. Evol.* 9:2987–3007.

Cornuault J, Bataillard A, Warren BH, Lootvoet A, Mirleau P, Duval T, Milá B, Thébaud C,

Heeb P. 2012. The role of immigration and in-situ radiation in explaining blood parasite assemblages in an island bird clade. *Mol. Ecol.* 21:1438–1452.

Covas R. 2012. Evolution of reproductive life histories in island birds worldwide. *Proc. R. Soc. B Biol. Sci.* 279:1531–1537.

van Dijk A, Veldhuizen EJ, Haagsman HP. 2008. Avian defensins. *Vet. Immunol. Immunopathol.* 124:1–18.

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

Doutrelant C, Paquet M, Renoult JP, Grégoire A, Crochet P-A, Covas R. 2016. Worldwide patterns of bird colouration on islands. *Ecol. Lett.* 19:537–545.

Ellegren H, Smeds L, Burri R, Olason PI, Backström N, Kawakami T, Künstner A, Mäkinen H, Nadachowska-Brzyska K, Qvarnström A. 2012. The genomic landscape of species divergence in Ficedula flycatchers. *Nature* 491:756–760.

Fijarczyk A, Babik W. 2015. Detecting balancing selection in genomes: limits and prospects. *Mol. Ecol.* 24:3529–3545.

Fijarczyk A, Dudek K, Babik W. 2016. Selective Landscapes in newt Immune Genes Inferred from Patterns of Nucleotide Variation. *Genome Biol. Evol.* 8:3417–3432.

Frankham R. 1997. Do island populations have less genetic variation than mainland populations? *Heredity* 78:311–327.

Garamszegi LZ. 2006. The evolution of virulence and host specialization in malaria parasites of primates. *Ecol. Lett.* 9:933–940.

Garrison E, Marth G. 2012. Haplotype-based variant detection from short-read sequencing. *ArXiv Prepr. ArXiv12073907*.

Gonzalez-Quevedo C, Phillips KP, Spurgin LG, Richardson DS. 2015a. 454 screening of individual MHC variation in an endemic island passerine. *Immunogenetics* 67:149–162.

Gonzalez-Quevedo C, Spurgin LG, Illera JC, Richardson DS. 2015b. Drift, not selection, shapes toll-like receptor variation among oceanic island populations. *Mol. Ecol.* 24:5852–5863.

Grant PR. 1965. The adaptive significance of some size trends in island birds. *Evolution*:355–367.

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. *Mol. Ecol.* 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.

Guéguen L, Gaillard S, Boussau B, Gouy M, Groussin M, Rochette NC, Bigot T, Fournier D, Pouyet F, Cahais V, et al. 2013. Bio++: Efficient Extensible Libraries and Tools for Computational Molecular Evolution. *Mol. Biol. Evol.* 30:1745–1750.

Haller BC, Messer PW. 2017. SLiM 2: Flexible, interactive forward genetic simulations. *Mol. Biol. Evol.* 34:230–240.

Hawley DM, Sydenstricker KV, Kollias GV, Dhondt AA. 2005. Genetic diversity predicts pathogen resistance and cell-mediated immunocompetence in house finches. *Biol. Lett.* 1:326–329.

Hill AV. 1991. HLA associations with malaria in Africa: some implications for MHC evolution. In: Molecular evolution of the major histocompatibility complex. Springer. p. 403–420.

Hochberg ME, Møller AP. 2001. Insularity and adaptation in coupled victim–enemy associations. *J. Evol. Biol.* 14:539–551.

Hughes AL, Nei M. 1988. Pattern of nucleotide substitution at major histocompatibility complex class I loci reveals overdominant selection. *Nature* 335:167–170.

Illera Cobo JC, Perera A. 2020. Where are we in the host-parasite relationships of native land vertebrates in Macaronesia? *Ecosistemas*.

Illera JC, Emerson BC, Richardson DS. 2008. Genetic characterization, distribution and prevalence of avian pox and avian malaria in the Berthelot's pipit (Anthus berthelotii) in Macaronesia. *Parasitol. Res.* 103:1435–1443.

Illera JC, Fernández-Álvarez Á, Hernández-Flores CN, Foronda P. 2015. Unforeseen biogeographical patterns in a multiple parasite system in Macaronesia. *J. Biogeogr.* 42:1858–1870.

Institute B. 2019. “Picard Toolkit”, Broad institute, GitHub repository. *Picard Toolkit*.

Ishtiaq F, Clegg SM, Phillimore AB, Black RA, Owens IP, Sheldon BC. 2010. Biogeographical patterns of blood parasite lineage diversity in avian hosts from southern Melanesian islands. *J. Biogeogr.* 37:120–132.

Kappes D, Strominger JL. 1988. Human class II major histocompatibility complex genes and proteins. *Annu. Rev. Biochem.* 57:991–1028.

Kassambara A. 2018. ggpubr:“ggplot2” based publication ready plots. *R Package Version 01* 7.

Kimura M. 1962. On the Probability of Fixation of Mutant Genes in a Population. *Genetics* 47:713–719.

Klein J. 1986. Natural history of the major histocompatibility complex. Wiley

Kutschera VE, Poelstra JW, Botero-Castro F, Dussex N, Gemmell NJ, Hunt GR, Ritchie MG, Rutz C, Wiberg RAW, Wolf JB. 2020. Purifying selection in corvids is less efficient on islands. *Mol. Biol. Evol.* 37:469–474.

Kuznetsova A, Brockhoff PB, Christensen RH. 2017. lmerTest package: tests in linear mixed effects models. *J. Stat. Softw.* 82:1–26.

Lamichhaney S, Berglund J, Almén MS, Maqbool K, Grabherr M, Martinez-Barrio A, Promerová M, Rubin C-J, Wang C, Zamani N, et al. 2015. Evolution of Darwin/’s finches and their beaks revealed by genome sequencing. *Nature* 518:371–375.

Lee KA. 2006. Linking immune defenses and life history at the levels of the individual and the species. *Integr. Comp. Biol.* 46:1000–1015.

Leroy T, Anselmetti Y, Tilak M-K, Berard S, Csukonyi L, Gabrielli M, Scornavacca C, Mila B, Thebaud C, Nabholz B. 2019. A bird’s white-eye view on avian sex chromosome evolution. *BioRxiv 505610 Ver 4 Peer-Rev. Recomm. Peer Community Evol. Biol. 101101505610* [Internet]. Available from: https://www.biorxiv.org/content/early/2019/01/23/505610

Leroy T, Rousselle M, Tilak M-K, Caizergues AE, Scornavacca C, Recuerda M, Fuchs J, Illera JC, De Swardt DH, Blanco G. 2021. Island songbirds as windows into evolution in small populations. *Curr. Biol.* 31:1303–1310.

Li H. 2013. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *ArXiv Prepr. ArXiv13033997*.

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R. 2009. The Sequence Alignment/Map format and SAMtools. *Bioinforma. Oxf. Engl.* 25:2078–2079.

Lindström KM, Foufopoulos J, Pärn H, Wikelski M. 2004. Immunological investments reflect parasite abundance in island populations of Darwin’s finches. *Proc. R. Soc. Lond. B Biol. Sci.* 271:1513–1519.

Lobato E, Doutrelant C, Melo M, Reis S, Covas R. 2017. Insularity effects on bird immune parameters: A comparison between island and mainland populations in West Africa. *Ecol. Evol.* 7:3645–3656.

Loire E, Chiari Y, Bernard A, Cahais V, Romiguier J, Nabholz B, Lourenço JM, Galtier N. 2013. Population genomics of the endangered giant Galapagos tortoise. *Genome Biol.* 14:R136.

Loiseau C, Melo M, Lobato E, Beadell JS, Fleischer RC, Reis S, Doutrelant C, Covas R. 2017. Insularity effects on the assemblage of the blood parasite community of the birds from the Gulf of Guinea. *J. Biogeogr.* 44:2607–2617.

Lomolino MV. 2005. Body size evolution in insular vertebrates: generality of the island rule. *J. Biogeogr.* 32:1683–1699.

Losos JB, Ricklefs RE. 2009. Adaptation and diversification on islands. *Nature* 457:830–836.

Lundberg M, Liedvogel M, Larson K, Sigeman H, Grahn M, Wright A, Åkesson S, Bensch S. 2017. Genetic differences between willow warbler migratory phenotypes are few and cluster in large haplotype blocks. *Evol. Lett.* 1:155–168.

Luo R, Liu B, Xie Y, Li Z, Huang W, Yuan J, He G, Chen Y, Pan Q, Liu Y. 2012. SOAPdenovo2: an empirically improved memory-efficient short-read de novo assembler. *Gigascience* 1:18.

MacArthur RH, Wilson EO. 1967. The theory of island biogeography. *Princet. NJ*.

Maria L, Svensson E, Ricklefs RE. 2009. Low diversity and high intra-island variation in prevalence of avian Haemoproteus parasites on Barbados, Lesser Antilles. *Parasitology* 136:1121–1131.

Martinez J, Vasquez RA, Venegas C, Merino S. 2015. Molecular characterisation of haemoparasites in forest birds from Robinson Crusoe Island: is the Austral Thrush a potential threat to endemic birds? *Bird Conserv. Int.* 25:139–152.

Matson KD. 2006. Are there differences in immune function between continental and insular birds? *Proc. Biol. Sci.* 273:2267–2274.

Matson KD, Beadell JS. 2010. Infection, immunity, and island adaptation in birds. *Biogeogr. Host–parasite Interact.*:217.

Minias P, Pikus E, Whittingham LA, Dunn PO. 2019. Evolution of copy number at the MHC varies across the avian tree of life. *Genome Biol. Evol.* 11:17–28.

Munoz-Mérida A, Viguera E, Claros MG, Trelles O, Pérez-Pulido AJ. 2014. Sma3s: a three-step modular annotator for large sequence datasets. *DNA Res.* 21:341–353.

Nguyen L-T, Schmidt HA, von Haeseler A, Minh BQ. 2014. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol. Evol.* 32:268–274.

Nieberding C, Morand S, Libois R, Michaux JR. 2006. Parasites and the island syndrome: the colonization of the western Mediterranean islands by Heligmosomoides polygyrus (Dujardin, 1845). *J. Biogeogr.* 33:1212–1222.

O'Connor EA, Hasselquist D, Nilsson J-Å, Westerdahl H, Cornwallis CK. 2020. Wetter climates select for higher immune gene diversity in resident, but not migratory, songbirds. *Proc. R. Soc. B Biol. Sci.* 287:20192675.

Ohta T. 1992. The nearly neutral theory of molecular evolution. *Annu. Rev. Ecol. Syst.* 23:263–286.

Ortutay C, Vihinen M. 2009. Identification of candidate disease genes by integrating Gene Ontologies and protein-interaction networks: case study of primary immunodeficiencies. *Nucleic Acids Res.* 37:622–628.

Padilla DP, Illera JC, Gonzalez-Quevedo C, Villalba M, Richardson DS. 2017. Factors affecting the distribution of haemosporidian parasites within an oceanic island. *Int. J. Parasitol.* 47:225–235.

Paradis E, Schliep K. 2019. ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics* 35:526–528.

Peona V, Blom MPK, Xu L, Burri R, Sullivan S, Bunikis I, Liachko I, Haryoko T, Jønsson KA, Zhou Q, et al. 2021. Identifying the causes and consequences of assembly gaps using a multiplatform genome assembly of a bird-of-paradise. *Mol. Ecol. Resour.* 21:263–286.

Peona V, Weissensteiner MH, Suh A. 2018. How complete are "complete" genome assemblies?—An avian perspective. *Mol. Ecol. Resour.* 18:1188–1195.

Pérez-Rodríguez A, Ramírez Á, Richardson DS, Pérez-Tris J. 2013. Evolution of parasite island syndromes without long-term host population isolation: Parasite dynamics in Macaronesian blackcaps Sylvia atricapilla. *Glob. Ecol. Biogeogr.* 22:1272–1281.

Pinheiro J, Bates D, DebRoy S, Sarkar D, Heisterkamp S, Van Willigen B, Maintainer R. 2017. Package 'nlme.' *Linear Nonlinear Mix. Eff. Models Version* 3.

R core team. 2018. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing

Rando JC, Alcover JA, Illera JC. 2010. Disentangling Ancient Interactions: A New Extinct Passerine Provides Insights on Character Displacement among Extinct and Extant Island Finches. *PLOS ONE* 5:e12956.

Ranwez V, Harispe S, Delsuc F, Douzery EJ. 2011. MACSE: Multiple Alignment of Coding SEquences accounting for frameshifts and stop codons. *PLoS One* 6:e22594.

Robinson JA, Ortega-Del Vecchyo D, Fan Z, Kim BY, Marsden CD, Lohmueller KE, Wayne RK. 2016. Genomic flatlining in the endangered island fox. *Curr. Biol.* 26:1183–1189.

Rogers RL, Slatkin M. 2017. Excess of genomic defects in a woolly mammoth on Wrangel island. *PLoS Genet.* 13:e1006601.

Rohland N, Reich D. 2012. Cost-effective, high-throughput DNA sequencing libraries for multiplexed target capture. *Genome Res.* 22:939–946.

Santonastaso T, Lighten J, van Oosterhout C, Jones KL, Foufopoulos J, Anthony NM. 2017. The effects of historical fragmentation on major histocompatibility complex class II β and microsatellite variation in the Aegean island reptile, Podarcis erhardii. *Ecol. Evol.* 7:4568–4581.

Sawyer SA, Hartl DL. 1992. Population genetics of polymorphism and divergence. *Genetics* 132:1161–1176.

She R, Chu JS-C, Uyar B, Wang J, Wang K, Chen N. 2011. genBlastG: using BLAST searches to build homologous gene models. *Bioinforma. Oxf. Engl.* 27:2141–2143.

Silva-Iturriza A, Ketmaier V, Tiedemann R. 2012. Prevalence of avian haemosporidian parasites and their host fidelity in the central Philippine islands. *Parasitol. Int.* 61:650–657.

Singhal S, Leffler EM, Sannareddy K, Turner I, Venn O, Hooper DM, Strand AI, Li Q, Raney B, Balakrishnan CN. 2015. Stable recombination hotspots in birds. *Science* 350:928–932.

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

Slowikowski K, Schep A, Hughes S, Lukauskas S, Irisson J-O, Kamvar ZN, Ryan T, Christophe D, Hiroaki Y, Gramme P. 2018. Package ggrepel. *Autom. Position Non-Overlapping Text Labels 'ggplot2*.

Spiess A-N, Spiess MA-N. 2018. Package 'qpcR.' *Model. Anal. Real-Time PCRdata Httpscran R-Proj. OrgwebpackagesqpcRqpcR Pdf*.

Spurgin LG, Van Oosterhout C, Illera JC, Bridgett S, Gharbi K, Emerson BC, Richardson DS. 2011. Gene conversion rapidly generates major histocompatibility complex diversity in recently founded bird populations. *Mol. Ecol.* 20:5213–5225.

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

Tange O. 2018. GNU parallel 2018. Lulu. com

Van Riper III C, Van Riper SG, Goff ML, Laird M. 1986. The epizootiology and ecological significance of malaria in Hawaiian land birds. *Ecol. Monogr.* 56:327–344.

Velová H, Gutowska-Ding MW, Burt DW, Vinkler M. 2018. Toll-like receptor evolution in birds: gene duplication, pseudogenization, and diversifying selection. *Mol. Biol. Evol.* 35:2170–2184.

Warren BH, Simberloff D, Ricklefs RE, Aguilée R, Condamine FL, Gravel D, Morlon H, Mouquet N, Rosindell J, Casquet J. 2015. Islands as model systems in ecology and evolution: Prospects fifty years after MacArthur-Wilson. *Ecol. Lett.* 18:200–217.

Waters JM, Emerson BC, Arribas P, McCulloch GA. 2020. Dispersal Reduction: Causes, Genomic Mechanisms, and Evolutionary Consequences. *Trends Ecol. Evol.* 35:512–522.

Wickham H. 2016. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York Available from: http://ggplot2.org

Wickham H, François R, Henry L, Müller K. 2019. dplyr: A Grammar of Data Manipulation. R package version 0.8. 0.1. *Retrieved January* 13:2020.

Wikelski M, Foufopoulos J, Vargas H, Snell H. 2004. Galápagos birds and diseases: invasive pathogens as threats for island species. *Ecol. Soc.* 9.

Wolf JBW, Künstner A, Nam K, Jakobsson M, Ellegren H. 2009. Nonlinear Dynamics of Nonsynonymous (dN) and Synonymous (dS) Substitution Rates Affects Inference of Selection. *Genome Biol. Evol.* 1:308–319.